SIAM J. Applied Dynamical Systems
Vol. 6, No. 1, pp. 1–28

Bifurcation Analysis of a Semiconductor Laser with Filtered Optical Feedback*

Hartmut Erzgräber†, Bernd Krauskopf‡, and Daan Lenstra§

Abstract. We study the dynamics and bifurcations of a semiconductor laser with delayed filtered optical feedback, where a part of the output of the laser reenters after spectral filtering. This type of coherent optical feedback is more challenging than the case of conventional optical feedback from a simple mirror, but it provides additional control over the output of the semiconductor laser by means of choosing the filter detuning and the filter width. This laser system can be modeled by a system of delay differential equations with a single fixed delay, which is due to the travel time of the light outside the laser. In this paper we present a bifurcation analysis of the filtered feedback laser. We first consider the basic continuous wave states, known as the external filtered modes (EFMs), and determine their stability regions in the parameter plane of feedback strength versus feedback phase. The EFMs are born in saddle-node bifurcations and become unstable in Hopf bifurcations. We show that for small filter detuning there is a single region of stable EFMs, which splits up into two separate regions when the filter is detuned. We then concentrate on the periodic orbits that emanate from Hopf bifurcations. Depending on the feedback strength and the feedback phase, two types of oscillations can be found. First, there are undamped relaxation oscillations, which are typical for semiconductor laser systems. Second, there are oscillations with a period related to the delay time, which have the remarkable property that the laser frequency oscillates while the laser intensity is almost constant. These frequency oscillations are only possible due to the interaction of the laser with the filter. We determine the stability regions in the parameter plane of feedback strength versus feedback phase of the different types of oscillations. In particular, we find that stable frequency oscillations are dominant for nonzero values of the filter detuning.

Key words. semiconductor laser, delayed feedback, optical filtering, delay differential equation, bifurcation analysis

AMS subject classifications. 15A15, 15A09, 15A23

DOI. 10.1137/060656656

1. Introduction.

Semiconductor lasers are very efficient sources of coherent light that are used in countless technical applications—most notably in optical data storage and optical communication systems. A single-mode semiconductor laser on its own is mathematically a damped nonlinear oscillator, where an important nonlinearity is a strong coupling between the light intensity and the output frequency. Furthermore, due to the high efficiency of the semiconductor material, the reflectivity of the mirrors of semiconductor lasers is very low (only

*Received by the editors April 7, 2006; accepted for publication (in revised form) by T. Kaper November 15, 2006; published electronically February 5, 2007.
http://www.siam.org/journals/siads/6-1/65665.html

†Afdeling Natuurkunde en Sterrenkunde, Vrije Universiteit Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands (h.erzgraber@few.vu.nl).

‡Department of Engineering Mathematics, University of Bristol, Bristol BS8 1TR, UK (b.krauskopf@bristol.ac.uk).
The work of this author was supported by an Engineering and Physical Sciences Research Council (EPSRC) Advanced Research Fellowship grant.
§Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, P.O. box 5031, 2600 GA Delft, The Netherlands (d.lenstra@tudelft.nl).

about 30%). Therefore, this type of laser is very susceptible to external influences: external light can enter easily and disturb the delicate balance inside the laser; see, for example, [21, 23] as entry points to the extensive literature on semiconductor laser dynamics.

From a dynamical systems point of view, semiconductor laser systems are very attractive because they show an intriguing variety of complicated dynamics. Furthermore, it is known that their description, by relatively simple mathematical models, is in very good agreement with experimental measurements [7, 17, 34].

In applications it is impossible to completely isolate a laser from external light. Therefore, one has to find ways to minimize or control the effects of external optical perturbations. Furthermore, one often wants to control the output characteristics of the laser, for example, to ensure constant intensity output at a desired frequency [20]. On the other hand, there is the possibility of making use of chaotic dynamics of semiconductor lasers, for example, as a chaotic carrier wave for secure communication or to produce highly efficient incoherent light sources.

Different schemes have been proposed for controlling the dynamics of semiconductor lasers, including optical injection of light from a second laser [34] and different types of optical feedback [21]. The simplest optical feedback scheme is conventional optical feedback (COF), where the laser receives feedback from a normal mirror. However, other types of feedback are also possible, including optical feedback from two different mirrors [30], incoherent feedback [8], phase-conjugate feedback (PCF) [1, 11], and optoelectronic feedback [27].

In this paper we consider a semiconductor laser subject to filtered optical feedback (FOF), where the reflected light is spectrally filtered before it reenters the laser. This coherent optical feedback system, which is known as the FOF laser, has recently been the subject of a number of experimental and theoretical studies [5, 6, 9, 10, 12, 18, 19, 29, 33, 36, 37]. As in any optical feedback system, important parameters are the delay time and the feedback rate. Moreover, for coherent feedback, there is also a feedback phase that controls the phase of the incident light. The interest in the FOF laser is due to the fact that filtering of the reflected light allows additional control over the behavior of the laser by means of choosing the filter detuning (the difference between the central filter frequency and the free-running laser frequency) and the filter width.

A particular motivation for the bifurcation analysis performed here was the discovery by Fischer et al. [10] of a new type of oscillations. These, the so-called frequency oscillations (FOs), are characterized by oscillations of the optical frequency of the laser while its intensity remains practically constant. Mathematically, this means that the dynamics of the laser takes place in a very small neighborhood of a cylinder in phase space. The existence of FOs is remarkable for several reasons. First, pure FOs are "unusual" for semiconductor lasers due to the strong amplitude-phase coupling in these lasers. Second, the period of the FOs is on the order of the delay time of the FOF system, while one would normally expect the undamping of the characteristic relaxation oscillations (ROs) to be the first instability to be encountered in semiconductor lasers. Note that ROs are a well-known feature of laser dynamics. Specifically, they are a periodic exchange of energy between the optical field (the number of photons) and the population inversion (the number of electron-hole pairs) of the laser. They have a characteristic frequency that depends on the laser and its operating conditions and is on the order of GHz [35].

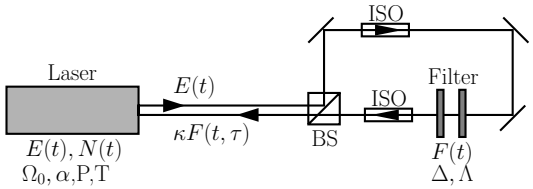


Figure 1. *FOF laser set-up with a Fabry–Pérot filter in a loop with optical isolators (ISO).*

We present here a detailed bifurcation study, where we identify the stability regions of the basic solutions of the FOF laser, the external filtered modes (EFMs), and the different types of bifurcating oscillations, ROs and FOs. Laser systems involving optical feedback, such as the FOF laser considered here, are quite challenging to analyze because they need to be modeled by delay differential equations (DDEs), which feature an infinite-dimensional phase space. In this work we use numerical continuation software for DDEs, namely the packages DDE-BIFTOOL [4] and PDDE-CONT [32], to find and follow EFMs and periodic orbits (corresponding to FOs and ROs) and to determine their stability and bifurcations. We present this information in the plane of feedback strength versus feedback phase, for different values of the filter detuning. This amounts to a study of a physically relevant part of a three-dimensional parameter space.

We finish this introduction with a brief review of the literature on the FOF laser. Experimental studies in comparison with results from numerical integration of the governing rate equations can be found in [9, 36, 37]. As was mentioned, FOs were first found in an experiment reported in [10]. A characterization of the FOs in comparison with a measurement is presented in our short paper [6], which also contains a single stability diagram. (Here we go much further and study how the bifurcation diagrams depend on the filter detuning, which is shown to have a large influence on the stability regions of EFMs, ROs, and FOs.) The connection between FOF and optical injection is the subject of [5, 18], while [19] considers different limits of the FOF laser equations. A reduced model for weak FOF is derived, analyzed, and compared with the full model in [29], and Hopf bifurcation curves giving rise to ROs and FOs are identified. All of these papers consider the case of a filter with a single maximum in its reflectivity (as a function of the frequency). A filter with a minimum at its center frequency is the subject of [33], where continuous wave solutions and bifurcating periodic orbits are determined in a rate equation model.

This paper is organized as follows. In section 2 we present details of the FOF laser and, in particular, the governing DDE model. Section 3 presents the stability of the EFMs, which includes a detailed analysis of how the stability region splits into two parts when the filter is detuned. In section 4 we characterize ROs and FOs and determine their stability regions in the plane of feedback strength versus feedback for different values of the detuning. Finally, we summarize and point to future work in section 6.

2. The FOF laser system. There are a number of ways to set up a frequency selective element in optics [16], including Michealson interferometers, optical gratings, or Fabry–Pérot cavities. Figure 1 shows a looped set-up that has been used in experiments [9]. A fraction of the laser's emission travels through a Fabry–Pérot filter before the light is fed back into the

laser. Optical isolators ensure that there are no unwanted reflections.

The FOF laser can be modeled by rate equations for the complex-valued optical field E of the laser, the real-valued population inversion N of the laser, and the complex-valued optical field F of the filter. In dimensionless form these equations can be written as

$$\frac{dE}{dt} = (1 + i\alpha)N(t)E(t) + \kappa F(t), \tag{2.1}$$

$$T\frac{dN}{dt} = P - N(t) - (1 + 2N(t))|E(t)|^2, \tag{2.2}$$

$$\frac{dF}{dt} = \Lambda E(t - \tau)e^{-iC_p} + (i\Delta - \Lambda)F(t). \tag{2.3}$$

Here the material properties of the laser are given by the linewidth enhancement factor α and the electron life-time T, while P is the pump rate.

The laser is coupled to the filter in (2.1) via the coupling term $\kappa F(t)$, where κ is the feedback rate. Equation (2.3) for the complex envelope F of the filter field is derived by assuming a single Lorentzian approximation for the Fabry–Pérot filter; see, for example, [12, 28] for details. Here τ is the delay time that arises from the finite propagation time of the light in the external feedback loop. The feedback phase C_p in (2.3) measures the exact phase relationship between the laser and the filter fields, whereas Δ is the detuning between the filter center frequency Ω_F and the solitary frequency Ω_0 of the laser, that is,

$$C_p = \Omega_0\tau, \quad \Delta = \Omega_F - \Omega_0. \tag{2.4}$$

Finally, the parameter Λ is the filter width (half-width at half-maximum).

The parameters κ and C_p are our main bifurcation parameters; that is, we consider the bifurcation diagram in the (κ, C_p)-plane. Furthermore, we study how the bifurcation diagram changes with the filter detuning Δ. Throughout this paper, we fix the other parameters at physically realistic values. First of all, we consider $\alpha = 5.0$ and $T = 100$, which are very typical values for describing the material properties of the laser. Furthermore, we set $P = 3.5$, which means that the laser is pumped well above threshold where it is quite robust; by contrast, a laser pumped close to its laser threshold is very sensitive to even small changes of the pump rate. We fix the delay time at $\tau = 500$, so that there are a moderate number of external filtered modes. Finally, we fix the filter width at $\Lambda = 0.007$, which is well within the range where frequency oscillations can be found; see also the discussion in section 6.

Mathematically, (2.1)–(2.3) are a system of DDEs with a single fixed delay τ. As such, they have as phase-space the infinite-dimensional space of continuous functions over the delay interval $[-\tau, 0]$ with values in (five-dimensional) (E, N, F)-space. This infinite-dimensionality of the system makes its analysis quite challenging. Nevertheless, stability and bifurcation theory for DDEs with fixed delays is well developed [2, 15]. Furthermore, numerical bifurcation tools are now becoming available that allow one to find and follow equilibria and periodic solutions and some of their bifurcations; see the review [22].

As is common for optical feedback systems (with the exception of phase conjugation) [25], (2.1)–(2.3) have an S^1-symmetry, given by any rotation of both E and F:

$$(E, N, F) \to (Ee^{ib}, N, Fe^{ib}). \tag{2.5}$$

In other words, solutions are not isolated, which must be taken into account in the numerical continuation; see [14, 22]. Furthermore, there is the trivial 2π-translational symmetry

$$(E, N, F, C_p) \to (E, N, F, C_p + 2\pi) \tag{2.6}$$

in the feedback phase C_p. We already remark that it is convenient to represent bifurcation diagrams on the covering space $\mathbb{R}$ of C_p, that is, over several 2π-intervals.

3. External filtered modes. The external filtered modes are the basic solutions of (2.1)–(2.3) that correspond to a constant-intensity output of the laser. They are conceptually the same as the external cavity modes of the COF laser [12, 31]. Mathematically, the EFMs are group orbits of the S^1-symmetry (2.5) of the form

$$(E(t),\, N(t),\, F(t)) = (E_s e^{i\omega_s t},\, N_s,\, F_s e^{i\omega_s t + i\phi}). \tag{3.1}$$

Here ω_s is a fixed frequency, $E_s \geq 0$ and $F_s \geq 0$ are fixed (real) values of the field amplitude of the laser field and the filtered field, N_s is a fixed level of inversion, and ϕ is a fixed phase shift between the laser field and the filtered field. In other words, the system outputs light with frequency ω_s and constant intensity $I_s = E_s^2$.

An analytical study of the EFMs and their dependence on the filter detuning Δ and filter width Λ has been performed in [12]. Depending on the values of these parameters, the ECMs lie on one or two closed curves in the (ω_s, N_s)-projection, which are known as EFM-components. The number of EFM-components can be derived from a fourth-degree polynomial. We remark that stability properties of EFMs were not considered in [12].

In this paper we perform a bifurcation analysis of EFMs; determine how their stability depends on parameters; find stability boundaries, in particular, Hopf bifurcations; and study the stability of bifurcating periodic orbits. To this end, we use numerical continuation, namely the software package DDE-BIFTOOL [4], to analyze the full DDE (2.1)–(2.3). Some stability curves of periodic orbits are computed with the package PDDE-CONT [32].

Figure 2(a) shows EFMs on a single EFM-component in the (ω_s, N_s)-projection (for $C_p = \pi$). This projection is popular in the laser physics literature, even though neither ω_s nor N_s is a parameter of the system. As C_p is decreased, the EFMs trace out the closed curve—the EFM-component—as is indicated by the arrows. The EFM-component has a "bulge" around the center frequency of the filter, which is detuned slightly in the negative direction with respect to the solitary laser frequency. EFMs are born in pairs in saddle-node (+) bifurcations in the low-inversion region, move along the EFM-component, and then disappear in a second saddle-node bifurcation in the high-inversion region. When C_p is decreased by 2π, each EFM has moved to the position of its right neighbor, and the initial picture is recovered; see also the accompanying animation (65665_01.gif [233KB]).

Figure 2 also shows the relevant stability information of the EFMs. First of all, EFMs are stable along the green parts and unstable along the red parts of the EFM-component in Figure 2(a). The two stability regions are bounded by Hopf bifurcations ($*$), which may lead either to stable relaxation oscillations or frequency oscillations, as will be discussed in section 4. However, there are additional Hopf bifurcations of already unstable EFMs. They are plotted in Figure 2(b) where the EFM-component is shown in terms of its two constituent branches (red and blue curves) that meet at the two saddle-node bifurcations.

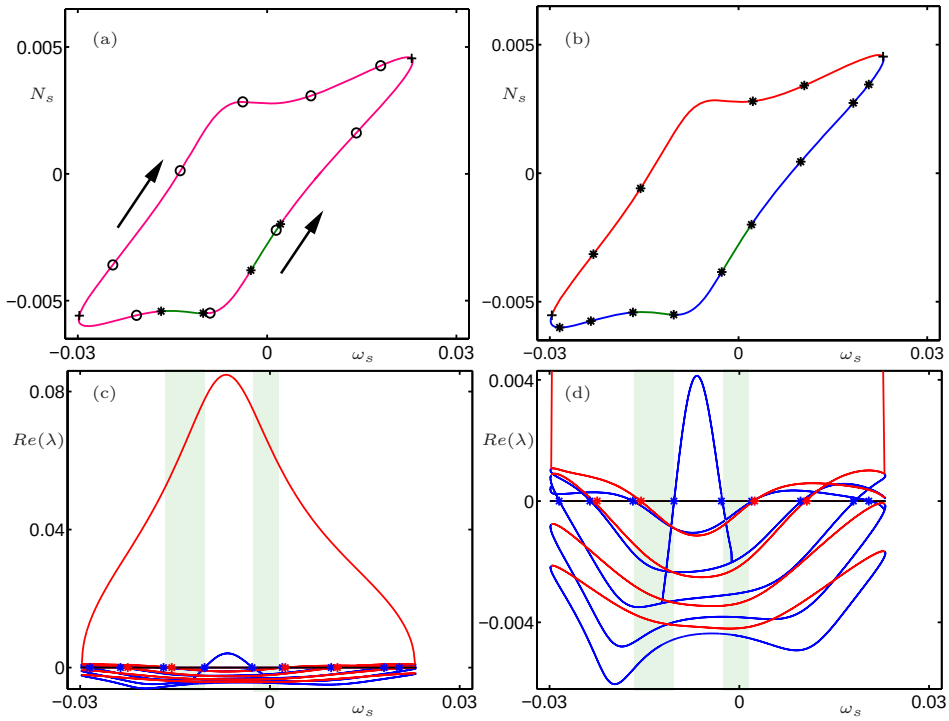


Figure 2. *EFM-component and real parts of the corresponding twelve largest eigenvalues for a small negative detuning of* $\Delta = -0.007$. *Panel* (a) *shows the EFM-component with two green regions of stable EFMs, which are bounded by Hopf bifurcations* $(*)$. *The circles* $(\circ)$ *are individual EFMs for* $C_p = \pi$. *The EFMs trace out the EFM-component as* C_p *is decreased, which is indicated by the arrows; see also the accompanying animation* (65665_01.gif [233KB]). *Panel* (b) *shows the same EFM-component split up into a red branch and a blue branch that meet at saddle-node bifurcations* $(+)$*; also shown are all Hopf bifurcations* $(*)$ *(including those of unstable EFMs). Panels* (c) *and* (d) *show the real parts of the twelve largest eigenvalues of the EFMs along the two branches; the stable regions are highlighted by the green shading, and the Hopf bifurcations appear as crossings of the zero axis.*

We now give an impression of how the stability information that is encoded in the bifurcation diagrams in this paper is generated. Recall that an equilibrium of a DDE with a finite number of fixed delays has an infinite but discrete spectrum. What is more, there are only finitely many eigenvalues that have a real part larger than a given fixed real number, while the limiting behavior of the infinitely many eigenvalues with negative real part is also known; see, for example, [2, 15, 22] for details. DDE-BIFTOOL computes accurately all eigenvalues of the dominant part of the spectrum (with real part above a user-specified value) by combining a heuristic scheme with Newton corrections [4]. In the specific case of the EFM-component in Figure 2, panels (c) and (d) show the real parts of the twelve largest eigenvalues of the EFMs along the red and the blue branch of the EFM-component. (Since there are many pairs of complex conjugate eigenvalues there are fewer than twelve curves.) Notice that all EFMs along the red branch are unstable due to the single (real) positive eigenvalue that emerges from the saddle-node bifurcations. As the enlargement in Figure 2(d) shows, the regions of stability on the blue branch can indeed be identified by the fact that all blue curves of eigen-

values are below the zero axis. The points where curves cross the zero axis are exactly the Hopf bifurcations in panel (b).

We finally remark that each EFM has a zero eigenvalue, which is due to the S^1-symmetry (2.5) of the system; see also [14, 22]. This trivial eigenvalue is computed by DDE-BIFTOOL as part of the spectrum. In fact, the accuracy of the stability information can be checked by monitoring the error of the trivial eigenvalue zero. In the continuation the number of calculated eigenvalues is influenced by the user who sets the boundary for the real part. Throughout the paper, we calculate all eigenvalues of EFMs with a real part above -0.1 (compare the scale of Figure 2(d)). These are many more than the twelve that are shown, but their real parts remain negative throughout the EFM component.

Figure 2(a) already indicates that there are values of C_p for which more than one EFM is stable. To bring out this point better, Figure 3 shows the EFMs in the (κ, N_s)-projection for six different values of the feedback phase C_p. In other words, we now plot actual branches of EFMs as a function of the parameter κ. For small feedback rates close to zero, only one EFM exists, which is actually (the continuation of) the stable solitary laser mode. As κ is increased, new EFMs are born in pairs in saddle-node bifurcations. One of these EFMs may be stable (green parts of the curve). As the feedback rate κ is increased further, stable EFMs destabilize in Hopf bifurcations. The different panels of Figure 3 show the EFM structure for six different values of C_p over one cycle of 2π. As a function of C_p the branches of EFMs move until the same situation is regained after C_p has been changed over 2π. During this cyclic process, the saddle-node bifurcations trace out the light blue curve, and the Hopf bifurcations the light red curves, respectively; see also the accompanying animation (65665_02.gif [304KB]). Note that tracing out all EFM branches over an interval of 2π in C_p is equivalent to tracing out a single EFM branch over several intervals of 2π in the covering space.

3.1. Stability regions of the external filtered modes. The stability information of EFMs for $\Delta = -0.007$ is presented in Figure 4 as a two-parameter bifurcation diagram in the (κ, C_p)-plane. In panel (a) the bifurcation diagram is shown in the covering space of C_p over several cycles of 2π. The green region is the stability region of EFMs, where all eigenvalues have negative real part. Its boundary is formed by saddle-node bifurcations (blue curves) and by Hopf bifurcations (red curves), where a real eigenvalue or a complex pair of eigenvalues, respectively, cross the imaginary axis. At codimension-two Bogdanov–Takens (BT) and saddle-node Hopf (SH) bifurcation points the nature of the stability boundary changes from saddle-node to Hopf bifurcation. At the BT point, characterized by a double zero eigenvalue of the EFM, the Hopf curve ends and the frequency of the associated periodic orbit goes to zero. On the other hand, at the SH point, characterized by a real zero eigenvalue and a pair of purely imaginary complex conjugate eigenvalues of the EFM, the saddle-node and Hopf curves are tangent and change from supercritical to subcritical; the frequency of the associated periodic orbits is finite. We refer to [13, 26] for background reading on bifurcation theory. The stability boundary is formed to the right by a total of four different Hopf bifurcation curves, which intersect at double-Hopf bifurcation points. A Hopf bifurcation curve that does not bound the stability region corresponds to a pair of purely imaginary complex conjugate eigenvalues of an already unstable EFM.

The bifurcation diagram shown in Figure 4(a) is only one of infinitely many copies under the 2π-translation symmetry of C_p. To reveal the considerable degree of multistability of

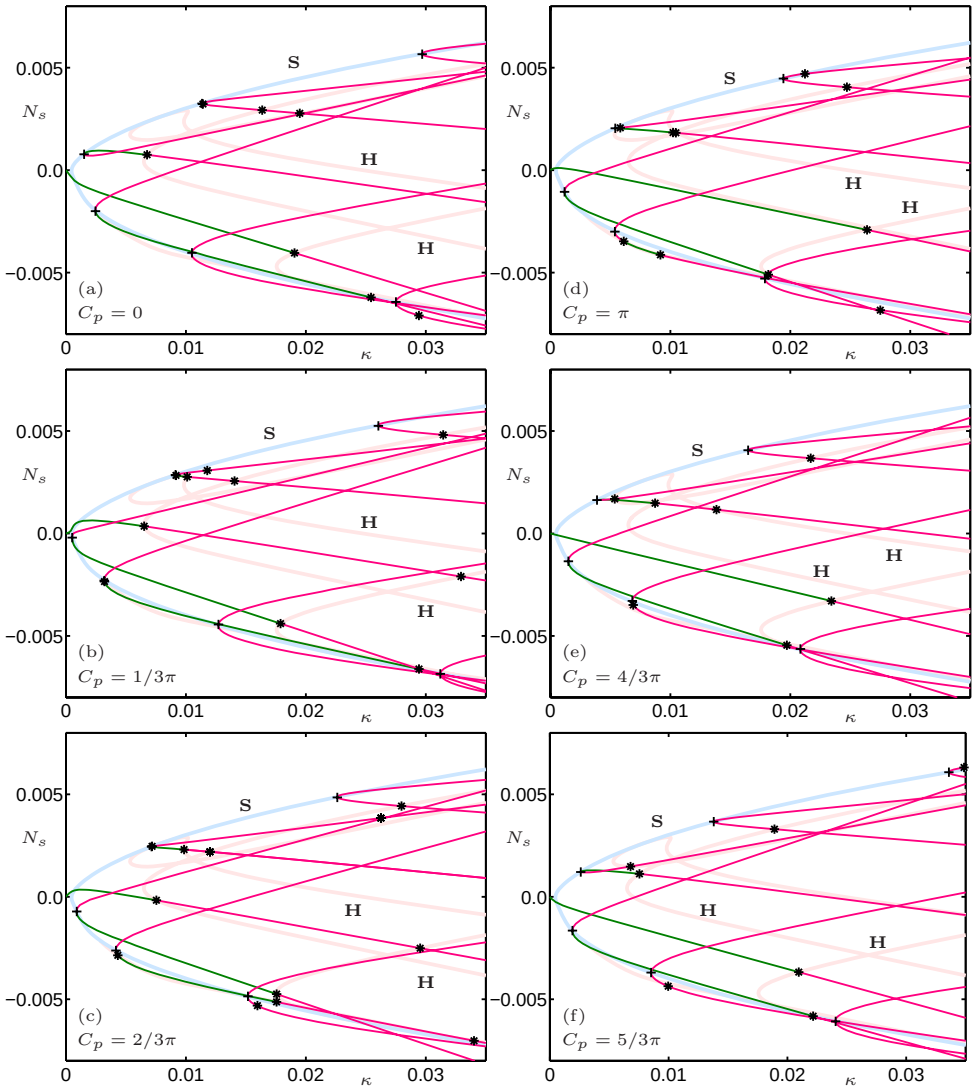


Figure 3. *The EFMs in the (κ, N_s)-projection for six different values of C_p, as indicated in the panels, and for a small negative detuning of $\Delta = -0.007$; same conventions for stability and bifurcations as in Figure 2. In addition, curves of saddle-node bifurcations (light blue) and Hopf bifurcations (light red) are plotted. As C_p is changed each bifurcation follows its respective curve; see also the accompanying animation* (65665_02.gif [304KB]).

EFMs that was already apparent from Figure 3, we show in Figure 4(b) the stability regions of all EFMs over a fundamental 2π-interval of C_p. (Note that the top and bottom of panel (b) can be glued together to obtain a bifurcation diagram on the half-cylinder $S^1 \times \mathbb{R}^+$.) While panel (b) clearly shows the multistability of the system, a representation of the bifurcation diagram in the covering space as in panel (a) is more convenient for distinguishing bifurcation curves that make up the stability boundary.

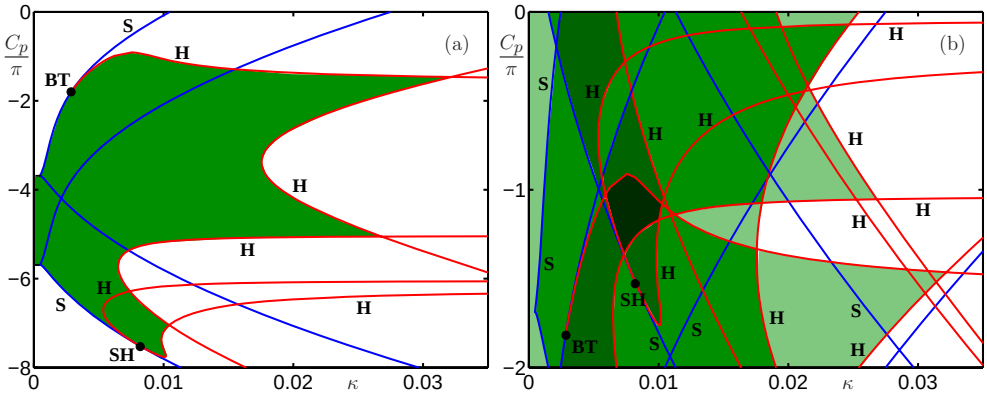


Figure 4. *Two-parameter bifurcation diagram in the (κ, C_p)-plane for a small nonzero detuning of $\Delta = -0.007$. Saddle-node bifurcation curves (S) are blue, Hopf bifurcation curves (H) are red, and regions of stable EFMs are green. Codimension-two BT and SH points are indicated by dots. Panel* (a) *shows the bifurcation diagram in the covering space (over several 2π-intervals of C_p), while panel* (b) *shows it on a fundamental 2π-interval of C_p. Different shades of green indicated regions of different numbers of simultaneously stable EFMs.*

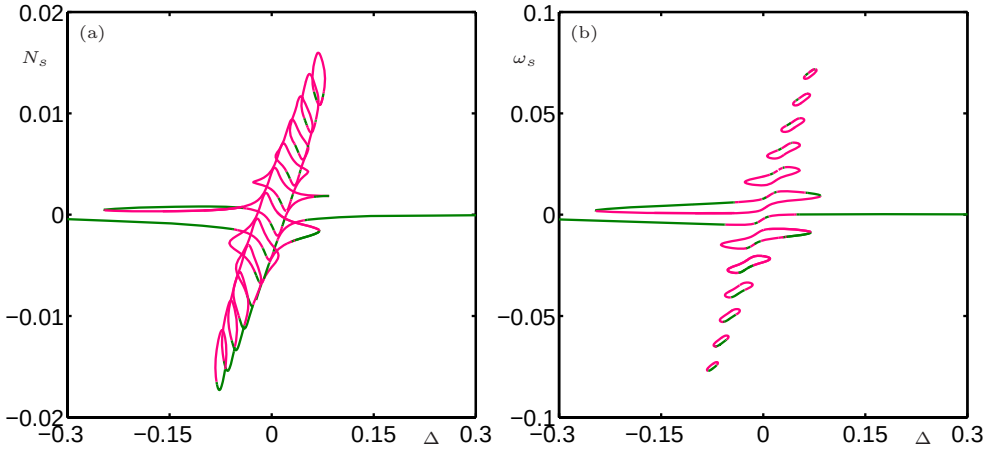


Figure 5. *One-parameter bifurcation diagrams showing branches of EFMs as a function of Δ for fixed $\kappa = 0.02$. Panel* (a) *shows the N_s-values, and panel* (b) *the ω_s-values of the EFMs; stable parts are green and unstable parts are red.*

3.2. Dependence of EFM stability on the detuning. We now consider what effect the influence of the detuning Δ has on the stability of the EFMs. We start with the one-parameter bifurcation diagram in Figure 5, where the N_s- and ω_s-values of the EFMs are shown as functions of Δ for fixed $\kappa = 0.02$. The EFM originating from the solitary laser solution exists for all values of Δ and forms a single branch. For large positive and negative detunings (when there is effectively no feedback any more) it approaches $N_s = 0$ and $\omega_s = 0$, respectively. Around zero detuning additional EFMs exist on isolas (closed curves) that are bounded by saddle-node bifurcations. For each fixed Δ there are finitely many EFMs that lie either on a single EFM component (as in Figure 2) or on two separate EFM-components; cf. [12].

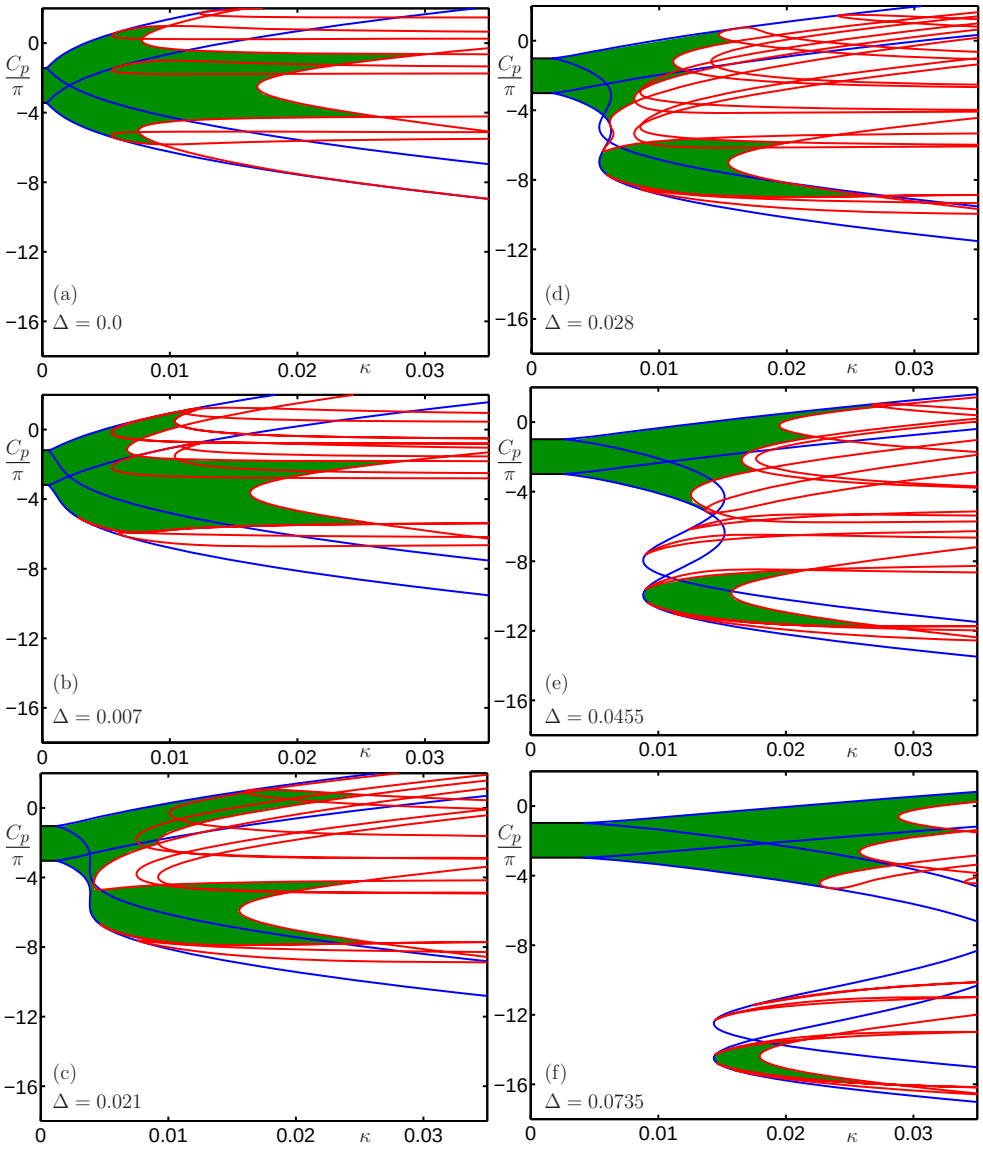


Figure 6. *Two-parameter bifurcation diagram in the (κ, C_p)-plane for increasing detuning Δ, as indicated in the panels; see also the accompanying animation* (65665_03.gif [617KB]). *Curves and regions are colored as in Figure 4.*

We now discuss global aspects that a change of the detuning Δ has on EFM stability regions. Representative bifurcation diagrams in the (κ, C_p) are shown in Figures 6 and 7; see also the accompanying animation (65665_03.gif [617KB]). Specific details of transitions through codimension-three points are discussed in section 3.3.

Figure 6 shows bifurcation diagrams in the (κ, C_p)-plane for six different values of positive detuning from $\Delta = 0$ to $\Delta = 0.0735$. In the individual panels only those Hopf bifurcation curves are shown that form part of the EFM stability boundary and those that become

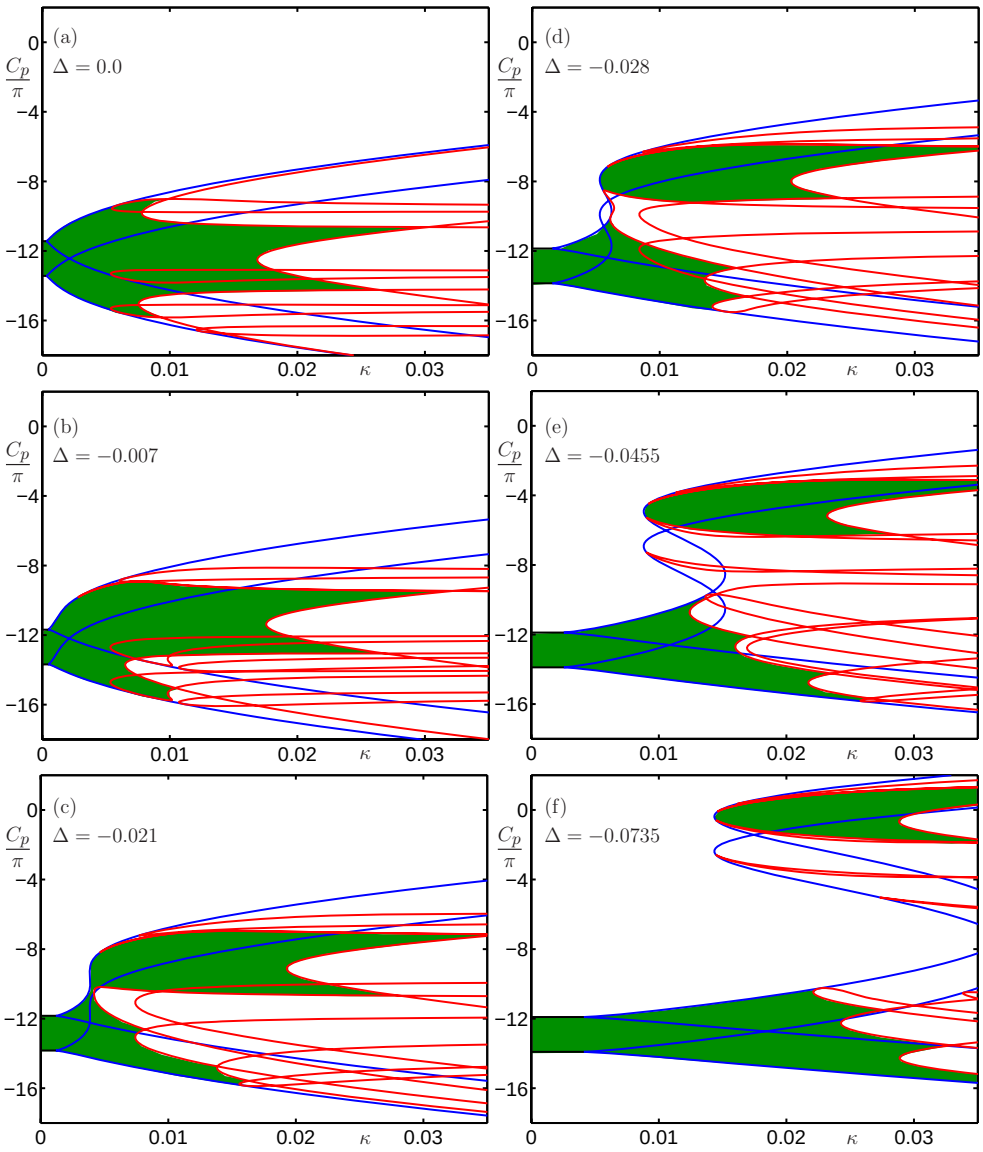


Figure 7. *Two-parameter bifurcation diagram in the (κ, C_p)-plane for decreasing detuning Δ, as indicated in the panels; see also the accompanying animation* (65665_03.gif [617KB]). *Curves and regions are colored as in Figure* 4.

relevant as the detuning changes. For $\Delta = 0$ in panel (a) the EFM stability region is almost symmetrical, which indicates only a weak influence (on the EFM structure) of the phase-amplitude coupling. Note that the boundary changes from a saddle-node bifurcation curve to a Hopf bifurcation curve at two saddle-node Hopf points. With increasing Δ the overall shape of the stability region changes and becomes less symmetrical. Already for $\Delta = 0.007$ in panel (b) the Hopf curve forming the lower boundary of the EFM stability region ends in a BT point, as was discussed above. The nature of this change from an SH to a BT point is

discussed in more detail in section 3.3.

Increasing Δ, the stability region deforms further until, at $\Delta \approx 0.021$, it splits into two parts; see Figure 6(c). This transition is due to a transition through a codimension-three degenerate SH bifurcation, as is discussed in more detail in section 3.3. For even higher values of Δ the EFM stability region consists of two parts; see panels (d)–(f). It is known that there is a boundary curve in the (Λ, Δ)-plane, inside of which one finds two distinct EFM-components; this curve scales linearly with the feedback rate κ [12]. The two separate EFM stability regions have different physical meaning: one corresponds to operation of the laser around its solitary frequency, and the other to operation around the filter frequency. Note that, irrespective of their stability, EFMs exist in the whole area bounded by the saddle-node curves.

However, for small κ stable EFMs are only possible around the solitary laser frequency (around $C_p = -2\pi$). For higher values of κ EFMs around the center frequency of the filter may also be stable. These form an "island" of stable EFMs. Further increasing Δ separates the two stability regions. Specifically, the island of EFMs located around the center frequency of the filter moves down in the (κ, C_p)-plane and simultaneously becomes smaller in size; see panels (d)–(f). Physically, the filter center is on the blue side of the laser; that is, it has a higher frequency than the laser. However, feedback causes a red-shift of the instantaneous laser frequency. Eventually, when Δ is too large the laser is unable to support EFMs located around the center frequency of the filter. Mathematically, this means that the island actually shrinks down to a point and disappears. On the other hand, the stability region around the solitary laser frequency extends with increasing Δ to higher feedback rates κ, because the system now operates at the tail of the filter where the effective feedback rate is smaller. As a consequence, instabilities, such as the Hopf bifurcations bounding the stability region to the right, arise only for quite high levels of feedback κ.

We find a similar global transition in Figure 7, where we decrease the detuning Δ from zero. In the individual panels of Figure 7 the detuning has the same modulus as in the respective panels of Figure 6 (only the sign has changed). This allows one to study the difference between positive and negative detuning. The effect of decreasing the detuning is qualitatively the same. In particular, the EFM region splits into two separate regions where the "island" located around the center frequency of the filter is now at the top of the figure. Furthermore, the region of stable EFMs centered around the solitary laser frequency extends to higher values of κ with decreasing Δ. However, there are quite significant quantitative differences between Figures 6 and 7. In particular, the separate island of stable EFMs around the center frequency of the filter is notably larger for negative detunings. This is due to the filter being on the red side of the laser. In combination with the red-shift of the solitary laser frequency caused by the feedback, this leads to a large region of stability of these EFMs.

3.3. Transitions through codimension-three bifurcations. The boundary of the stable EFM regions is formed by either saddle-node bifurcation curves or by Hopf bifurcation curves. The switch-over points from one type of boundary to the other are codimension-two bifurcation points, namely either a BT bifurcation or an SH point. As was already mentioned in the previous section, we find that the BT point appears as the "switch-over" point only for sufficiently large Δ. The transition from an SH point to a BT switch-over point involves codimension-three bifurcations, as is shown in Figure 8. Panel (a) shows a situation as in

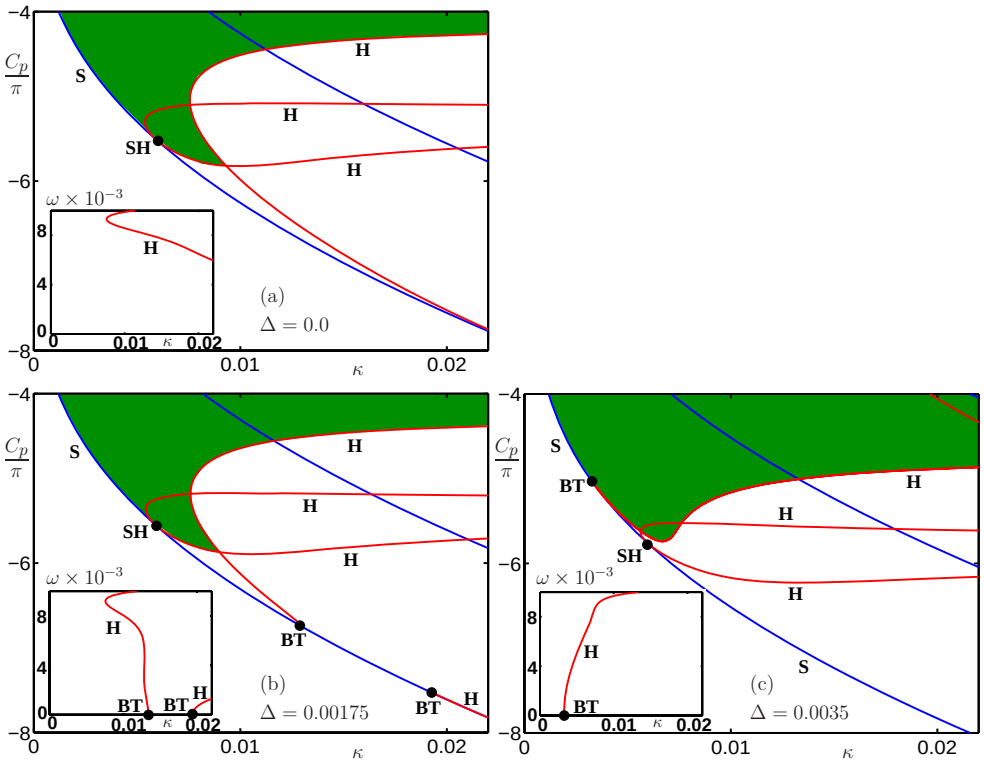


Figure 8. *Enlarged views of bifurcation diagrams in the (κ, C_p)-plane near the creation of two BT points. The insets of each panel show the frequency of the periodic orbits along branches of Hopf bifurcations, which go to zero at the BT points.*

Figure 6(a), where the switch-over from saddle-node to Hopf bifurcation is due to an SH point. Notice a second Hopf bifurcation curve that passes close to the saddle-node curve. The inset shows the frequency of the periodic orbit along this curve. As Δ is increased, the second Hopf curve becomes tangent to the saddle-node curve at a degenerate BT point [3], after which it splits up and ends at two newly created BT points. That these points are indeed BT points is evidenced by the inset, which shows that the frequency of the respective periodic orbits goes to zero at the BT points, as is demanded by theory. As Δ is increased further, one BT point moves to the right and out of the region of consideration, while the other BT point moves to the left. The latter eventually passes through the SH point at another codimension-three bifurcation, namely a Bogdanov–Takens–Hopf bifurcation, where there is a semisimple double eigenvalue and a pair of complex eigenvalues with zero real part. To our knowledge, this bifurcation with a center manifold of dimension four has not been described in the literature. In the present situation the BT point "moves through" the SH point, so that it forms the switch-over point from now on; see Figure 6(c).

The second transition involving codimension-three bifurcations that we discuss here in detail concerns the mechanism in which the single region of EFM stability splits up into two separate regions. Enlarged views near this splitting are shown in Figure 9. In panel (a) there is still a narrow channel connecting the two parts of the EFM stability region. As Δ is increased,

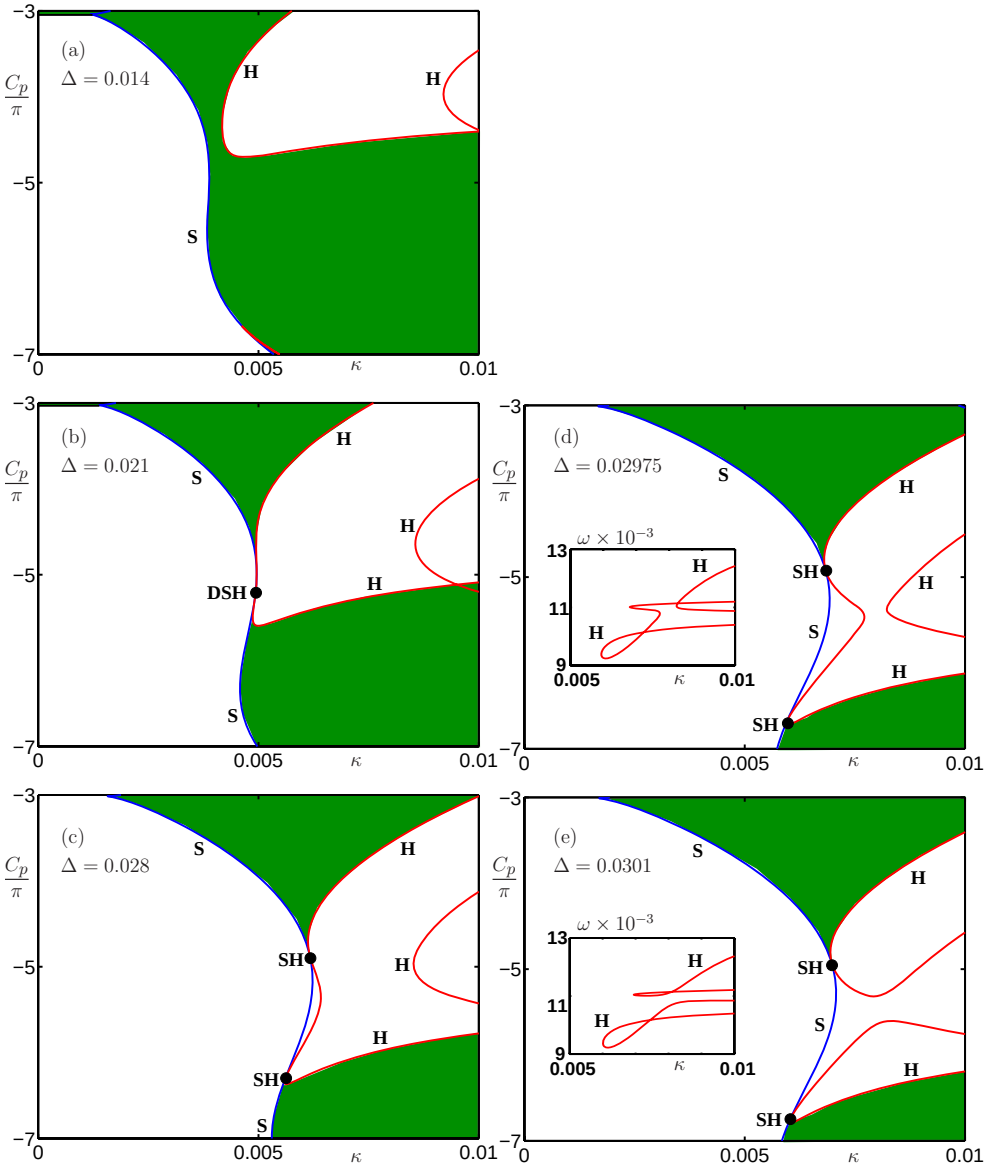


Figure 9. *Enlarged views of bifurcation diagrams in the (κ, C_p)-projection near the splitting-off of the EFM stability region. The insets of panels* (d) *and* (e) *show the frequencies of the periodic orbits along the Hopf curves involved in a singularity transition.*

a saddle-node curve and a Hopf curve become tangent at a degenerate SH point DSH [24]; see panel (b). When Δ is increased further, two codimension-two SH points are created, each of which is now a switch-over point; see panel (c). This creates two separate EFM stability regions, one around the solitary laser frequency and one around the center frequency of the filter. Physically this corresponds to a level of detuning that is so large that the system cannot support "mixed" EFMs with frequencies in between the solitary laser frequency and the center frequency of the filter any longer.

When Δ is increased even further, we find a transition through a saddle in the surface of Hopf bifurcations in (κ, C_p, Δ)-space. This manifests itself as a change in how four branches of Hopf bifurcations connect; see the transition shown in Figure 9(d) and (e). Specifically, two Hopf curves approach each other, connect in a different way, and then separate again. As a result, the two SH points now lie on two different (unconnected) Hopf bifurcation curves. This "completes" the splitting of the EFM stability islands that was discussed in the previous section.

4. Different types of oscillations. In laser systems with delayed feedback one can identity different characteristic time scales. From the physical point of view, one expects these time scales to show up as frequency components in the dynamics. In the FOF laser at least two time scales can be identified, one related to the solitary laser and a second related to the delay time.

A typical type of oscillation that one expects in any laser system are the ROs—a periodic exchange of energy between the optical field and the population inversion. In a solitary laser they are damped but can be excited by a perturbation, after which the laser relaxes down to constant output; hence, the name relaxation oscillation. ROs are fast oscillations. In the solitary laser their frequency is given by $\omega_{RO} = \sqrt{2P/T}$, where P is the pump rate and T the photon life-time.

Since we are dealing with a laser with delayed optical feedback, one also expects to find oscillations with a frequency of $\omega_{FO} \approx 2\pi/\tau$. Recall that τ is the delay time, that is, the time it takes for one external roundtrip of the light. This type of external roundtrip oscillation has been found experimentally in the FOF laser [10]. Surprisingly, they have the feature that oscillations occur only in the laser frequency, while the laser intensity remains practically constant. FOs are "untypical" for semiconductor lasers, which are characterized by a strong coupling between amplitude and phase of the electric field (as expressed by large values of the parameter α, typically well above 1). In other words, one would expect any FO to be accompanied by intensity oscillations of a similar amplitude.

In this section we perform a detailed bifurcation analysis of the periodic orbits that are born when the EFMs become unstable in Hopf bifurcations. Similar to the case for EFMs, the Floquet spectrum of a periodic orbit of a DDE with a finite number of fixed delays is discrete, with zero as its only accumulation point [2, 15, 22]. DDE-BIFTOOL uses this fact and accurately calculates the finitely many Floquet multipliers that are larger than a user-specified value. This allows us to study which type of oscillations, ROs and FOs, appear and where they are stable. In particular, we are able to show that FOs are stable in large, experimentally accessible regions of the (κ, C_p)-plane. In fact, they are the dominant type of oscillations for moderate levels of detuning Δ.

Figure 10(a) shows the bifurcation diagram in the (κ, C_p)-plane for $\Delta = 0$. As expected, the Hopf bifurcations bounding the EFM stability region give rise to oscillations, which are stable in the colored regions, where all of their Floquet multipliers lie within the unit circle of the complex plane. There is a region (orange) of stable ROs, which can be accessed simply by following the solitary laser solution (around $C_p = 0$) towards higher levels of feedback strength κ. The region of stable ROs is bounded to the right by torus bifurcation curves, where DDE-BIFTOOL detects a complex conjugate pair of Floquet multiplier on the unit circle. Furthermore, there are two regions (cyan and purple) of stable FOs in Figure 10(a).

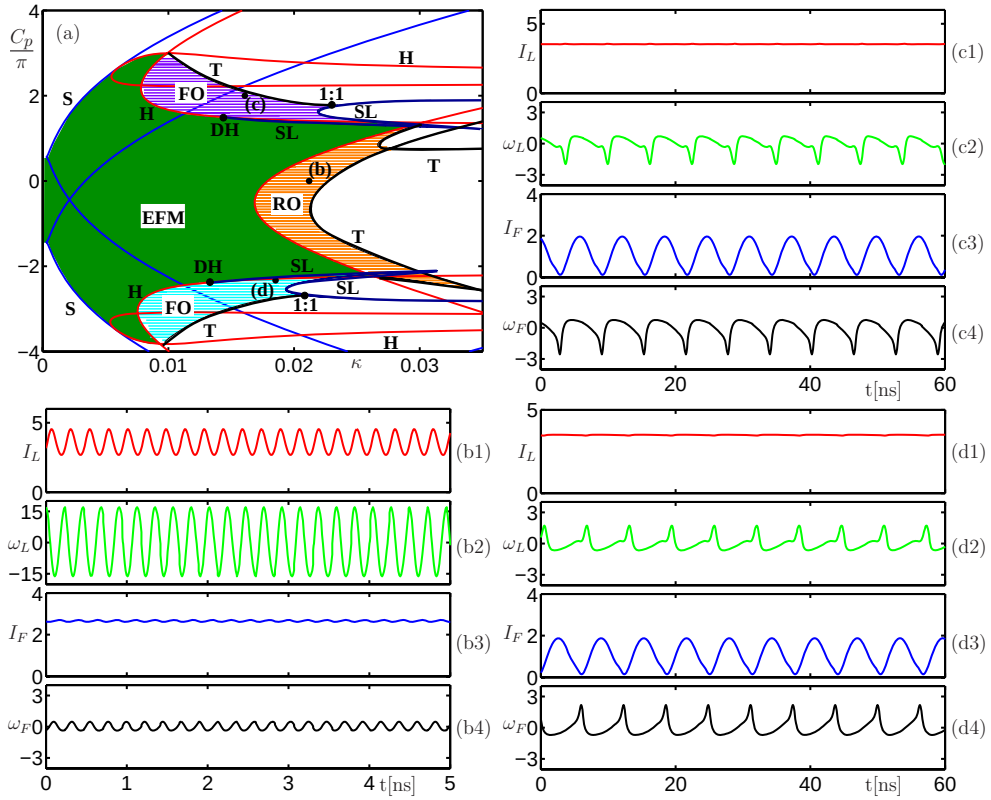


Figure 10. *Bifurcation diagram and examples of time series for* $\Delta = 0$. *Panel* (a) *shows the two-parameter bifurcation diagram in the* (κ, C_p)*-plane with different stability regions, namely, of EFMs (green), of ROs (orange), and of FOs (cyan and purple); boundaries are formed by saddle-node (S), Hopf (H), saddle-node of limit cycles (SL), and torus (T) bifurcations; shown also are* 1:1*-resonance and degenerate Hopf (DH) points. The black dots labeled* (b), (c), *and* (d) *indicate the parameter values of the time series in panels* (b)–(d), *which show the laser intensity* I_L, *the laser frequency* ω_L *(in units of* 10^9*), the filter intensity* I_F, *and the filter frequency* ω_F *(in units of* 10^9*). Case* (b) *shows ROs, and* (c) *and* (d) *show two examples of FOs.*

We note that the two regions of FOs are almost mirror images of each other under reflection around the central frequency of the filter. This indicates that the amplitude-phase coupling (the parameter α) has only a small influence; see also [29]. However, note that the stability region of ROs is actually quite asymmetrical in Figure 10(a). Another important feature is that stable FOs occur for much lower levels of κ than ROs. The FO stability regions are bounded to the right by a torus bifurcation curve, or by a curve of saddle-node of limit cycle bifurcations. The latter is detected by DDE-BIFTOOL as a real Floquet multiplier crossing the unit circle at 1. The torus curves T emerge from double Hopf points and meet the SL curves at 1:1 resonance points.

The stability regions of the different oscillations in Figure 10(a) have been determined by computing many one-parameter cross sections for fixed C_p and continuing with DDE-BIFTOOL the respective periodic solutions and their Floquet multipliers in κ; see also section 5. In fact, the horizontal colored lines in Figure 10 are actual branches of stable periodic orbits. The codimension-one bifurcation curves of periodic orbits, T and SL, were found

as bifurcation points on the one-parameter branches and continued directly with PDDE-CONT [32].

Figure 10 also shows examples of oscillations in the three different stability regions. In each case we show time series of the laser intensity I_L, the laser frequency ω_L, the filter intensity I_F, and the filter frequency ω_F. Panel (b) shows a typical example of ROs, where both the laser intensity I_L and the laser frequency ω_L oscillate. For the (typical) laser parameters considered here the ROs have a frequency of 4.21 GHz. Note that ROs hardly show any dynamics of the filter—it remains more or less passive.

Figure 10 (c) and (d) are examples of FOs, which are clearly external roundtrip oscillations with a frequency on the order of $1/\tau$; that is, they are much slower than ROs (note the different scale on the time axis). Furthermore, the laser intensity is almost constant for FOs (certainly compared to ROs). However, we find dynamics in the filter field—in both the filter intensity I_F and the filter frequency ω_F. In other words, the filter effectively compensates for the intensity dynamics that one would normally expect in a semiconductor laser. The difference between the FOs shown in Figure 10(c) and (d) is in the phase relationship between the laser frequency ω_L and the feedback intensity I_F. For the FOs shown in panel (c), from the upper (purple) FO stability region of panel (a), ω_L and I_F are almost in phase. This is in accordance with the fact that they oscillate around the left flank of the filter, where an increase of the frequency ω_L results in increased transmission of the intensity I_F. For the FOs shown in panel (d), from the lower (cyan) FO region of Figure 10(a), on the other hand, ω_L and I_F are almost in antiphase. This time the oscillations are around the right flank of the filter, so that I_F decreases as ω_L increases.

4.1. Dependence of RO and FO stability on the detuning. The strong influence of the detuning Δ on the EFM stability regions that was discussed in section 3.2 is mirrored by a strong influence on the stability regions of ROs and FOs.

Figure 11 shows the bifurcation diagram in the (κ, C_p)-plane for three positive values of Δ. As the detuning is increased, the RO stability region (orange) becomes smaller and moves towards lower values of C_p, but overall does not change very much. Notice that ROs appear only when EFMs that are located around the central frequency of the filter become unstable. The influence on the FO stability regions of changing Δ is much more dramatic. The lower (cyan) FO stability region decreases significantly in size and then disappears entirely; see Figure 11(b) and (c). On the other hand, the upper (purple) FO stability region grows substantially in size; see panel (b). For larger values of Δ this FO stability region occupies a substantial area between the two separate EFM stability regions; see panel (c). It is bounded in the "EFM stability gap" by a curve SL of saddle-node bifurcations of limit cycles on the left. Its right boundary also becomes much more complicated. For sufficiently large Δ we find that FOs may become unstable in period-doubling bifurcations, so that we find regions of stable period-doubled FOs.

Throughout the whole region the approximate in-phase relationship between ω_L and I_F is preserved. This is because the stable FO region never extends to the other side of the filter flank. However, in Figure 11(b), new oscillations become stable and coexist with the upper stable FO region. There is a small region around $(\kappa, C_p) = (0.01, 1)$ where faster FOs are stable; they feature a period $T_{FO} \approx \tau/2$. This region becomes larger for even larger Δ; see Figure 11(c). Initially, for $\Delta = 0.014$ in Figure 11(b), these fast FOs are created in

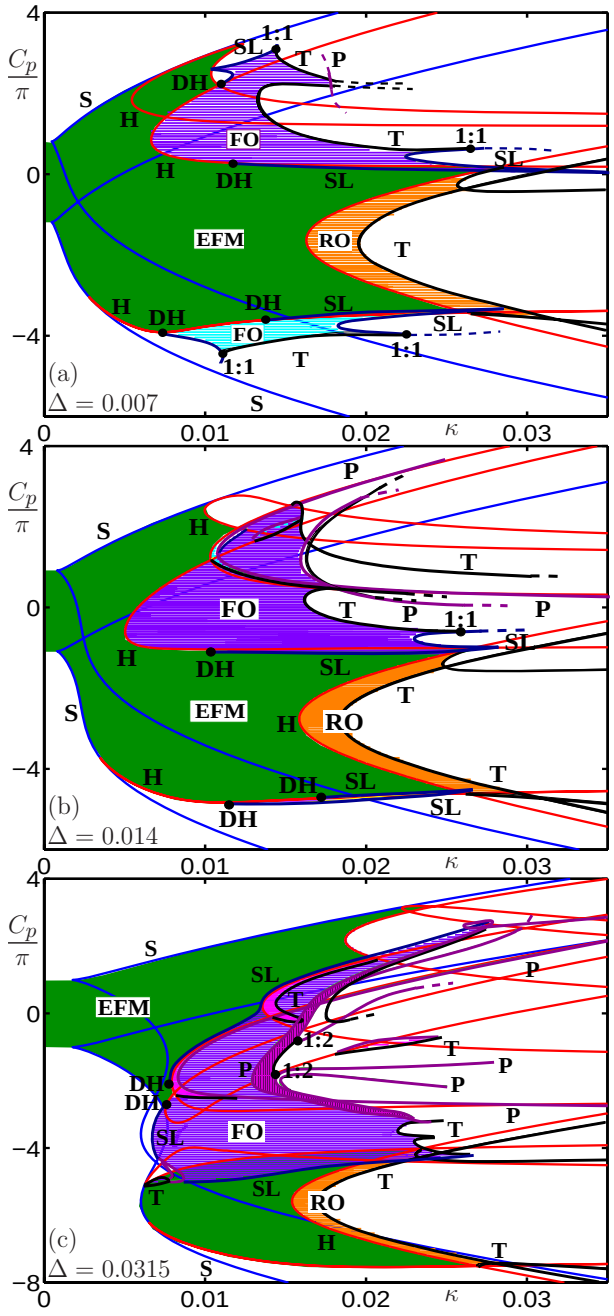


Figure 11. *Two-parameter bifurcation diagram in the (κ, C_p)-plane for increasing detuning Δ as indicated in the panels; notation is as in Figure 10 with the addition of period-doubling (P) bifurcations.*

supercritical Hopf bifurcations of EFMs. They destabilize in torus bifurcations T as C_p is decreased and in subcritical period-doubling bifurcations (P) as κ is increased. For larger detuning, $\Delta = 0.0315$ in Figure 11(c), there are already two regions of fast FOs. The upper one features a period of $T_{FO} \approx \tau/4$ and undergoes a torus bifurcation for increasing κ. The

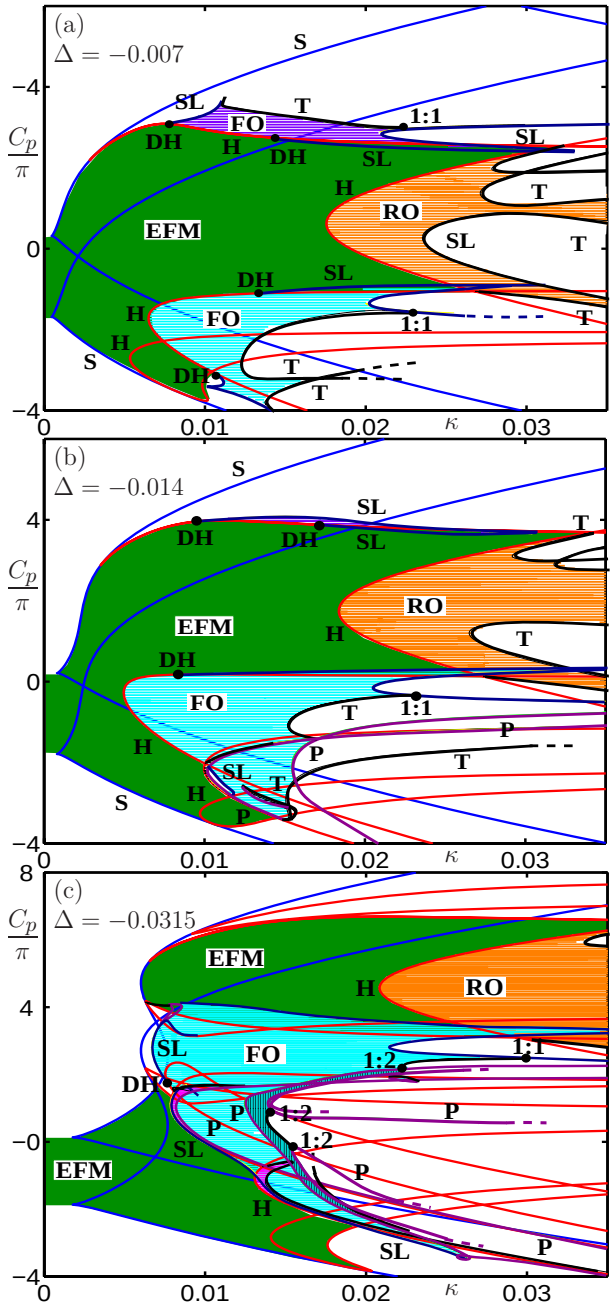


Figure 12. *Two-parameter bifurcation diagram in the (κ, C_p)-plane for decreasing detuning Δ as indicated in the panels; notation is as in Figure 10 with the addition of period-doubling (P) bifurcations.*

lower one features a period of $T_{FO} \approx \tau/2$, undergoes a subcritical period-doubling bifurcation and connects to the large FO region. This will be discussed in more detail in section 5.

Figure 12 shows the bifurcation diagram in the (κ, C_p)-plane for three different values of negative detuning. Again we find a large region of stable ROs around the filter center. Similar

to Figure 11(a), there are still two regions of stable FOs; in the upper (purple) region there is an in-phase relationship between ω_L and I_F (left flank of the filter profile), while in the lower (cyan) region there is an antiphase relationship between ω_L and I_F (right flank of the filter). However, when Δ is decreased, the (purple) region of stable in-phase FOs increases in size, whereas the (cyan) region of stable antiphase FOs becomes smaller and then disappears; see Figure 12(b) and (c). As a result, the regions of stable antiphase FOs extend over a large area in the (κ, C_p)-plane, between the regions of stable EFMs. As is the case for positive Δ, this large FO stability region is characterized by the appearance of period-doubled solutions. In terms of the EFM and FO stability regions, the panels of Figure 12 for negative Δ are almost reflected images of those in Figure 11 positive Δ. This again highlights the weak influence of the amplitude-phase coupling parameter α on these dynamics.

5. Beyond stable frequency oscillations. Evidence of possible period-doublings of FOs has been found experimentally [6], and our bifurcation analysis shows that regions of stable period-doubled FOs do exist for sufficiently large detuning Δ. We now consider this transition of FOs in more detail. Figure 13(a1) shows a one-parameter bifurcation diagram, where we plot branches of EFMs and FOs as functions of κ for fixed $(\Delta, C_p) = (0.0315, -2.26\pi)$. Figure 13(a2) is an enlarged view around a period-doubling cascade of FOs. The appearance of stable period-doubled FOs is actually not so straightforward. A stable EFM undergoes a Hopf bifurcation H, from which a branch of FOs emerges with a period of around $\tau/2$. The branch of periodic orbits then undergoes a subcritical period-doubling bifurcation P. (Note that, after further bifurcations, this branch connects to a second Hopf bifurcation on a different EFM branch.) The initially unstable period-doubled periodic orbit stabilizes in a saddle-node of limit cycle bifurcation SL. The period of these stable FOs is now approximately τ, and this part of the branch lies in the (purple) FO stability region in Figure 11(c). This branch destabilizes in another period-doubling bifurcation P, which is the first in a cascade of supercritical period-doublings.

Figure 13(b1)–(b4) shows time series and Figure 13(b5) the corresponding trajectory in the (I_L, N)-plane for the point on the main stable FO branch; labeled (b) in panel (a2). As is typical for FOs, the laser intensity I_L is almost constant. Note the small scale of the x-axis in Figure 13(b5). The inversion N of the laser is directly related to the laser frequency ω_L. Figures 13(c1)–(c4) are time series, and Figure 13(c5) is a trajectory in the (I_L, N)-plane for a point on the stable period-doubled FO branch; labeled (c) in panel (a2). The time series looks quite similar to those in Figure 13(b) and, in particular, the intensity I_L is still virtually constant. However, as is most prominent in the frequency dynamics of the laser, it can be seen that any two consecutive dips of ω_L are no longer identical. Indeed the period of oscillations is now about twice what it was before in panels (b). Also the phase portrait in Figure 13(c5) shows that the periodic orbit has period-doubled.

We now discuss another way in which FOs can become unstable. This is related to the question of whether it is possible to have a "mixing" of features of FOs and of ROs. One would expect to find such mixed types of oscillation near the double-Hopf points, that is, where the boundaries of the stability regions of FOs and ROs come together. Figure 14 shows an example of a stable FO (an external roundtrip oscillation of period about τ) that shows small oscillations with the typical and much faster RO frequency. Panel (a) is a one-parameter bifurcation diagram, where we plot branches of EFMs and FOs as a function

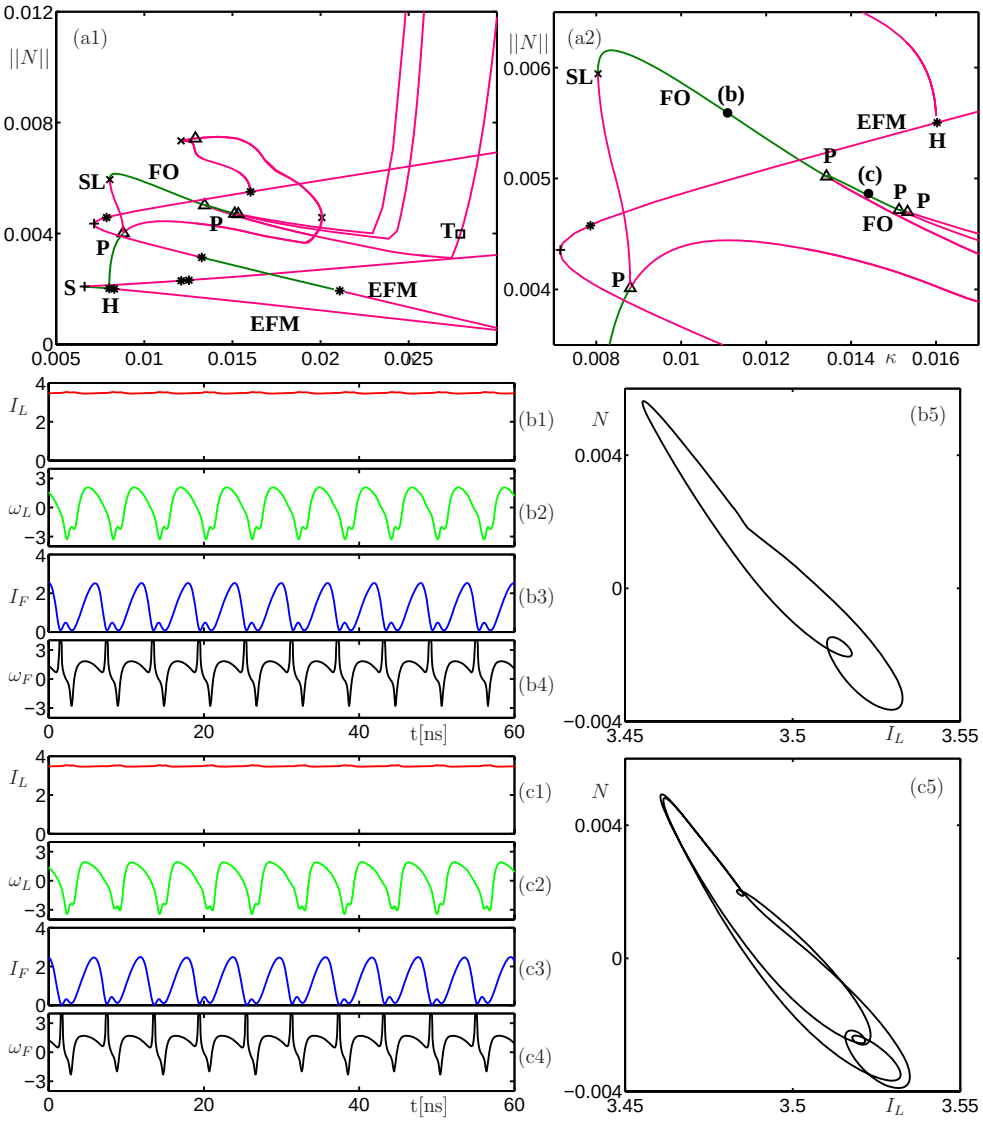


Figure 13. *Panel* (a1) *shows a one-parameter bifurcation diagram in the parameter* κ *for fixed* $(\Delta, C_p) = (0.0315, -2.26\pi)$. *As is clear from the enlargement in panel* (a2)*, there is a region of stable period-doubled FOs after the first period-doubling bifurcations (P) of a period-doubling cascade. Labels* (b) *and* (c) *indicate points on the FO branch for which time series and phase portraits are shown in panels* (b) *and* (c).

of κ for fixed $(\Delta, C_p) = (0.0315, -4\pi)$. The inset shows the period of the periodic orbit along the FO branch. Starting from a Hopf bifurcation H, of an unstable EFM around $(\kappa, ||N||) \approx (0.009, 0.0041)$, a branch of unstable FO emerges. After undergoing several saddle-node of limit cycle bifurcations, SL, it eventually becomes stable. The stable part of the FO branch is almost horizontal until, for a higher feedback rate around $\kappa = 0.0225$, the norm $||N||$ increases dramatically. This is a smooth transition and not a bifurcation, even though it looks very sudden in this projection. (Note that the FO period does not show such a sudden

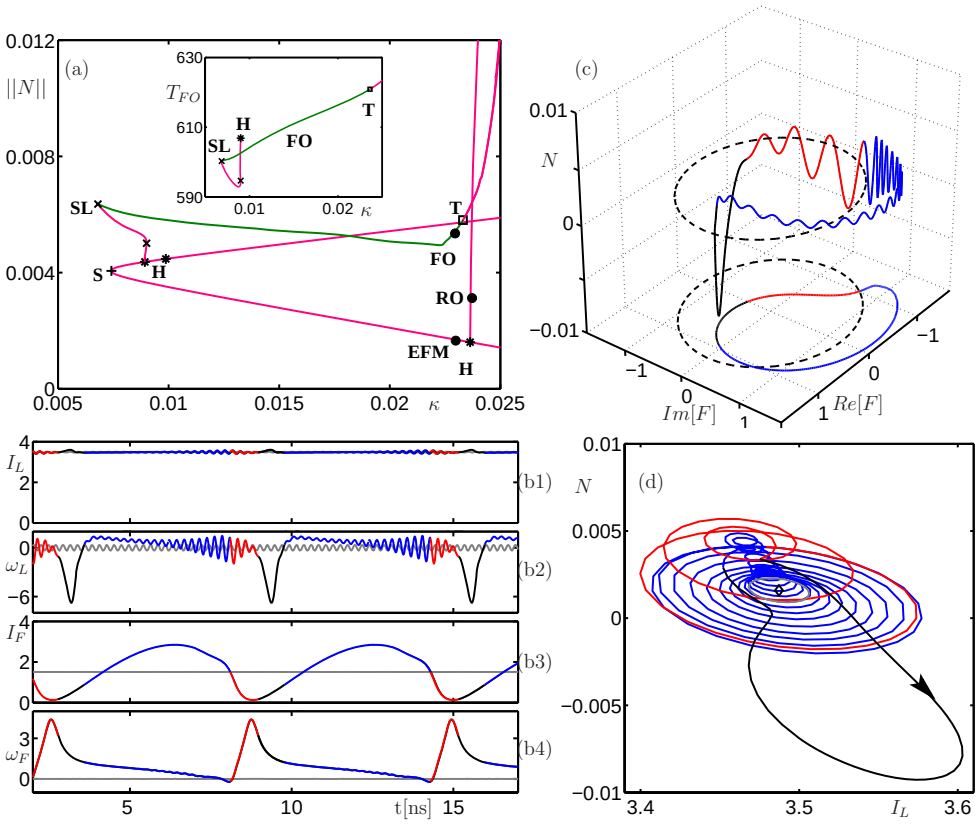


Figure 14. *Mixed FO/RO dynamics for* $(\Delta, C_p) = (0.0315, -4\pi)$. *Panel* (a) *shows a one-parameter bifurcation diagram in the parameter* κ. *The inset shows the period* T_{FO} *along the branch of FOs, which are stable between a saddle-node of limit cycle (SL) and a torus (T) bifurcation. Panel* (b) *shows time series of the stable periodic orbit for* $\kappa = 0.0226$ *(black dot in panel* (a) *near the end of the FO stability region). Along the blue part there is a build-up and along the red part a decay of fast oscillations, and the black part is a global excursion; for comparison an unstable RO (that appears for slightly higher* κ*) is shown as the gray time series in panels* (b). *Panel* (c) *shows the trajectory in the* (F, N)*-space and its projection on the* F*-plane; the dashed line in panel* (c) *indicates the unstable EFM. Panel* (d) *shows the projection on the* (I_L, N)*-plane. The diamond is the unstable EFM, and the gray circle is the trajectory of the unstable ROs.*

change.)

To explain the increase of the norm $||N||$ we note a Hopf bifurcation on the lower part of the EFM branch in Figure 14(a), which gives rise to unstable ROs. The corresponding EFM of saddle type is, therefore, characterized by leading complex conjugate unstable eigenvalues whose imaginary parts are close to the RO frequency. The idea is that, when it comes near this EFM, the FO "picks up" the additional RO frequency. Consequently, these oscillations of increasing amplitude make a substantial and increasing contribution to the norm $||N||$ along the FO branch, which leads to its increase.

To illustrate this further, Figure 14(b) shows time series of the FOs for parameter values identified by the black dot close to the torus bifurcation T in panel (a). Note again that the power is still almost constant, but additional fast oscillations are clearly seen in the frequency ω_L of the laser. That these are indeed on the scale of the ROs is demonstrated by the gray

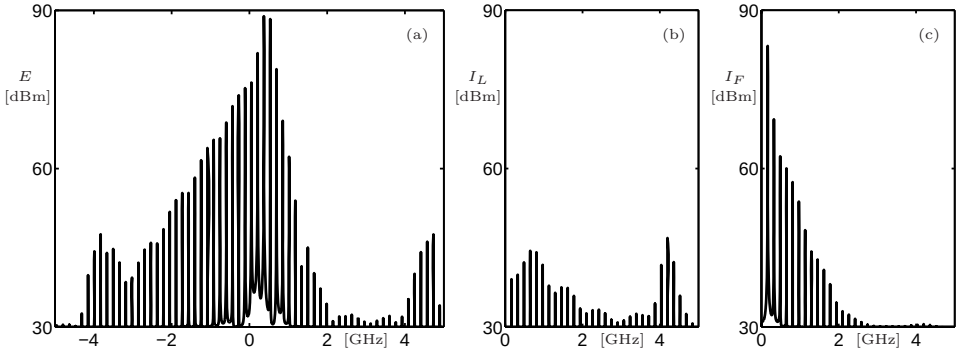


Figure 15. *Spectra of the time series from Figure 14. Panel* (a) *shows the spectrum of the laser field E, panel* (b) *the spectrum of the laser intensity I_L, and panel* (c) *the spectrum of the feedback intensity I_F. Notice that the relaxation oscillation component of the signal appears in the laser intensity but not in the feedback intensity.*

time series in panels (b). These are for parameter values identified by the black dot on the RO curve in panel (a). Moreover, the FO time series are split up into blue, red, and black parts to highlight different sections of the trajectory. During the blue interval a fast oscillation is building up, which decays again during the red interval. The build-up rate is noticeably slower than the decay rate. During the black interval the trajectory makes a large excursion in phase-space.

The FO periodic orbit is shown in Figure 14(c) in the (F, N)-space and in the projection onto the F-plane, and in Figure 14(d) in the projection onto the (I_L, N)-plane. The dotted circle in panel (c) indicates the position of the unstable EFM. (We remark that we fixed the phase of the periodic orbit, but still show the entire S^1 group orbit of the EFM for convenience.) Panel (c) shows that the change from build-up to decay of fast oscillations occurs when the trajectory is closest to the dotted circle representing the saddle EFM. This can be seen even better in the (I_L, N)-projection of Figure 14(d). Now the EFM is indicated by a diamond, and the orbit of the nearby unstable RO is shown in gray. The arrow indicates the direction along the periodic orbit.

A possible interpretation of the overall FO dynamics is the following. During the blue build-up of oscillations the trajectory also comes closer to the saddle EFM. After its closest approach to the EFM the trajectory leaves again, but with a faster rate. This indicates that this motion is along certain directions on the stable and unstable manifolds of the EFM that are characterized by complex conjugate eigenvalues. The frequency of the oscillations remains practically constant and of the order of the ROs, which is arguably because the EFM is close to a Hopf bifurcation of an unstable RO. The black part of the trajectory is a large excursion in phase space that leads to a reinjection into the vicinity of the EFM. It seems that the FO periodic orbit is close to a heteroclinic connection with the EFM, but such details of the dynamics are beyond the scope of this paper.

Figure 14 is a clear example of mixed FO/RO dynamics in the FOF laser system. To make the connection with the possible detection of this type of dynamics in an experiment, Figure 15 shows the relevant spectra of the dynamics as determined by Fourier transform from the numerical data. Figure 15(a) shows the optical spectrum of the laser field, which is the

Fourier transform of the complex-valued electrical field E. The sharp equidistant peaks are due to the FO frequency of about 161.4 MHz (and its higher harmonics). Peaks are noticeably enhanced around ± 4.2 GHz, which is due to the RO component of the signal.

To distinguish between the actual FO and RO dynamics it is best to consider the spectra of the laser intensity I_L and the filter intensity I_F as shown in Figure 15(b) and (c). These spectra are known as relative intensity noise (RIN) spectra, and they can be obtained experimentally by recording the light with a fast photo-diode and analyzing the signal with a spectrum analyzer. Notice that the RIN spectrum of I_F contains only the FO frequency and its harmonics and no RO component. The RIN spectrum of I_L, on the other hand, clearly shows the RO component at about 4 GHz. Notice also that the FO component in the laser field is weaker than that of the filter field by about 40 dBm. Overall, the example in Figure 15 shows that it should be possible to distinguish the dynamics presented here in terms of its FO and RO contents by means of recording RIN spectra of the laser intensity and the filter intensity.

6. Summary and outlook. We have investigated the bifurcation structure of a semiconductor laser with filtered optical feedback (FOF) as a function of three main parameters, namely the feedback rate κ, the feedback phase C_p, and the detuning Δ between the solitary laser frequency and the center frequency of the filter. The emphasis here was on stable external filtered modes and bifurcating stable oscillations, which can be either relaxation oscillations or frequency oscillations. This stability information was presented in two-parameter bifurcation diagrams in the (κ, C_p)-plane for different, representative values of Δ. Depending on the value of Δ, we found a single or two separate large regions of stable EFMs. Similarly, the detuning has a strong influence of the stability regions of ROs and FOs. Our results show clearly that the feedback phase C_p is a key parameter in the FOF laser. In other words, both the *magnitude* and the *phase* of the feedback must be considered to reveal a comprehensive picture of the dynamics.

Our theoretical study produced specific results concerning the two types of oscillations, which can be checked experimentally. Specifically, we expect that the regions of FOs presented here are large enough to be detected experimentally. Importantly, FOs can be found already for quite low values of κ. We demonstrated that even quite complicated dynamics, such as mixed FO/RO oscillations, show up clearly in the spectra of the laser and filter intensities. Therefore, we expect that careful sweeps in the feedback strength κ for fixed values of the feedback phase C_p are a means of detecting the regions of stability in the (κ, C_p)-plane as reported here. For Δ close to zero we predict two separate regions of stable FOs, namely, in-phase and antiphase dynamics between laser and filter, which occur at the two opposite flanks of the filter profile. For larger detunings Δ, on the other hand, there is only a single region of stable FOs, which is located in between the two separate EFM stability regions. Hence, determining experimentally the stability of FOs in the (κ, C_p)-plane for different values of Δ is a means of confirming the predicted dependence on the filter detuning. In fact, the bifurcation diagrams in Figures 11 and 12 form the basis for ongoing measurements of the FOF laser.

Our results were obtained for physically representative values of the parameters of a semiconductor laser with filtered optical feedback. The electron life-time T is effectively a constant for the class of semiconductor lasers, while the linewidth enhancement factor α may differ (in about the range $[1 \ldots 10]$) between semiconductor lasers. The results presented here are for

semiconductor lasers with a moderately large α as used in recent experiments; the exact value of $\alpha \approx 5$ should not play a significant role. Similarly, the pump rate P is not expected to have much influence as long as it stays well above the laser threshold, and moderate changes of the delay time τ chiefly change the number of EFMs. More interesting is the dependence of the FOF laser dynamics on the filter characteristic. We already considered the role of the filter detuning Δ. The filter width Λ is also important for the existence of FOs, because they depend on an interaction with the flanks of the filter profile. If the filter is too narrow, then the flank is too steep for FOs to develop; in fact, in the limit of an $\Lambda \to 0$ the FOF laser effectively becomes a laser with optical injection [5, 12, 18]. On the other hand, when the filter is too wide, then the effect of filtering is lost and the laser acts as a laser with conventional optical feedback [12, 19]. However, there appears to be quite a range of intermediate values of Λ where FOs do occur stably, and the fixed value $\Lambda = 0.007$ chosen here is representative in this respect. A study of how the stability regions of FOs are bounded for very small and very large Λ is a topic of ongoing theoretical and experimental research.

Another immediate question that arises when linking a theoretical bifurcation study with experiments is that of the influence of noise. The work reported in [36] shows that the FOF laser system may be very sensitive to noise. However, our study shows that the stability regions of EFMs and different types of oscillations are actually quite large. Furthermore, we consider the case of a laser that is pumped well above threshold, in which case the spontaneous emission noise of the laser is (relatively) small. We therefore expect that the stability regions reported here are sufficiently robust to noise. This is confirmed by the fact that EFMs, ROs, and FOs have been measured experimentally in the FOF laser (even though their stability regions have not been mapped out) [9, 10, 6]. On the other hand, the dynamics beyond stable ROs and FOs may be very complicated, and we expect noise to play a more significant role in this regime of the FOF laser.

We presented the bifurcation diagrams in the covering space of C_p. This is very convenient because, as one would do in an experiment, one can follow a particular solution, for example, the solitary laser solution for $\kappa = 0$, to the different stability boundaries. However, it does not convey the strong multistability of the FOF laser. We finish by giving an idea of the overall complexity of the FOF laser system in Figure 16, where all stability regions of EFMs, ROs, and FOs are shown on a fundamental 2π-interval of C_p. Effectively, every dynamical state can coexist with any other one in some region of the (κ, C_p)-plane. Figure 16 also demonstrates the advantage of bifurcation analysis of DDEs with numerical continuation tools over numerical simulation. By tracking individual solutions in the relevant parameters we were able to detect multistability reliably.

Laser systems with delay have been a class of motivating testbed examples for the recent development of computational tools for DDEs [22]. The bifurcation study presented here is much in this spirit in that it demonstrates the state-of-the-art of numerical continuation techniques for DDEs arising in applications. We hope that numerical methods of bifurcation analysis will find use in other areas of application where delays are important, such as control theory, ecological systems, and network dynamics, to name just a few.

Acknowledgment. We thank Kirk Green for helpful discussions and comments on a manuscript of this work.

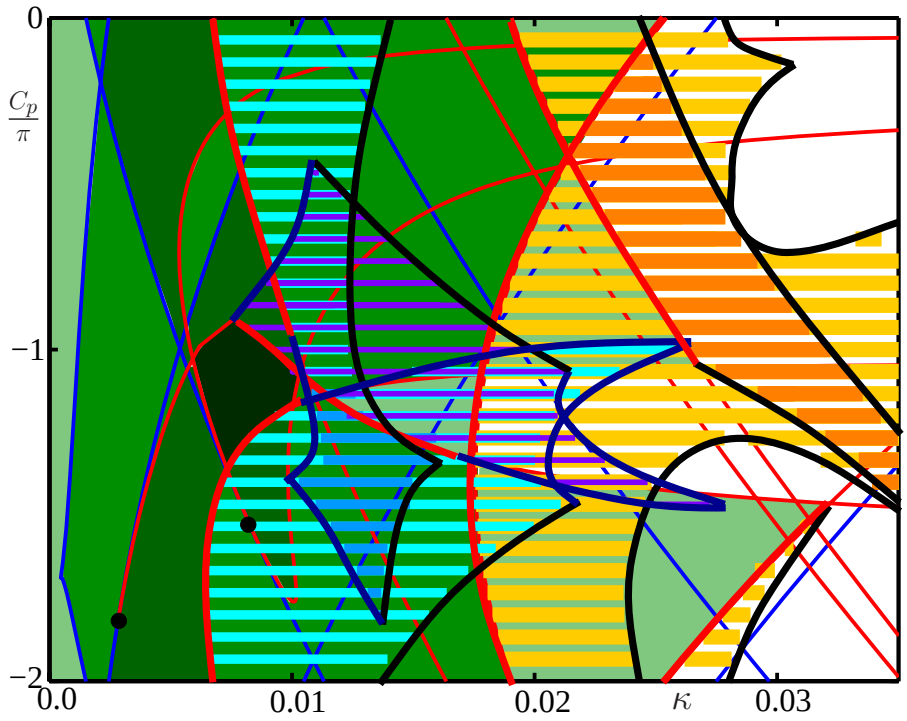


Figure 16. *Two-parameter bifurcation diagram in the (κ, C_p)-plane for $\Delta = -0.007$, where all stability regions of EFMs, ROs, and FOs are shown on a fundamental 2π-interval of C_p; compare with Figures* 4(a) *and* 10(a). *The image summarizes the large amount of multistability of the FOF laser; the multistability of EFMs is indicated by different shades of green, that of RO by different shades of orange, and that of FO by different shades of blue.*

REFERENCES

[1] D. H. DeTienne, G. R. Gray, G. P. Agrawal, and D. Lenstra, *Semiconductor laser dynamics for feedback from a finite-penetration-depth phase-conjugate mirror*, IEEE J. Quantum Electron., 33 (1997), pp. 838–844.

[2] O. Diekmann, S. A. V. Gils, and S. M. V. Lunel, *Delay Equations: Functional-, Complex-, and Nonlinear Analysis*, Springer-Verlag, New York, 1995.

[3] F. Dumortier, R. Roussarie, J. Sotomayor, and H. Zoladek, *Bifurcations of Planar Vector Fields. Nilpotent Singularities and Abelian Integrals*, Lecture Notes in Math. 1480, Springer-Verlag, Berlin, 1991.

[4] K. Engelborghs, T. Luzyanina, and G. Samaey, *DDE-BIFTOOL v.* 2.00 *User Manual: A Matlab Package for Bifurcation Analysis of Delay Differential Equations*, Technical Report TW-330, Department of Computer Science, K. U. Leuven, Leuven, 2001.

[5] T. Erneux, M. Yousefi, and D. Lenstra, *The injection laser limit of lasers subject to filtered optical feedback*, in Proceedings of the European Quantum Electronics Conference, 2003.

[6] H. Erzgräber, B. Krauskopf, D. Lenstra, A. P. A. Fischer, and G. Vemuri, *Frequency versus relaxation oscillations in a semiconductor laser with coherent filtered optical feedback*, Phys. Rev. E (3), 73 (2006), paper 055201(R).

[7] H. Erzgräber, D. Lenstra, B. Krauskopf, E. Wille, M. Peil, I. Fischer, and W. Elsässer, *Mutually delay-coupled semiconductor lasers: Mode bifurcation scenarios*, Opt. Commun., 255 (2005), pp. 286–296.

[8] B. Farias, T. P. de Silans, M. Chevrollier, and M. Oriá, *Frequency bistability of a semiconductor laser under a frequency-dependent feedback*, Phys. Rev. Lett., 94 (2005), paper 173902.
[9] A. P. A. Fischer, O. K. Andersen, M. Yousefi, S. Stolte, and D. Lenstra, *Experimental and theoretical study of filtered optical feedback in a semiconductor laser*, IEEE J. Quantum Electron., 36 (2000), pp. 375–384.
[10] A. P. A. Fischer, M. Yousefi, D. Lenstra, M. W. Carter, and G. Vemuri, *Experimental and theoretical study of semiconductor laser dynamics due to filtered optical feedback*, Phys. Rev. Lett., 92 (2004), pp. 23901–23904.
[11] K. Green and B. Krauskopf, *Bifurcation analysis of a semiconductor laser subject to non-instantaneous phase-conjugate feedback*, Opt. Commun., 231 (2004), pp. 383–393.
[12] K. Green and B. Krauskopf, *Mode structure of a semiconductor laser subject to filtered optical feedback*, Opt. Commun., 258 (2006), pp. 243–255.
[13] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1986.
[14] B. Haegeman, K. Engelborghs, D. Rose, D. Pieroux, and T. Erneux, *Stability and rupture of bifurcation bridges in semiconductor lasers subject to optical feedback*, Phys. Rev. E (3), 66 (2002), paper 046216.
[15] J. K. Hale and S. M. V. Lunel, *Introduction to Functional Differential Equations*, Springer-Verlag, New York, 1993.
[16] E. Hecht, *Optics*, Addison-Wesley, London, UK, 2002.
[17] T. Heil, I. Fischer, W. Elsässer, and A. Gavrielides, *Dynamics of semiconductor lasers subject to delayed optical feedback: The short cavity regime*, Phys. Rev. Lett., 87 (2001), paper 243901.
[18] G. Hek and V. Rottschäfer, *Semiconductor laser with filtered optical feedback: Bridge between conventional feedback and optical injection*, in Proceedings of ENOC, 2005.
[19] G. Hek and V. Rottschäfer, *Semiconductor laser with filtered optical feedback: From optical injection to conventional feedback*, submitted.
[20] E. Jaatinen, *Single frequency and tunable single frequency semiconductor laser systems*, in Unlocking Dynamical Diversity: Optical Feedback Effects on Semiconductor Lasers, D. Kane and K. Shore, eds., Wiley, New York, 2005, pp. 257–305.
[21] D. M. Kane and K. A. Shore, eds., *Unlocking Dynamical Diversity: Optical Feedback Effects on Semiconductor Lasers*, Wiley, New York, 2005.
[22] B. Krauskopf, *Bifurcation analysis of lasers with delay*, in Unlocking Dynamical Diversity: Optical Feedback Effects on Semiconductor Lasers, D. Kane and K. Shore, eds., Wiley, New York, 2005, pp. 147–183.
[23] B. Krauskopf and D. Lenstra, eds., *Fundamental Issues of Nonlinear Laser Dynamics*, AIP Conference Proceedings 548, Melville, New York, 2000.
[24] B. Krauskopf and C. Rousseau, *Codimension-three unfoldings of reflectionally symmetric planar vector fields*, Nonlinearity, 10 (1997), pp. 1115–1150.
[25] B. Krauskopf, G. V. Tartwijk, and G. Gray, *Symmetry properties of lasers subject to optical feedback*, Opt. Commun., 177 (2000), pp. 347–353.
[26] Yu. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag, New York, 1995.
[27] L. Larger, P.-A. Lacourt, S. Poinsot, and M. Hanna, *From flow to map in an experimental high-dimensional electro-optic nonlinear delay oscillator*, Phys. Rev. Lett., 95 (2005), paper 043903.
[28] D. Lenstra and M. Yousefi, *Theory of delayed optical feedback in lasers*, in Fundamental Issues of Nonlinear Laser Dynamics, AIP Conference Proceedings 548, B. Krauskopf and D. Lenstra, eds., AIP-American Institute of Physics, New York, 2000, pp. 66–86.
[29] M. Nizette and T. Erneux, *Optical frequency dynamics and relaxation oscillations of a semiconductor laser subject to filtered optical feedback*, in Semiconductor Lasers and Laser Dynamics II, Proceedings of SPIE, D. Lenstra, M., Pessa, and I. H. White, eds., SPIE, Bellingham, WA, 2006, pp. 6184–32.
[30] F. Rogister, P. Mégret, O. Deparis, M. Blondel, and T. Erneux, *Suppression of low frequency fluctuations and stabilization of a semiconductor subject to optical feedback from a double cavity: Theoretical results*, Opt. Lett., 24 (1999), pp. 1218–1220.
[31] V. Rottschäfer and B. Krauskopf, *The ECM-backbone of the Lang-Kobayashi equations: A geometric picture*, Int. J. Bif. Chaos, in press.

[32] R. Szalai, *PDDE-CONT: A Continuation and Bifurcation Software for Delay-Differential Equations*, Department of Applied Mechanics, Budapest University of Technology and Economics, Budapest, Hungary, 2005, http://www.mm.bme.hu/~szalai/pdde.

[33] V. Z. Tronciu, H.-J. Wünsche, M. Radziunas, and M. Wolfrum, *Semiconductor lasers under resonant feedback from a Fabry–Perot: Stability of continuous wave operation*, Phys. Rev. E (3), 73 (2006), paper 046205.

[34] S. M. Wieczorek, B. Krauskopf, T. B. Simpson, and D. Lenstra, *The dynamical complexity of optically injected semiconductor lasers*, Phys. Reports, 416 (2005), pp. 1–128.

[35] A. Yariv, *Quantum Electronics*, Wiley, New York, 1988.

[36] M. Yousefi, D. Lenstra, and G. Vemuri, *Nonlinear dynamics of a semiconductor laser with filtered optical feedback and the influence of noise*, Phys. Rev. E (3), 67 (2003), paper 046213.

[37] M. Yousefi, D. Lenstra, G. Vemuri, and A. P. A. Fischer, *Control of nonlinear dynamics of a semiconductor laser with filtered optical feedback*, in Proceedings of the Conference on Optoelectronics, Vol. 148, 2001, pp. 223–237.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 1, pp. 29–60

An Efficient Method for Studying Weak Resonant Double Hopf Bifurcation in Nonlinear Systems with Delayed Feedbacks*

Jian Xu†, Kwok-Wai Chung‡, and Chuen-Lit Chan‡

Abstract. An efficient method, called the perturbation-incremental scheme (PIS), is proposed to study, both qualitatively and quantitatively, the delay-induced weak or high-order resonant double Hopf bifurcation and the dynamics arising from the bifurcation of nonlinear systems with delayed feedback. The scheme is described in two steps, namely, the perturbation and the incremental steps, when the time delay and the feedback gain are taken as the bifurcation parameters. As for applications, the method is employed to investigate the delay-induced weak resonant double Hopf bifurcation and dynamics in the van der Pol–Duffing and the Stuart–Landau systems with delayed feedback. For bifurcation parameters close to a double Hopf point, all solutions arising from the resonant bifurcation are classified qualitatively and expressed approximately in a closed form by the perturbation step of the PIS. Although the analytical expression may not be accurate enough for bifurcation parameters away from the double Hopf point, it is used as an initial guess for the incremental step which updates the approximate expression iteratively and performs parametric continuation. The analytical predictions on the two systems show that the delayed feedback can, on the one hand, drive a periodic solution into an amplitude death island where the motion vanishes and, on the other hand, create complex dynamics such as quasi-periodic and coexisting motions. The approximate expression of periodic solutions with parameter varying far away from the double Hopf point can be calculated to any desired accuracy by the incremental step. The validity of the results is shown by their consistency with numerical simulations. We show that as an analytical tool the PIS is simple but efficient.

Key words. delay differential system, double Hopf bifurcation, delayed feedback control, nonlinear dynamics

AMS subject classifications. 34K06, 37H20, 70K30, 74G10, 93B52

DOI. 10.1137/040614207

1. Introduction.

Time delay is ubiquitous in many physical systems. In particular, delay differential equations (DDEs) are used to model various phenomena, such as neural [1, 2], ecological [3], biological [4], mechanical [5, 6, 7, 8], controlling chaos [9], secure communication via chaotic synchronization [10, 11], and other natural systems due to finite propagation speeds of signals, finite reaction times, and finite processing times [12]. This research shows that the time delay in various systems has not only a quantitative but also a qualitative effect on dynamics even for a small time delay [13]. Therefore, the investigation of the mechanism of how the delay induces various dynamics of a system becomes important.

*Received by the editors August 31, 2004; accepted for publication (in revised form) by S. van Gils September 29, 2006; published electronically February 9, 2007. This work was supported by the strategic research grant 7001338 of the City University of Hong Kong and National Nature and Science Foundation of China under grants 10625211 and 10532050.

http://www.siam.org/journals/siads/6-1/61420.html

†School of Aerospace Engineering and Applied Mechanics, Tongji University, Shanghai 200092, P. R. China (xujian@mail.tongji.edu.cn).

‡Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong (makchung@cityu.edu.hk, mapeter@cityu.edu.hk).

The qualitative and quantitative theories for DDEs have been well developed recently. Many methods and techniques of geometric theory of dynamical systems on ordinary differential equations (ODEs) have been extended to investigate DDEs, such as stability analysis, bifurcation theory, and perturbation techniques. In stability analysis, Kalas and Baráková [14] set up a theoretical base for the stability and asymptotic behavior of a two-dimensional DDE. Using a suitable Lyapunov–Krasovskii functional, they transformed a real two-dimensional system into a single equation with complex coefficients and proved the stability and asymptotic behavior of the system under consideration. Gopalsamy and Leung [15] used the perturbation method to study the stability of periodic solutions in DDEs. On the aspect of the bifurcation theory of DDEs, the center manifold of a flow is usually adopted to reduce DDEs to finite-dimensional systems. Redmond, LeBlanc, and Longtin [16] completely studied the bifurcations of the trivial equilibrium and computed the normal form coefficients of the reduced vector field on the center manifold. The analysis reveals a Hopf bifurcation curve terminating on a pitchfork bifurcation line at a codimension-2 Takens–Bogdanov point in the parameter space. Campbell and LeBlanc [17] used center manifold analysis to investigate a 1:2 resonant double Hopf bifurcation in a DDE. It is possible to find period doubling bifurcations near such a point for a system without the time delay [18]. Correspondingly, the perturbation method is also extended to determine the analytical solutions of DDEs arising from bifurcations. Instead of center manifold reduction (CMR), Das and Chatterjee [19] employed the method of multiple scales (MMS) to obtain analytical solutions close to Hopf bifurcation points for DDEs. In applications, the main steps of the analysis are schemed as follows [20]:

a. consider an equilibrium at a critical parameter;
b. solve the eigenvalue problem for this equilibrium to find its linear stability;
c. localize the critical point in parameter space where bifurcation occurs;
d. calculate the eigenvalues and eigenfunctions at the bifurcation point and reduce DDEs on the center manifold;
e. compute the appropriate normal form coefficients.

Some of the literature mentioned above deals with controlled systems with delayed feedbacks. In our previous research [21, 22], it was found that controlled systems with delayed feedbacks may undergo a double Hopf bifurcation with the time delay and feedback gain varying. However, the mechanism has not been explained in detail. Such a phenomenon has also been observed by Reddy, Sen, and Johnston [12] and Campbell et al. [20]. Recently, Buono and Bélair [23] employed the methods developed by Faria and Magalhães [24] to investigate the normal form and universal unfolding of a vector field at nonresonant double Hopf bifurcation points for particular classes of retarded functional differential equations (RFDEs). They represented restrictions on the possible flows on a center manifold for certain singularities. Buono and LeBlanc [25] extended Arnold's theory on universal unfolding of matrices to the case of parameter-dependent linear RFDEs. In studying delay systems governed by DDEs, many authors recommend a CMR as the first step due to the fact that DDEs and RFDEs are infinite-dimensional. However, the CMR has its disadvantages. On the one hand, computation of the CMR with normal form is very tedious for a codimension-2 bifurcation. On the other hand, the CMR is invalid for values of the bifurcation parameters far away from the bifurcation point. This constitutes the motivation of the present paper.

Our goal is to propose a simple but efficient method which not only inherits the advan-

tages of the CMR and MMS, but also overcomes the disadvantages of them, to explain the mechanism of the delay-induced double Hopf bifurcation. To this end, this paper is focused on controlled systems with delayed feedbacks governed by a type of two-dimensional DDE and proposes a method called the *perturbation-incremental scheme* (PIS). The scheme is described in two steps, namely, the perturbation and the incremental steps, when the time delay and the feedback gain are taken as the bifurcation parameters. The perturbation step of the PIS provides not only accurate qualitative prediction and classification but also an analytical expression with high accuracy for both periodic and quasi-periodic motions when bifurcation parameters are closed to the weak resonant point. For those values of bifurcation parameters far away from the resonant point, the analytical expression may not be accurate enough, but it can be considered as an initial guess in the incremental step. Thus this overcomes the disadvantage of the incremental harmonic balance (IHB) method. From the incremental step, the approximate analytical expression is updated iteratively and can reach any desired accuracy. Parametric continuation can also be performed by using the PIS. As for applications of the method, the delay-induced weak resonant double Hopf bifurcation and dynamics in the van der Pol–Duffing and the Stuart–Landau systems with delayed feedback are investigated. The analytical predictions show that the delayed feedback can lead not only to the vanishing of periodic motion in an "amplitude death island" but also to more complex dynamics such as quasi-periodic and coexisting motions. The validity of the results is shown by their consistency with numerical simulations. As an analytical tool, the PIS is simple but efficient. The approximate expression obtained from the PIS can achieve any required accuracy which is not possible by the CMR and other perturbation methods.

It should be noted that this work is also an extension of our previous studies [22, 26, 27, 28].

The paper is organized as follows. In section 2, we discuss in detail the corresponding linearized systems when two pairs of purely imaginary eigenvalues occur at a critical value of time delay, giving rise to weak resonant double Hopf bifurcations. In section 3, a new method, called the perturbation-incremental scheme (PIS), is proposed in two steps to investigate harmonic solutions derived from weak resonant double Hopf bifurcation. As illustrative examples, the van der Pol–Duffing and the Stuart–Landau systems are considered in sections 4 and 5, where the analytical and numerical results are compared. Finally, we close the paper with a summary of our results.

2. Weak resonant double Hopf bifurcation. We consider two first-order DDEs with linear delayed feedback and nonlinearities in the general form

$$\dot{Z}(t) = C\,Z(t) + D\,Z(t-\tau) + \epsilon\,F(\,Z(t)\,,Z(t-\tau)\,), \tag{2.1}$$

where $Z(t) = \{x(t), y(t)\}^T \in R^2$, C and D are 2×2 real constant matrices such that $C = \left[\begin{smallmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{smallmatrix}\right]$ and $D = \left[\begin{smallmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{smallmatrix}\right]$, F is a nonlinear function in its variable with $F(0,0) = 0$, ϵ is a parameter representing strength of nonlinearities, and τ is the time delay.

Now we derive some formulae relating to resonant double Hopf bifurcation points. It can be seen from (2.1) that $Z = 0$ is always an equilibrium point or trivial solution of the system. To determine the stability of the trivial solution for $\tau \neq 0$, we linearize system (2.1) around $Z = 0$ to obtain the characteristic equation

$$\det(\lambda\,I - C - D\,e^{-\lambda\,\tau}) = 0, \tag{2.2}$$

where I is the identity matrix and the perturbation is assumed to have a time dependence proportional to $e^{\lambda\tau}$. The characteristic equation (2.2) can be rewritten in the form

$$\lambda^2 - \lambda\left(c_1 + d_1\, e^{-\lambda\tau}\right) + c_d\, e^{-\lambda\tau} + c_2 + \det(D)\, e^{-2\lambda\tau} = 0, \tag{2.3}$$

where

$$\begin{aligned} c_1 &= c_{11} + c_{22}, \\ d_1 &= d_{11} + d_{22}, \\ c_2 &= c_{11}\, c_{22} - c_{12}\, c_{21}, \\ c_d &= c_{22}\, d_{11} - c_{21}\, d_{12} - c_{12}\, d_{21} + c_{11}\, d_{22}. \end{aligned} \tag{2.4}$$

The roots of the characteristic equation (2.3) are commonly called the eigenvalues of the equilibrium point of system (2.1). The stability of the trivial equilibrium point will change when the system under consideration has zero or a pair of imaginary eigenvalues. The former occurs if $\lambda = 0$ in (2.3) or $c_d + c_2 + \det(D) = 0$, which can lead to the static bifurcation of the equilibrium points such that the number of equilibrium points changes when the bifurcation parameters vary. The latter deals with the Hopf bifurcation such that the dynamical behavior of the system changes from a static stable state to a periodic motion or vice versa. The dynamics become quite complicated when the system has two pairs of pure imaginary eigenvalues at a critical value of time delay. We will concentrate on such cases. For this, we let $c_d + c_2 + \det(D) \neq 0$. Thus, $\lambda = 0$ is not a root of the characteristic equation (2.3) in the present paper. Such an assumption can be realized in engineering as long as one chooses a suitable feedback controller.

It is easy to find explicit expressions for the critical stability boundaries of the following two cases:

a. $\det(D) = 0$, $c_d + c_2 \neq 0$;

b. $\det(D) \neq 0$, $c_{11} = c_{22}$, $c_{12} = -c_{21}$, $d_{11} = d_{22}$, $d_{12} = d_{21} = 0$, $c_{11} > 0$, $c_{12} > 0$, $d_{11} \neq 0$.

For the case with $\det(D) = 0$ but $c_d + c_2 \neq 0$, substituting $\lambda = a + i\omega$ into (2.3) and equating the real and imaginary parts to zero yield

$$\begin{aligned} a^2 - \omega^2 - a\, c_1 + c_2 - e^{-a\tau}\,\omega \sin(\tau\,\omega)\, d_1 + e^{-a\tau}\cos(\tau\,\omega)\,(c_d - a\, d_1) = 0, \\ 2\, a\,\omega - \omega\, c_1 - e^{-a\tau}\,\omega\cos(\tau\,\omega)\, d_1 + e^{-a\tau}\sin(\tau\,\omega)\,(a\, d_1 - c_d) = 0. \end{aligned} \tag{2.5}$$

One can derive the explicit expressions for the critical stability boundaries by setting $a = 0$ in (2.5) and obtain

$$\begin{aligned} -\omega^2 + c_2 + c_d\cos(\tau\,\omega) - \omega\, d_1 \sin(\tau\,\omega) = 0, \\ -\omega\, c_1 - c_d \sin(\tau\,\omega) - \omega\, d_1\cos(\tau\,\omega) = 0. \end{aligned} \tag{2.6}$$

Eliminating τ from (2.6), we have

$$\omega_{\pm} = \frac{\sqrt{{d_1}^2 - {c_1}^2 + 2\, c_2 \pm \sqrt{\left({d_1}^2 - {c_1}^2 + 2\, c_2\right)^2 - 4\,({c_2}^2 - {c_d}^2)}}}{\sqrt{2}} \tag{2.7}$$

when the following conditions hold:

$$\begin{aligned} & c_2^2 - c_d^2 > 0, \\ & \left(d_1{}^2 - c_1{}^2 + 2\,c_2\right)^2 > 4\left(c_2{}^2 - c_d{}^2\right). \end{aligned} \tag{2.8}$$

Then, two families of surfaces, denoted by τ_- and τ_+ in terms of c_d and d_1 corresponding to ω_- and ω_+, respectively, can be derived from (2.5) and be given by

$$\begin{aligned} \cos(\,\omega_-\tau_-) &= \frac{\omega_-^2\,c_d - c_2\,c_d - \omega_-^2\,c_1\,d_1}{c_d{}^2 + \omega_-^2\,d_1{}^2}, \\ \cos(\,\omega_+\tau_+) &= \frac{\omega_+^2\,c_d - c_2\,c_d - \omega_+^2\,c_1\,d_1}{c_d{}^2 + \omega_+^2\,d_1{}^2}. \end{aligned} \tag{2.9}$$

It should be noted that $\omega_- < \omega_+$. Thus, a possible double Hopf bifurcation point occurs when two such families of surfaces intersect each other where

$$\tau_- = \tau_+. \tag{2.10}$$

Equation (2.10) not only determines the linearized system around the trivial equilibrium which has two pairs of pure imaginary eigenvalues $\pm i\omega_-$ and $\pm i\omega_+$, but also gives a relation between ω_- and ω_+. If

$$\omega_- : \omega_+ = k_1 : k_2, \tag{2.11}$$

then a possible double Hopf bifurcation point appears with frequencies in the ratio $k_1 : k_2$. If $k_1, k_2 \in Z^+$, $k_1 < k_2$, $k \neq 1$, and $k_2 \neq 1$, then such a point is called the $k_1 : k_2$ weak or no low-order resonant double Hopf bifurcation point. Equations (2.10) and (2.11) form the necessary conditions for the occurrence of a resonant double Hopf bifurcation point. Equation (2.11) yields

$$d_1{}^2 = c_1{}^2 - 2\,c_2 + \frac{k_1{}^2 + k_2{}^2}{k_1\,k_2}\sqrt{c_2{}^2 - c_d{}^2} \tag{2.12}$$

if conditions (2.8) are satisfied. Substituting (2.12) into (2.7), one can obtain the frequencies in the simple expressions given by

$$\omega_- = \sqrt{\frac{k_1}{k_2}\sqrt{c_2{}^2 - c_d{}^2}}, \qquad \omega_+ = \sqrt{\frac{k_2}{k_1}\sqrt{c_2{}^2 - c_d{}^2}}. \tag{2.13}$$

The other parameters can be determined by (2.10) or the equation

$$\begin{aligned} & \arccos\left(\frac{-(c_2\,c_d\,k_2) + \sqrt{c_2{}^2 - c_d{}^2}\,(c_d - c_1\,d_1)\,k_1}{c_d{}^2\,k_2 + \sqrt{c_2{}^2 - c_d{}^2}\,d_1{}^2\,k_1}\right) \\ & \quad = \frac{k_1}{k_2}\arccos\left(\frac{-(c_2\,c_d\,k_1) + \sqrt{c_2{}^2 - c_d{}^2}\,(c_d - c_1\,d_1)\,k_2}{c_d{}^2\,k_1 + \sqrt{c_2{}^2 - c_d{}^2}\,d_1{}^2\,k_2}\right), \end{aligned} \tag{2.14}$$

where d_1 is given in (2.12). The corresponding value of the time delay at the resonant double Hopf bifurcation point is given by

$$\begin{aligned}\tau_c &= \tau_- = \tau_+ \\ &= \sqrt{\frac{k_1}{k_2\sqrt{{c_2}^2-{c_d}^2}}}\arccos\left(\frac{-(c_2\,c_d\,k_1)+\sqrt{{c_2}^2-{c_d}^2}\,(c_d-c_1\,d_1)\,k_2}{{c_d}^2\,k_1+\sqrt{{c_2}^2-{c_d}^2}\,{d_1}^2\,k_2}\right).\end{aligned} \tag{2.15}$$

For the case with $\det(D)\neq 0$ and $c_{11}=c_{22}$, $c_{12}=-c_{21}$, $d_{11}=d_{22}$, $d_{12}=d_{21}=0$, $c_{11}>0$, $c_{12}>0$, $d_{11}<0$, the characteristic equation (2.2) becomes

$$\lambda = c_{11}\pm ic_{12}+d_{11}e^{-\lambda\tau}. \tag{2.16}$$

By substituting $\lambda=a+i\omega$ in (2.16), we get

$$\begin{aligned}&\omega=\omega_\pm=c_{12}\pm\sqrt{d_{11}^2\,e^{-2a\tau}-(a-c_{11})^2},\\ &a=c_{11}-\tfrac{\omega-c_{21}}{\tan(\omega\tau)},\end{aligned} \tag{2.17}$$

where we consider only one set of curves by choosing $\omega=c_{12}\pm\sqrt{\cdot}$. The other set of curves arising due to $\omega=-c_{12}\pm\sqrt{\cdot}$ is implicit in the above since the eigenvalues always occur in complex conjugate pairs. It follows from (2.17) that ω is real only when $d_{11}^2\geq(a-c_{11})^2\,e^{2a\tau}$. To obtain the critical boundary, set $a=0$. This gives

$$\begin{aligned}\omega_- &= c_{12}-\sqrt{{d_{11}}^2-{c_{11}}^2},\\ \omega_+ &= c_{12}+\sqrt{{d_{11}}^2-c_{11}^2},\end{aligned} \tag{2.18}$$

and

$$\begin{aligned}\tau_-[j] &= \tfrac{1}{\omega_-}\left(2j\pi-\cos^{-1}(-c_{11}/d_{11})\right),\\ \tau_+[j] &= \tfrac{1}{\omega_+}\left(2j\pi+\cos^{-1}(-c_{11}/d_{11})\right),\end{aligned} \tag{2.19}$$

where $j=1,2,\ldots$. The necessary conditions for the $k_1:k_2$ occurrence of the resonant double Hopf bifurcation can be obtained by setting $\tau_-[j]=\tau_+[j]$ and $\frac{\omega_-}{\omega_+}=\frac{k_1}{k_2}$. This yields

$$(d_{11})_c=-\frac{c_{11}}{\cos(2j\pi(k_2-k_1)/(k_1+k_2))},\quad \tau_c=\frac{4jk_1\pi}{(k_1+k_2)\omega_-}. \tag{2.20}$$

Finally, it should be noted that the parameters d_1 and c_d cannot be solved in a closed form from (2.15) due to the trigonometric function. However, the values can be obtained numerically. Such parameters denoted by $(\cdot)_c$ are called the critical values at the resonant double Hopf bifurcation point with frequencies in the ratio $k_1:k_2$. Thus, for given physical parameters τ and D in (2.1), one can obtain

$$\epsilon\,\tau_\epsilon=\tau-\tau_c,\qquad \epsilon\,D_\epsilon=D-D_c, \tag{2.21}$$

such that (2.1) can be rewritten as

$$\dot{Z} = C\,Z + D_c\,Z_{\tau_c} + \widetilde{F}(Z, Z_{\tau_c}, Z_{\tau_c+\epsilon\tau_\epsilon}, \epsilon), \tag{2.22}$$

where $Z = Z(t)$, $Z_\tau = Z(t-\tau)$, τ_c is given by (2.15), and

$$\widetilde{F}(\cdot) = D_c\,[Z_{\tau_c+\epsilon\tau_\epsilon} - Z_{\tau_c}] + \epsilon\,[D_\epsilon Z_{\tau_c+\epsilon\tau_\epsilon} + F(Z, Z_{\tau_c+\epsilon\tau_\epsilon})]\,. \tag{2.23}$$

Clearly, $\widetilde{F} = 0$ for $\epsilon = 0$ and the resonant double Hopf bifurcation point may occur in the system when $\epsilon = 0$.

3. Perturbation-incremental scheme. Various harmonic solutions with distinct topological structures can occur in a system due to nonresonant and high-order double Hopf bifurcation [32]. When the time delay is absent, the perturbation method [5] can be applied directly to (2.22) for small ϵ, and the IHB method [7] to the system for large ϵ. The key problem of IHB is to find an initial value and in general it is quite difficult. The perturbation-incremental method proposed in [26, 27] can efficiently overcome this disadvantage of the IHB method. However, the above methods must be re-examined and extended to investigate the system with time delay.

In this section, we propose a new method, called the perturbation-incremental scheme (PIS), to investigate harmonic solutions derived from the weak resonant double Hopf bifurcation of (2.1). Our goal is to obtain the harmonic solutions with any desired accuracy and consider the continuation of these solutions when the time delay and feedback gain in (2.1) are taken as bifurcation parameters. The scheme is described in two steps, namely, the perturbation step (noted as step one) for bifurcation parameters close to the weak resonant point and the incremental step (noted as step two) for those far away from the bifurcation point.

3.1. Perturbation step of PIS. It can be seen from the previous section that a double Hopf bifurcation with k_1:k_2 weak resonance occurs in (2.1) at $\tau = \tau_c$ and $D = D_c$. If τ and D are considered as two bifurcation parameters, then (D_c, τ_c) is a double Hopf bifurcation point with weak resonance. In this subsection, we derive the analytic expression of harmonic solutions arising from the weak resonant double Hopf bifurcation in (2.1) or (2.22) when τ and D are close to τ_c and D_c, respectively.

For $\epsilon = 0$, it can be seen from (2.21) and (2.23) that $\tau = \tau_c$, $D = D_c$, and $\widetilde{F}(\cdot) = 0$. Thus, the solution of (2.22) may be supposed as

$$Z_0(t) = \sum_{i=1}^{2} \left[\begin{Bmatrix} a_{k_i} \\ c_{k_i} \end{Bmatrix} \cos(k_i\,\phi) + \begin{Bmatrix} b_{k_i} \\ d_{k_i} \end{Bmatrix} \sin(k_i\,\phi) \right], \tag{3.1}$$

which results in

$$Z_0(t-\tau_c) = \sum_{i=1}^{2} \left[\begin{Bmatrix} a_{k_i} \\ c_{k_i} \end{Bmatrix} \cos(k_i\,\phi - k_i\omega\tau_c) + \begin{Bmatrix} b_{k_i} \\ d_{k_i} \end{Bmatrix} \sin(k_i\,\phi - k_i\omega\tau_c) \right], \tag{3.2}$$

where τ_c is given by (2.15), $\phi = \omega t$, and $\omega = \frac{\omega_-}{k_1} = \frac{\omega_+}{k_2}$ is determined by (2.13). Substituting (3.1) and (3.2) into (2.22) for $\epsilon = 0$ and using the harmonic balance, one obtains that

$$M_{k_i} \begin{Bmatrix} b_{k_i} \\ d_{k_i} \end{Bmatrix} = N_{k_i} \begin{Bmatrix} a_{k_i} \\ c_{k_i} \end{Bmatrix} \tag{3.3}$$

and

$$-M_{k_i}\begin{Bmatrix} a_{k_i} \\ c_{k_i} \end{Bmatrix} = N_{k_i}\begin{Bmatrix} b_{k_i} \\ d_{k_i} \end{Bmatrix}, \tag{3.4}$$

where $M_{k_i} = k_i\omega I + D_c \sin(k_i\omega\tau)$, $N_{k_i} = C + D_c\cos(k_i\omega\tau)$, and I is the 2×2 identity matrix. The equations (3.3) and (3.4) are in fact identical. Let $\begin{Bmatrix} \widetilde{a}_{k_i} \\ \widetilde{b}_{k_i} \end{Bmatrix} = \frac{N_{k_i}}{\det(N_{k_i})}\begin{Bmatrix} a_{k_i} \\ b_{k_i} \end{Bmatrix}$ and note that $\det(N_{k_i}) = \det(M_{k_i})$. It follows from (3.3) that the harmonic solution of (2.22) for $\epsilon = 0$ is given by

$$Z_0(t) = \sum_{i=1}^{2}\left[\widetilde{N}_{k_i}\cos(k_i\,\phi) + \widetilde{M}_{k_i}\sin(k_i\,\phi)\right]\begin{Bmatrix} \widetilde{a}_{k_i} \\ \widetilde{b}_{k_i} \end{Bmatrix} \tag{3.5}$$

with $\widetilde{N}_{k_i} = N_{k_i}^{-1}\det(N_{k_i})$ and $\widetilde{M}_{k_i} = M_{k_i}^{-1}\det(M_{k_i})$ for $i = 1, 2$.

Based on the expression in (3.5), we now consider the solution of (2.22) for small ϵ. The harmonic solution of (2.22) can be considered to be a perturbation to that of (3.5), given by

$$Z(t) = \sum_{i=1}^{2}\left[\widetilde{N}_{k_i}\cos(k_i\,\omega t + \epsilon\sigma_i t) + \widetilde{M}_{k_i}\sin(k_i\,\omega t + \epsilon\sigma_i t)\right]\begin{Bmatrix} a_{k_i}(\epsilon) \\ b_{k_i}(\epsilon) \end{Bmatrix}, \tag{3.6}$$

where $a_{k_i}(0) = \widetilde{a}_{k_i}$, $b_{k_i}(0) = \widetilde{b}_{k_i}$, and σ_1 and σ_2 are detuning parameters. The following theorem provides a new method to determine $\begin{Bmatrix} a_{k_i}(\epsilon) \\ b_{k_i}(\epsilon) \end{Bmatrix}$ in (3.6).

Theorem. *If $W(t)$ is a periodic solution of the equation*

$$\dot{W}(t) = -C^T\,W(t) - {D_c}^T\,W(t+\tau_c), \tag{3.7}$$

and $W(t) = W(t + 2\pi/\omega)$, then

$$\begin{aligned} &\textstyle\int_{-\tau_c}^{0}\left[{D_c}^T\,W(t+\tau_c)\right]^T\left[Z(t) - Z(t+2\pi/\omega)\right]dt \\ &-\left[W(0)\right]^T\left[Z(2\pi/\omega) - Z(0)\right] + \textstyle\int_0^{2\pi/\omega}\left[W(t)\right]^T\widetilde{F}(Z, Z_{\tau_c}, Z_{\tau_c+\epsilon\tau_\epsilon}, \epsilon)\,dt = 0. \end{aligned} \tag{3.8}$$

Proof. Multiplying both sides of (2.22) by $[W(t)]^T$ and integrating with respect to t from zero to $2\pi/\omega$, one has

$$\begin{aligned} &\textstyle\int_0^{2\pi/\omega}\left[W(t)\right]^T\dot{Z}(t)dt \\ &= \textstyle\int_0^{2\pi/\omega}\left[W(t)\right]^T\left[C\,Z(t) + D_c\,Z(t-\tau_c) + \widetilde{F}(Z, Z_{\tau_c}, Z_{\tau_c+\epsilon\tau_\epsilon}, \epsilon)\right]dt, \end{aligned} \tag{3.9}$$

where $2\pi/\omega$ is a period of $W(t)$ in t. The equation (3.9) yields

$$\begin{aligned} &\textstyle\int_0^{2\pi/\omega}\left[\dot{W}(t) + C^T\,W(t) + {D_c}^T\,W(t+\tau_c)\right]^T Z(t)\,dt \\ &+\textstyle\int_{-\tau_c}^{0}\left[{D_c}^T\,W(t+\tau_c)\right]^T\left[Z(t) - Z(t+2\pi/\omega)\right]dt \\ &-\left[W(0)\right]^T\left[Z(2\pi/\omega) - Z(0)\right] + \textstyle\int_0^{2\pi/\omega}\left[W(t)\right]^T\widetilde{F}(Z, Z_{\tau_c}, Z_{\tau_c+\epsilon\tau_\epsilon}, \epsilon)\,dt = 0. \end{aligned} \tag{3.10}$$

The theorem follows from (3.7).

To apply the theorem to determine $\left\{ \begin{smallmatrix} a_{k_i}(\epsilon) \\ b_{k_i}(\epsilon) \end{smallmatrix} \right\}$, one must obtain the expression of $W(t)$ in (3.10).

It is easily seen that the periodic solution of (3.7) can been written as

$$W(t) = \sum_{i=1}^{2} \left[(-\widetilde{N}_{k_i})^T \cos(k_i\,\phi) + (\widetilde{M}_{k_i})^T \sin(k_i\,\phi) \right] \begin{Bmatrix} p_{k_i} \\ q_{k_i} \end{Bmatrix}, \tag{3.11}$$

where p_{k_i} and q_{k_i} are independent constants. Substituting (3.6) and (3.11) into (3.8), neglecting two order terms in power ϵ, and noting the independence of p_{k_i} and q_{k_i} yield four algebraic equations in $a_{k_i}(\epsilon)$, $b_{k_i}(\epsilon)$, σ_1, and σ_2. For $\epsilon \neq 0$, $a_{k_i}(\epsilon)$ and $b_{k_i}(\epsilon)$ are dependent. Therefore, we change this four algebraic equations in polar form by setting

$$\begin{aligned} a_{k_i}(\epsilon) &= \frac{-\,r_{k_i}(\epsilon)\,(c_{12}\sin(\theta_i) + d_{12}\sin(k_i\omega\tau_c + \theta_i))}{c_{12}\,(k_i\omega + d_{22}\sin(k_i\omega\tau_c)) + d_{12}\,(k_i\omega\cos(k_i\omega\tau_c) - c_{22}\sin(k_i\omega\tau_c))}, \\ b_{k_i}(\epsilon) &= \frac{-\,r_{k_i}(\epsilon)\,(c_{22}\sin(\theta_i) + d_{22}\sin(k_i\omega\tau_c + \theta_i) + k_i\omega\cos(\theta_i))}{c_{12}\,(k_i\omega + d_{22}\sin(k_i\omega\tau_c)) + d_{12}\,(k_i\omega\cos(k_i\omega\tau_c) - c_{22}\sin(k_i\omega\tau_c))}, \end{aligned} \tag{3.12}$$

where $i = 1, 2$ and $(r_{k_1}, r_{k_2}, \theta_1, \theta_2)$ is a polar coordinate system. Using the fact that $\cos^2(\theta_i) + \sin^2(\theta_i) = 1$ $(i = 1, 2)$, one can solve $r_{k_1}(\epsilon)$, $r_{k_2}(\epsilon)$, σ_1, and σ_2 from these four algebraic equations. Thus, when the time delay and feedback gain is very close to the double Hopf bifurcation point (i.e., $\epsilon\tau_c$ and ϵD_ϵ are very small), the approximate solution in $O(\epsilon)$ is expressed as

$$Z(t) = \begin{Bmatrix} r_{k_1}\cos((k_1\omega + \epsilon\sigma_1)t + \theta_1) + r_{k_2}\cos((k_2\omega + \epsilon\sigma_2)t + \theta_2) \\ (\cdot) \end{Bmatrix}, \tag{3.13}$$

in terms of (3.6) and (3.12), where $(\cdot)$ is a complicated expression. Furthermore, when θ_1 and θ_2 are determined from the initial conditions, $a_{k_i}(\epsilon)$ and $b_{k_i}(\epsilon)$ in (3.6) can be obtained from (3.12), denoted as $a^*_{k_i}(\epsilon)$ and $b^*_{k_i}(\epsilon)$. Thus, (3.6) becomes

$$Z(t) = \sum_{i=1}^{2} \left[\widetilde{N}_{k_i}\cos(k_i\,\omega t + \epsilon\sigma_i t) + \widetilde{M}_{k_i}\sin(k_i\,\omega t + \epsilon\sigma_i t) \right] \begin{Bmatrix} a^*_{k_i}(\epsilon) \\ b^*_{k_i}(\epsilon) \end{Bmatrix}. \tag{3.14}$$

Up to now, an approximate solution is obtained from the perturbation step of the PIS.

It should be noted that the solution (3.6) is a perturbation to (3.5) in an ϵ-order magnitude. Besides, the four algebraic equations are obtained in $O(\epsilon)$. Therefore, the solution of these four algebraic equations is accurate for small ϵ. However, the case for large ϵ should be investigated further.

A large ϵ perturbation yields one of the following cases:

(a) both $\epsilon\tau_\epsilon$ and ϵD_ϵ are small but the nonlinearity in (2.1) is strong;
(b) ϵD_ϵ is small but $\epsilon\tau_\epsilon$ is large and the nonlinearity in (2.1) is strong;
(c) $\epsilon\tau_\epsilon$ is small but ϵD_ϵ is large and the nonlinearity in (2.1) is strong;
(d) both $\epsilon\tau_\epsilon$ and ϵD_ϵ are large and the nonlinearity in (2.1) is strong.

As one will see in sections 4 and 5, the perturbation step of the PIS is still valid for case (a) but invalid for cases (b), (c), and (d). It implies that the approximate solution given in (3.14) is accurate enough to represent the motions near a double Hopf bifurcation point with weak resonance as long as the time delay and feedback gain are close to the point even for very large ϵ. Thus, the second step of the PIS is required for cases (b), (c), and (d) for which the perturbation step is invalid. For these three cases, the values of (τ, D) are far away from (τ_c, D_c).

3.2. Incremental step of the PIS. In this subsection, the incremental step of the PIS is proposed in detail for case (b) since we focus on effects of the time delay on (2.1) in the present paper. It is an extension of our previous work in [26, 27].

Similar to the formulation in [26, 27], a time transformation is first introduced as

$$\frac{d\phi}{dt} = \Phi(\phi), \qquad \Phi(\phi + 2\pi) = \Phi(\phi), \tag{3.15}$$

where ϕ is the new time. Thus, $\Phi(\phi)$ can be approximately expanded in a truncated Fourier series about ϕ as

$$\Phi(\phi) = \sum_{j=0}^{m} (p_j \cos j\,\phi + q_j \sin j\,\phi). \tag{3.16}$$

In the ϕ domain, (2.1) or (2.22) is rewritten as

$$\Phi Z' = C\,Z + D\,Z_\tau + \epsilon\,F(\,Z\,, Z_\tau\,), \tag{3.17}$$

where prime denotes differentiation with respect to ϕ, $D = D_c + \epsilon\,D_\epsilon$, $\tau = \tau_c + \epsilon\tau_\epsilon$.

If ϕ_1 corresponds to $t - \tau$, it follows from (3.15) that

$$\begin{aligned} dt &= \frac{d\phi}{\Phi(\phi)} = \frac{d\phi_1}{\Phi(\phi_1)} \\ \Longrightarrow \quad \Phi(\phi)&\frac{d\phi_1}{d\phi} = \Phi(\phi_1). \end{aligned} \tag{3.18}$$

We note that $\phi_1 - \phi$ is a periodic function in ϕ with period 2π. Similarly, $\phi_1 - \phi$ can also be expanded in a truncated Fourier series about ϕ as

$$\phi_1 = \phi + \sum_{j=0}^{m} (r_j \cos j\phi + s_j \sin j\phi). \tag{3.19}$$

The integration constant of (3.18) provides information about the delay τ. Since ϕ_1 is the new time corresponding to $t - \tau$, it follows from (3.18) that

$$\begin{aligned} \int_{t-\tau}^{t} dt_1 &= \int_{\phi_1}^{\phi} \frac{d\theta}{\Phi(\theta)} \\ \Longrightarrow \quad \tau &= \int_{\phi_1}^{\phi} \frac{d\theta}{\Phi(\theta)}. \end{aligned} \tag{3.20}$$

To consider the continuation with the delay τ as the bifurcation parameter, ϵ and D are kept fixed such that D is close to D_c. If (3.17) possesses a periodic solution at $\tau = \tau_0 = \tau_c + \epsilon\tau_\epsilon$ and the expression (3.14) provides a sufficiently accurate representation, where either $\left\{ \begin{smallmatrix} a^*_{k_1}(\epsilon) \\ b^*_{k_1}(\epsilon) \end{smallmatrix} \right\} = 0$ or $\left\{ \begin{smallmatrix} a^*_{k_2}(\epsilon) \\ b^*_{k_2}(\epsilon) \end{smallmatrix} \right\} = 0$, then a periodic solution at $\tau = \tau_0 + \Delta\tau$ can be expressed in a truncated Fourier series as

$$Z = \sum_{j=0}^{m} (a_j \cos j\,\phi + b_j \sin j\,\phi), \tag{3.21}$$

where $a_j, b_j \in R^2$, $m \geq k_2 > k_1$. Correspondingly, one has

$$Z_\tau = \sum_{j=0}^{m} (a_j \cos j\phi_1 + b_j \sin j\phi_1). \tag{3.22}$$

For $\Delta\tau = 0$, one can easily obtain that

$$a_j = \begin{cases} \widetilde{N}_{k_i} \begin{Bmatrix} a^*_{k_i}(\epsilon) \\ b^*_{k_i}(\epsilon) \end{Bmatrix}, & j = k_i, \\ 0, & j \neq k_i, \end{cases}$$
$$b_j = \begin{cases} \widetilde{M}_{k_i} \begin{Bmatrix} a^*_{k_i}(\epsilon) \\ b^*_{k_i}(\epsilon) \end{Bmatrix}, & j = k_i, \\ 0, & j \neq k_i, \end{cases} \tag{3.23}$$

where $i = 1, 2$. The coefficients of $\Phi(\phi)$ in (3.16) are given by

$$p_0 = \begin{cases} \omega + \epsilon\sigma_1/k_1 & \text{for } \begin{Bmatrix} a^*_{k_2}(\epsilon) \\ b^*_{k_2}(\epsilon) \end{Bmatrix} = 0, \\ \omega + \epsilon\sigma_2/k_2 & \text{for } \begin{Bmatrix} a^*_{k_1}(\epsilon) \\ b^*_{k_1}(\epsilon) \end{Bmatrix} = 0, \end{cases} \tag{3.24}$$

and $p_j = 0$, $q_j = 0$ for all $j > 0$. From (3.19) and (3.20), one has also

$$r_0 = -p_0\,\tau, \tag{3.25}$$

and $r_j = 0$, $s_j = 0$ for all $j > 0$.

A small increment of τ from the initial value of τ_0 to $\tau_0 + \Delta\tau$ yields small increments of the following quantities:

$$Z \to Z + \Delta Z,\ Z_\tau \to Z_\tau + \Delta Z_\tau,\ \Phi \to \Phi + \Delta\Phi,\ \text{and}\ \phi_1 \to \phi_1 + \Delta\phi_1. \tag{3.26}$$

Substituting (3.26) into (3.17) and (3.18), and expanding in Taylor's series about an initial solution, one can obtain linearized incremental equations by ignoring all the nonlinear terms

of small increments as below:

$$
\begin{aligned}
&Z'\Delta\Phi(\phi)+\Phi(\phi)\Delta Z'-C\,\Delta Z-D\,\Delta Z_\tau\\
&-\epsilon\left(\frac{\partial F(Z,Z_\tau)}{\partial Z}\bigg|_0\Delta Z-\frac{\partial F(Z,Z_\tau)}{\partial Z_\tau}\bigg|_0\Delta Z_\tau\right)\\
&=C\,Z+D\,Z_\tau+\epsilon\,F(\,Z\,,Z_\tau\,)-\Phi(\phi)Z',
\end{aligned}
\tag{3.27}
$$

$$
\phi_1'\Delta\Phi(\phi)+\Phi(\phi)\Delta\phi_1'-\Delta\Phi(\phi_1)-\Phi'(\phi_1)\Delta\phi_1=\Phi(\phi_1)-\Phi(\phi)\phi_1',
\tag{3.28}
$$

where the subscript 0 represents the evaluation of the relevant quantities corresponding to the initial solution. From (3.16), (3.19), (3.21), and (3.22), one has

$$
\begin{aligned}
\Delta\Phi(\phi)&=\textstyle\sum_{j=0}^{m}(\Delta p_j\cos j\phi+\Delta q_j\sin j\phi),\\
\Delta\Phi'(\phi)&=\textstyle\sum_{j=1}^{m}j(\Delta q_j\cos j\phi-\Delta p_j\sin j\phi),
\end{aligned}
\tag{3.29}
$$

$$
\begin{aligned}
\Delta\phi_1&=\textstyle\sum_{j=0}^{m}(\Delta r_j\cos j\phi+\Delta s_j\sin j\phi),\\
\Delta\phi_1'&=\textstyle\sum_{j=1}^{m}j(\Delta s_j\cos j\phi-\Delta r_j\sin j\phi),
\end{aligned}
\tag{3.30}
$$

$$
\begin{aligned}
\Delta Z&=\textstyle\sum_{j=0}^{m}(\Delta a_j\cos j\phi+\Delta b_j\sin j\phi),\\
\Delta Z'&=\textstyle\sum_{j=1}^{m}j(\Delta b_j\cos j\phi-\Delta a_j\sin j\phi),
\end{aligned}
\tag{3.31}
$$

and

$$
\Delta Z_\tau=\sum_{j=0}^{m}(\Delta a_j\cos j\phi_1+\Delta b_j\sin j\phi_1)+\frac{\partial Z_\tau}{\partial\phi_1}\Delta\phi_1.
\tag{3.32}
$$

In addition, a small increment of τ_0 to $\tau_0+\Delta\tau$ also yields the linearized incremental equation of (3.20) as

$$
\int_{\phi_1}^{\phi}\frac{\Delta\Phi(\theta)}{\Phi^2(\theta)}d\theta+\frac{\Delta\phi_1}{\Phi(\phi_1)}=\int_{\phi_1}^{\phi}\frac{d\theta}{\Phi(\theta)}-\tau_0-\Delta\tau,
\tag{3.33}
$$

which implies, for $\phi=0$,

$$
\int_{\xi}^{0}\frac{\Delta\Phi(\theta)}{\Phi^2(\theta)}d\theta+\frac{\Delta\phi_1(0)}{\Phi(\alpha)}=\int_{\xi}^{0}\frac{d\theta}{\Phi(\theta)}-\tau_0-\Delta\tau,
\tag{3.34}
$$

where $\xi=\phi_1(0)$.

Substituting (3.29)–(3.32) into (3.27) and using the harmonic balance method, one obtains the linearized equation (3.27) in terms of the increments Δa_j, Δb_j, Δp_j, Δq_j, Δr_j, and Δs_j as

$$
\begin{aligned}
&\textstyle\sum_{j=0}^{m}\left(\Psi_{1,j}\Delta a_j+\Psi_{2,j}\Delta b_j+\Psi_{3,j}\Delta p_j+\Psi_{4,j}\Delta q_j+\Psi_{5,j}\Delta r_j+\Psi_{6,j}\Delta s_j\right)\\
&=\Lambda_1,
\end{aligned}
\tag{3.35}
$$

where

$$\Psi_{1,j} = -j\Phi(\phi)\sin j\phi - C\cos j\phi - D\cos j\phi_1 - \epsilon\left(\left.\frac{\partial F}{\partial Z}\right|_0 \cos j\phi + \left.\frac{\partial F}{\partial Z_\tau}\right|_0 \cos j\phi_1\right),$$

$$\Psi_{2,j} = j\Phi(\phi)\cos j\phi - C\sin j\phi - D\sin j\phi_1 - \epsilon\left(\left.\frac{\partial F}{\partial Z}\right|_0 \sin j\phi + \left.\frac{\partial F}{\partial Z_\tau}\right|_0 \sin j\phi_1\right),$$

$$\Psi_{3,j} = Z'\cos j\phi,$$

$$\Psi_{4,j} = Z'\sin j\phi, \tag{3.36}$$

$$\Psi_{5,j} = -D\frac{\partial Z_\tau}{\partial \phi_1}\cos j\phi - \left.\frac{\partial F}{\partial Z_\tau}\right|_0 \frac{\partial Z_\tau}{\partial \phi_1}\cos j\phi,$$

$$\Psi_{6,j} = -D\frac{\partial Z_\tau}{\partial \phi_1}\sin j\phi - \left.\frac{\partial F}{\partial Z_\tau}\right|_0 \frac{\partial Z_\tau}{\partial \phi_1}\sin j\phi,$$

$$\Lambda_1 = C\,Z + D\,Z_\tau + \epsilon\,F(\,Z\,,Z_\tau\,) - \Phi(\phi)Z'.$$

Similarly, from (3.28) and (3.34), we obtain, respectively,

$$\sum_{j=0}^{m}[\Psi_{7,j}\Delta p_j + \Psi_{8,j}\Delta q_j + \Psi_{9,j}\Delta r_j + \Psi_{10,j}\Delta s_j] = \Lambda_2 \tag{3.37}$$

and

$$\sum_{j=0}^{m}[\Psi_{11,j}\Delta p_j + \Psi_{12,j}\Delta q_j + \Psi_{13,j}\Delta r_j] = \Lambda_3, \tag{3.38}$$

where

$$\begin{aligned}
\Psi_{7,j} &= \phi_1'\cos j\phi - \cos j\phi_1,\\
\Psi_{8,j} &= \phi_1'\sin j\phi - \sin j\phi_1,\\
\Psi_{9,j} &= -j\Phi(\phi)\sin j\phi - \Phi'(\phi_1)\cos j\phi,\\
\Psi_{10,j} &= j\Phi(\phi)\cos j\phi - \Phi'(\phi_1)\sin j\phi,\\
\Psi_{11,j} &= \textstyle\int_\xi^0 \frac{\cos j\theta}{\Phi^2(\theta)}d\theta,\\
\Psi_{12,j} &= \textstyle\int_\xi^0 \frac{\sin j\theta}{\Phi^2(\theta)}d\theta,\\
\Psi_{13,j} &= \textstyle\frac{1}{\Phi(\xi)},\\
\Lambda_2 &= \Phi(\phi_1) - \Phi(\phi)\phi_1',\\
\Lambda_3 &= \textstyle\int_\xi^0 \frac{d\theta}{\Phi(\theta)} - \tau - \Delta\tau.
\end{aligned} \tag{3.39}$$

Since $\Psi_{i,j}$ $(1 \le i \le 13,\ 1 \le j \le m)$ and Λ_k $(1 \le k \le 3)$ are periodic functions in ϕ, they can be expressed in a Fourier series, the coefficients of which can easily be obtained by the method of fast Fourier transform (FFT). Let $a_{ij}, b_{ij} \in R$ $(1 \le i \le 2,\ 0 \le j \le m)$ be the ith element in a_j and b_j, respectively. By comparing the coefficients of $2(2m+1)+1$ harmonic terms of (3.35) and $2m+1$ of (3.37) and (3.38), a system of linear equations is thus obtained with unknowns Δa_{ij}, Δb_{ij}, Δp_j, Δq_j, Δr_j, and Δs_j in the form

$$\begin{aligned} &\textstyle\sum_{i=1}^{2}\sum_{j=0}^{m}(A_{k,ij}\Delta a_{ij} + B_{k,ij}\Delta b_{ij}) \\ &+ \textstyle\sum_{j=0}^{m}(P_{k,j}\Delta p_j + Q_{k,j}\Delta q_j + R_{k,j}\Delta r_j + S_{k,j}\Delta s_j) = T_k, \end{aligned} \tag{3.40}$$

where $k = 1, 2, \ldots, 3(2m+1)+2$ and T_k are the residue terms. The values of a_j, b_j, p_j, q_j, r_j, and s_j are updated by adding the original values and the corresponding incremental values. The iteration process continues until $T_k \to 0$ for all k (in practice, $|T_k|$ is less than a desired degree of accuracy). The entire incremental process proceeds by adding the $\Delta\tau$ increment to the converged value of τ, using the previous solution as the initial approximation until a new converged solution is obtained. We note that the value of m may be changed during the continuation so as to ensure sufficient accuracy of the solution.

The stability of a periodic solution can be determined by the Floquet method [34, 35]. Let $\zeta \in R^2$ be a small perturbation from a periodic solution of (2.1). Then,

$$\frac{d\zeta}{d\varphi} = \frac{1}{\Phi}[A(\varphi,\varphi_1)\zeta + B(\varphi,\varphi_1)\zeta_\tau] + O(\zeta^2, {\zeta_\tau}^2), \tag{3.41}$$

where $A(\varphi,\varphi_1) = C + \epsilon\frac{\partial F(Z,Z_\tau)}{\partial Z}$ and $B(\varphi,\varphi_1) = D + \epsilon\frac{\partial F(Z,Z_\tau)}{\partial Z_\tau}$. The entities of A and B are all periodic functions of φ with period 2π, which can be determined by using the incremental procedure. The time delay interval $I_1 = [-\tau, 0]$ corresponds to $I_2 = [\alpha, 0]$ in the φ domain. Discrete points in I_2 are selected for the computation of Floquet multipliers.

From the incremental procedure, the Fourier coefficients of φ_1 in (3.30) are obtained. Assume that $\varphi = \beta$ when $\varphi_1 = 0$ and let $I_3 = [0, \beta]$. For each $\varphi \in I_3$, there corresponds a unique $\varphi_1 \in I_2$. We choose a mesh size $h = \frac{\beta}{N-1}$ and discrete points $\varphi^{(i)} = ih$ $(0 \le i \le N-1)$ in I_3, which correspond to $\varphi_1^{(i)} = \varphi_1(\varphi^{(i)})$ in I_2. Let $\zeta(\varphi_1^{(i)})$ be the $(i+1)$th unit vector in R^2. By applying numerical integration to (3.1), we obtain the monodromy matrix M as

$$M = [\zeta(\varphi_1^{(0)} + 2\pi), \zeta(\varphi_1^{(1)} + 2\pi), \ldots, \zeta(\varphi_1^{(N-1)} + 2\pi)]. \tag{3.42}$$

The eigenvalues of M are used to determine the stability of the periodic solution. One of the eigenvalues or Floquet multipliers of M must be unity which provides a check for the accuracy of the calculation. If all the other eigenvalues are inside the unit circle, the periodic solution under consideration is stable; otherwise, it is unstable.

Although the incremental step described above is for case (b), it can be extended in a similar way to cases (c) and (d).

In this section, we have proposed the two steps of the PIS in details. As for applications, two examples will be investigated in the next two sections to illustrate the validity of the PIS.

4. Weak resonant double Hopf bifurcation for van der Pol–Duffing system with delayed feedback. First, we consider the van der Pol–Duffing oscillator with linear delayed position feedback governed by

$$\ddot{u} - (\alpha - \epsilon\,\gamma\,u^2)\,\dot{u} + \omega_0^2\,u + \epsilon\,\beta\,u^3 = A(u_\tau - u), \tag{4.1}$$

where α, γ, and β are positive constants, $A < 0$, τ is the time delay, ϵ is the small perturbation parameter, and $u_\tau = u(t-\tau)$. The system (4.1) can be expressed in the form of (2.1), where

$$\begin{aligned} &c_{11} = 0, \quad c_{12} = 1, \quad c_{21} = -\omega_0^2 - A, \quad c_{22} = \alpha, \\ &d_{11} = 0, \quad d_{12} = 0, \quad d_{21} = A, \quad d_{22} = 0, \end{aligned} \tag{4.2}$$

which imply, from (2.4),

$$c_1 = \alpha, \quad c_2 = A + {\omega_0}^2, \quad d_1 = 0, \quad c_d = -A. \tag{4.3}$$

Substituting (4.2) into (2.7)–(2.9), one obtains that if

$$\begin{aligned} 2\,A + {\omega_0}^2 &> 0, \\ \left(-2\,A + \alpha^2\right)^2 - 4\,\alpha^2\,{\omega_0}^2 &> 0, \end{aligned} \tag{4.4}$$

then

$$\omega_\pm = \sqrt{A - \frac{\alpha^2}{2} + {\omega_0}^2 \pm \sqrt{\left(A - \frac{\alpha^2}{2}\right)^2 - \alpha^2\,{\omega_0}^2}}, \tag{4.5}$$

$$\cos(\omega_\pm\,\tau) = 1 + \frac{{\omega_0}^2 - {\omega_\pm}^2}{A}, \quad \sin(\omega_\pm\,\tau) = \frac{\alpha}{A}\,\omega_\pm. \tag{4.6}$$

For $A < 0$ and $\alpha \geq 0$, it follows from (4.6) that

$$\tau_\pm[j] = \frac{1}{\omega_\pm}\left[2j\pi - \cos^{(-1)}\left(1 + \frac{\omega_0^2 - \omega_\pm^2}{A}\right)\right], \tag{4.7}$$

where $j = 1, 2, 3, \ldots$, and $\omega_\pm$ are determined by (4.5). With aids of (2.12) and (2.15), the necessary conditions in terms of the critical values A_c and τ_c for the occurrence of resonant double Hopf points with frequencies in the ratio $k_1 : k_2$ are given by

$$\alpha^2 - 2\,(A_c + {\omega_0}^2) + \frac{({k_1}^2 + {k_2}^2)\omega_0}{k_1\,k_2}\sqrt{2A_c + {\omega_0}^2} = 0, \tag{4.8}$$

$$\tau_c = \tau_+ = \tau_-, \tag{4.9}$$

and the corresponding frequencies are

$$\omega_- = \sqrt{\frac{k_1}{k_2}\,\omega_0\,\sqrt{2A_c + {\omega_0}^2}}, \quad \omega_+ = \sqrt{\frac{k_2}{k_1}\,\omega_0\,\sqrt{2A_c + {\omega_0}^2}}, \tag{4.10}$$

Table 1

Some critical values at possible weak resonant double Hopf bifurcations for the system (4.1).

j	$\omega_- : \omega_+$	α	A_c	ω_-	τ_c
1	3:4	0.301697	-0.306536	0.683027	5.67888
	3:5	0.17468	-0.336338	0.585895	5.88976
	4:5	0.32397	-0.301196	0.710246	5.64733
	5:6	0.33501	-0.298532	0.727308	5.63222
2	4:5	0.073289	-0.1892405	0.79416	12.261248
	5:6	0.106497	-0.176272	0.818864	12.1415189
3	6:7	0.035351	-0.13606	0.855147	18.63072
	7:8	0.0555963	-0.1268220	0.869442	18.5164772
4	8:9	0.020823	-0.106556	0.8879772	24.9618595
	9:10	0.0343324	-0.099884	0.8972743	24.85826

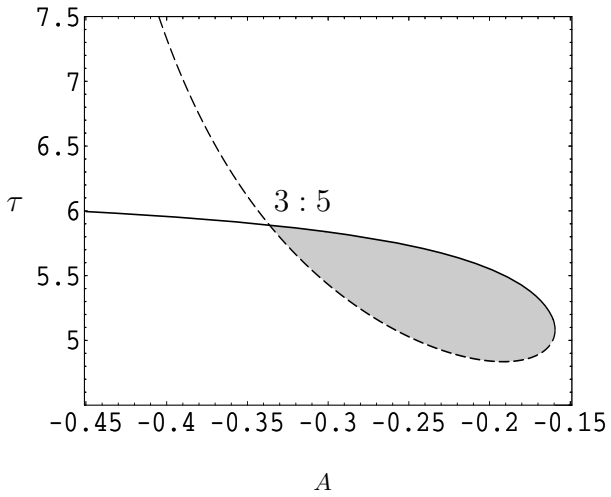


Figure 1. *High-order resonant double Hopf bifurcation diagram with frequencies in the ratio $\omega_- : \omega_+ = k_1 : k_2$ for $\alpha = 0.17468$ and $\omega_0 = 1.0$, where the solid line represents τ_+, the dashed line τ_-, and the grey region amplitude death.*

where ω_0 is a constant.

For $\alpha \neq 0$, A_c cannot be solved from (4.9) in a closed form, but they can be easily solved numerically for a fixed ω_0. Some values of the possible weak resonant double Hopf bifurcation are shown in Table 1 for $\omega_0 = 1$. In Figure 1, we plot the diagrams for the case with $\alpha = 0.17468$ in Table 1. The grey color regions show the stable trivial solutions for system (4.1), i.e., amplitude death regions. The intersection points located on the amplitude death regions are two double Hopf bifurcation points with 3:5 resonance. To obtain the neighboring solutions derived from such double Hopf bifurcation, we let $A = A_c + \epsilon A_\epsilon$ and $\tau = \tau_c + \epsilon\tau_\epsilon$ for a given α, where ϵA_ϵ and $\epsilon\tau_\epsilon$ are very small. Thus, system (4.1) can be rewritten as

$$\dot{Z}(t) = C\,Z(t) + D_c\,Z(t-\tau_c) + \widetilde{F}(Z(t), Z(t-\tau_c), Z(t-\tau_c-\epsilon\tau_\epsilon)\,, \epsilon)\,, \tag{4.11}$$

where

$$
\begin{aligned}
&Z(t) = \left\{ \begin{array}{c} u(t) \\ v(t) \end{array} \right\}, \quad C = \left[\begin{array}{cc} 0 & 1 \\ -(A_c + {\omega_0}^2) & \alpha \end{array} \right], \quad D_c = \left[\begin{array}{cc} 0 & 0 \\ A_c & 0 \end{array} \right], \\
&\widetilde{F} = \left\{ \begin{array}{c} 0 \\ A_c \left(u_{\tau_c + \epsilon\tau_\epsilon} - u_{\tau_c} \right) + \epsilon \left(A_\epsilon \left(u_{\tau_c + \epsilon\tau_\epsilon} - u \right) - u^2 \left(\beta u + \gamma v \right) \right) \end{array} \right\},
\end{aligned} \tag{4.12}
$$

where $u_\tau = u(t - \tau)$.

As an example, we consider the case shown in Figure 1 with $\omega_0 = 1$, $\alpha = 0.17468$, where a 3:5 weak resonant double Hopf bifurcation occurs in system (4.1) at $A_c = -0.336338$ and $\tau_c = 5.88976$ (cf. Lemma 8.15 in [36]). It follows from (3.11), (3.12), and (3.13) that

$$
W(t) = \left\{ \begin{array}{c} w_1(t) \\ w_2(t) \end{array} \right\}, \tag{4.13}
$$

with

$$
\begin{aligned}
w_1(t) = \; & -0.17468 p_{k_1} \cos(0.585895t) - 0.17468 p_{k_2} \cos(0.976492t) \\
& - 0.343274 q_{k_1} \cos(0.585895t) - 0.953538 q_{k_2} \cos(0.976492t) \\
& + 0.585895 p_{k_1} \sin(0.585895t) - 0.102345 q_{k_1} \sin(0.585895t) \\
& + 0.976492 p_{k_2} \sin(0.976492t) - 0.170574 q_{k_2} \sin(0.976492t), \\
w_2(t) = \; & p_{k_1} \cos(0.585895t) + p_{k_2} \cos(0.976492t) \\
& + 0.585895 q_{k_1} \sin(0.585895t) + 0.976492 q_{k_2} \sin(0.976492t),
\end{aligned} \tag{4.14}
$$

$$
\begin{aligned}
a_{k_1} &= -1.70679 r_{k_1} \sin(\theta_1), \quad b_{k_1} = -r_{k_1} \cos(\theta_1) - 0.298142 r_{k_1} \sin(\theta_1), \\
a_{k_2} &= -1.02407 r_{k_2} \sin(\theta_2), \quad b_{k_2} = -r_{k_2} \cos(\theta_2) - 0.178885 r_{k_2} \sin(\theta_2),
\end{aligned} \tag{4.15}
$$

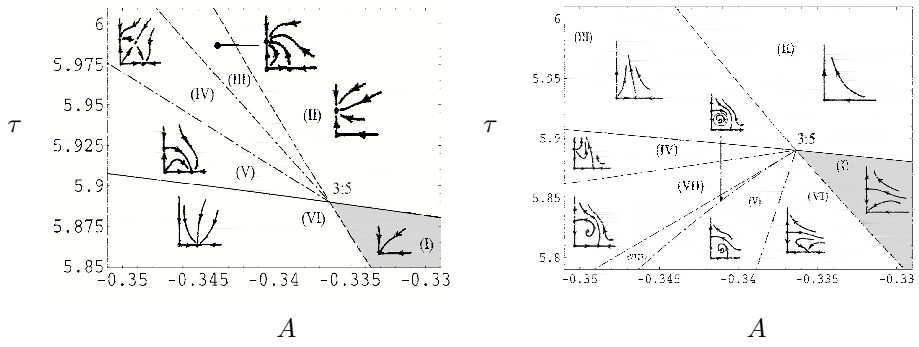
and

$$
\begin{aligned}
u(t) &= r_{k_1} \cos(\theta_1 + t\,(0.585895 + \epsilon\sigma_1)) + r_{k_2} \cos(\theta_2 + t\,(0.976492 + \epsilon\sigma_2)), \\
v(t) &= -0.585896 r_{k_1} \sin(\theta_1 + t\,(0.585895 + \epsilon\sigma_1)) \\
&\quad - 0.976492 r_{k_2} \sin(\theta_2 + t\,(0.976492 + \epsilon\sigma_2)).
\end{aligned} \tag{4.16}
$$

Substituting (4.12), (4.13), (4.14), and (4.16) into (3.8), noting that p_1, p_2, q_1, and q_2 are independent and $\cos^2(\theta_i) + \sin^2(\theta_i) = 1$ $(i = 1, 2)$, yields

$$
\begin{aligned}
& r_{k_1} \epsilon \left(A_\epsilon + 0.384107 {r_{k_1}}^2 \beta + 0.768215 {r_{k_2}}^2 \beta - 0.291413\sigma_1 + 0.0307097\tau_\epsilon \right) = 0, \\
& r_{k_1} \epsilon \left(A_\epsilon - 0.481362 {r_{k_1}}^2 \gamma - 0.962723 {r_{k_2}}^2 \gamma - 5.62729\sigma_1 - 0.616891\tau_\epsilon \right) = 0, \\
& r_{k_2} \epsilon \left(A_\epsilon + 10.8584 {r_{k_1}}^2 \beta + 5.42919 {r_{k_2}}^2 \beta - 6.865\sigma_2 + 1.20574\tau_\epsilon \right) = 0, \\
& r_{k_2} \epsilon \left(A_\epsilon - 0.962726 {r_{k_1}}^2 \gamma - 0.481363 {r_{k_2}}^2 \gamma + 3.71089\sigma_2 + 0.558141\tau_\epsilon \right) = 0.
\end{aligned} \tag{4.17}
$$

It can be seen from (4.16) that r_{k_1}, r_{k_2}, σ_1, and σ_2 determine the feature of motions of the system (4.11) when the double Hopf point at (A_c, τ_c) is perturbed by A_ϵ and τ_ϵ for given values of ϵ, β, and γ. Such motion can be amplitude death ($r_{k_1} = 0$ and $r_{k_2} = 0$), periodic ($r_{k_1} = 0$ and $r_{k_2} \neq 0$ or $r_{k_1} \neq 0$ and $r_{k_2} = 0$), or quasi-periodic ($r_{k_1} \neq 0$ and $r_{k_2} \neq 0$). Therefore, it is



(a) *simple case:* $\gamma = 1$, $\beta = 0$ (b) *difficult case:* $\gamma = 0$, $\beta = 4$

Figure 2. *Classification and bifurcation sets of the solution for system* (4.11) *due to* 3:5 *resonant double Hopf bifurcation where solid lines, dashed lines, and dot-dashed lines represent boundaries and amplitude death region is in grey for* (a) *the simple case:* $\gamma = 1$, $\beta = 0$ *and* (b) *the difficult case:* $\gamma = 0$, $\beta = 4$.

necessary to classify the solutions of the algebraic equation (4.17) in the neighborhood of the double Hopf point (A_c, τ_c).

To this end, one has first to distinguish so called "simple" and "difficult" double Hopf cases (see, for example, section 8.6 in [36]). In the simple case, the truncated cubic amplitude system (4.17) has no periodic orbits and addition of any fourth- and fifth-order terms does not change its bifurcation diagram. In the difficult case, one has to consider the truncated fifth-order amplitude system to "stabilize" the Hopf bifurcation there. This Hopf bifurcation implies the existence of three-dimensional invariant tori in the full four-dimensional system on the central manifold. Both cases can be seen when one takes $\beta = 0$, $\gamma \neq 0$ and $\beta \neq 0$, $\gamma = 0$, respectively, as shown in Figures 2(a) and 2(b).

We do not discuss the simple case shown in Figure 2(a) here as a similar case will be discussed at length in the next section.

As for the difficult case with $\gamma = 0$, solving (4.17) yields that $(r_{k_1}, r_{k_2}) = (r_{k_10}, r_{k_20}) = (0, 0)$ is always a root and that up to three other roots (in the positive quadrant) can appear, as follows:

$$
\begin{aligned}
&(r_{k_1}, r_{k_2}) = (r_{k_11}, r_{k_20}) = \left(\, 1.57118\sqrt{\tfrac{-A_\epsilon - 0.0660777\tau_\epsilon}{\beta}},\, 0\,\right) \quad \text{for} \quad \tfrac{A_\epsilon + 0.0660777\tau_\epsilon}{\beta} < 0,\\
&(r_{k_1}, r_{k_2}) = (r_{k_10}, r_{k_21}) = \left(\, 0,\, 0.724522\sqrt{\tfrac{-A_\epsilon - 0.785373\tau_\epsilon}{\beta}}\,\right) \quad \text{for} \quad \tfrac{A_\epsilon + 0.785373\tau_\epsilon}{\beta} < 0,\\
&(r_{k_1}, r_{k_2}) = (r_{12}, r_{22}) = \left(\, 0.68769\sqrt{\tfrac{A_\epsilon - 0.46195\tau_\epsilon}{\beta}},\, 1.21275\sqrt{\tfrac{-A_\epsilon - 0.019497\tau_\epsilon}{\beta}}\,\right)\\
&\qquad \text{for} \quad \tfrac{A_\epsilon - 0.46195\tau_\epsilon}{\beta} > 0, \quad \tfrac{A_\epsilon + 0.019497\tau_\epsilon}{\beta} < 0,
\end{aligned}
\tag{4.18}
$$

$$
\sigma_1 = 0.177705\,(A_\epsilon - 0.616891\tau_\epsilon)\,, \quad \sigma_2 = -0.269477\,(A_\epsilon + 0.558141\tau_\epsilon)\,.
$$

The stability of the solution (4.16) determined by (4.18) can be easily analyzed with the aid of (4.17). Thus, the parameter plane (A, τ) in the neighborhood of (A_c, τ_c) is divided into seven regions (I)–(VII) bounded by (4.18), as shown in Figure 2(b), which is very similar to that produced by Guckenheimer and Holmes (cf. Figure 7.5.5 in [32]).

In Figure 2(b), there are a stable trivial solution $(0,0)$ and an unstable periodic solution $(0, r_{k_2 1})$ in region (I) which is an amplitude death region. With (A, τ) changing to region (II), the trivial solution loses its stability and no local solution appears. Two unstable solutions at $(0,0)$ and $(r_{k_1 1}, 0)$ exist in region (III). When (A, τ) enters into region (IV), there are three unstable solutions given by $(0,0)$, $(r_{k_1 1}, 0)$, and $(0, r_{k_2 1})$. The stable nontrivial solutions of system (4.17) occur in regions (V) and (VI), determined by $(r_{k_1 2}, r_{k_2 2})$ and $(r_{k_1 1}, 0)$, respectively. The other unstable solutions in region (V) are at $(0,0)$, $(r_{k_1 1}, 0)$, and $(0, r_{k_2 1})$, and those in region (VI) at $(0,0)$, $(0, r_{k_2 1})$. It is easily seen from (4.16) that the stable solution in region (V) is quasi-periodic as σ_1/σ_2 is not a rational number. Particularly, the Hopf bifurcation of the nontrivial equilibrium at $(r_{k_1}, r_{k_2}) = (r_{12}, r_{22})$ occurs in the cubic amplitude system (4.17). This leads to the boundary between regions (VII) and (VIII), and a three-dimensional invariant torus in the full four-dimensional system on the central manifold occurs in region (VIII). Thus, one has to consider the truncated fifth-order amplitude system to observe the cycle generated by the Hopf bifurcation. Following section 8.6 in [36], one can obtain the border between regions (VIII) and (V), given by

$$\tau = -15.377 - 134.333A - 211.40A^2, \quad A \le A_c, \tag{4.19}$$

along which the cycle coexists with three saddles at $(r_{k_1 0}, r_{k_2 0})$, $(r_{k_1 1}, r_{k_2 0})$, and $(r_{k_1 0}, r_{k_2 1})$. Thus, the cycle disappears via a heteroclinic bifurcation when the parameter is varied from regions (VIII) to (V). These results are also sketched in Figure 2(b).

Now, the numerical simulation is employed to examine the validity and accuracy of step one of the PIS (or perturbation step of the PIS) for the difficult case as shown in Figure 2(b). To this end, the following two cases are considered. The values of the delay and gain are chosen to be, first, close to, and second, far away from the double Hopf bifurcation point at (A_c, τ_c) in Figure 2(b). The Runge–Kutta scheme is adopted to produce the numerical results, where $\beta = 4$, $\gamma = 0$, and the other parameters are the same as those in Figure 1. The gain A is kept fixed for the two cases as the effect of the time delay on the system under consideration is the prime concern.

First, we consider the case of A kept fixed and τ varying in region (VI). Figure 3 shows a comparison between the approximate solution (4.16) represented by a solid line and the numerical simulation from system (4.1) represented by crossing symbols for (a) $A = -0.3348$, $\tau = 5.852$; (b) $A = -0.33$, $\tau = 5.7$; (c) $A = -0.31$, $\tau = 5.4$; and (d) $A = -0.33$, $\tau = 5.2$. The values of (A, τ) in Figures 3(a) and (b) are close to (A_c, τ_c), while those of Figures 3(c) and (d) are far away. In Figures 3(a) and (b), the analytical prediction is in good agreement with the numerical result. This implies that the periodic solutions obtained from the presented method are accurate even for large ϵ as long as (A, τ) are near the double Hopf bifurcation point at (A_c, τ_c). With (A, τ) drifting away from (A_c, τ_c), the accuracy decreases as shown in Figures 3(b) and (d). This can be seen clearly from Figures 4(a) and (b), where τ is decreased from τ_c while A is kept fixed.

A similar conclusion can be obtained for (A, τ) in region (V). The time history of the quasi-periodic solution of system (4.1) is illustrated in Figures 5(a) and (b) with small and large ϵ, where the approximate solution (4.16) is represented by a solid line and the numerical simulation by crossing symbols. It follows from Figure 5, where good agreement is observed between the analytical prediction and the numerical simulation, that step one of the PIS can

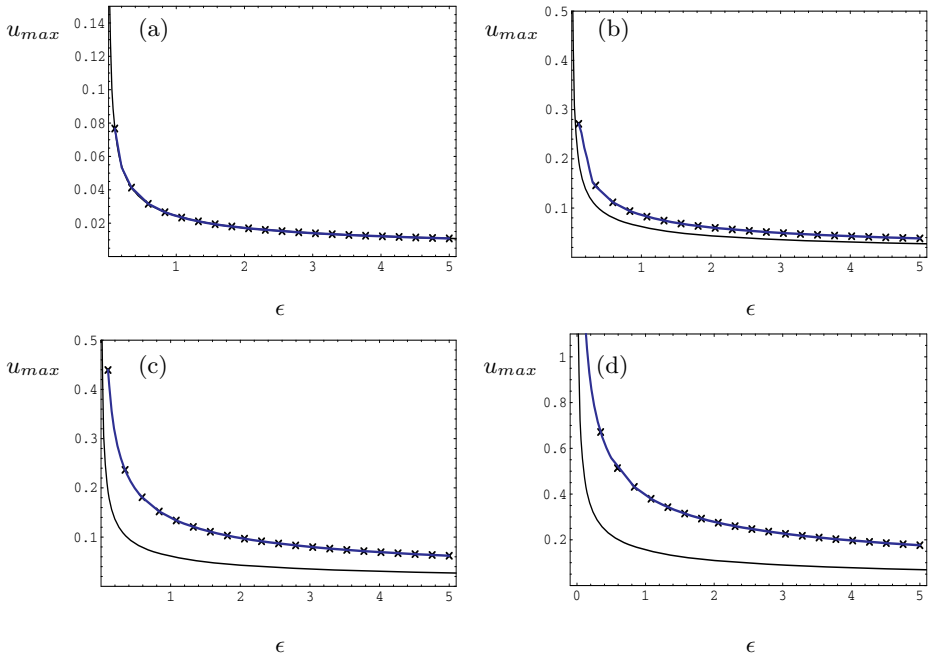


Figure 3. *A comparison among the approximate solution* (4.16) *(solid), the solution from step two of the PIS (thick blue), and the numerical simulation (crossing symbols) in* Max(u) *versus* ϵ *for the periodic solution of system* (4.1) *with* (A, τ) *located in region* (VI) *in Figure* 2*:* (a) $A = -0.3348$, $\tau = 5.852$, (b) $A = -0.33$, $\tau = 5.7$, (c) $A = -0.31$, $\tau = 5.4$, *and* (d) $A = -0.33$, $\tau = 5.2$.

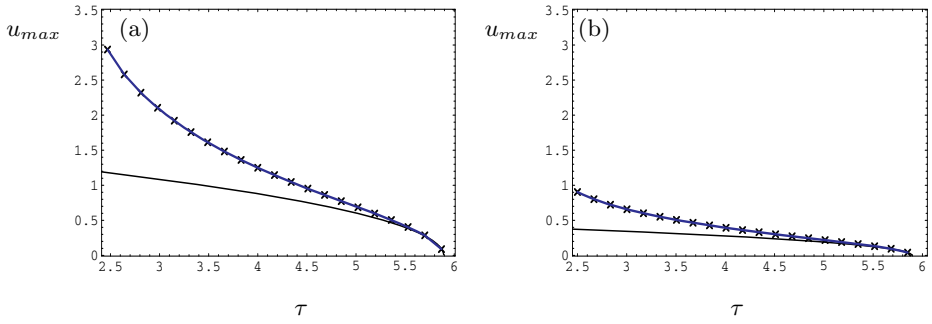


Figure 4. *A comparison between the approximate solution* (4.16) *(solid), the solution from step two of the PIS (thick blue), and the numerical simulation (crossing symbols) in* Max(u) *versus* τ *for the periodic solution of the system* (4.1) *when* (A, τ) *is located in region* (VI) *of Figure* 2*, where* (a) $\epsilon = 0.1$ *and* (b) $\epsilon = 1.0$ *and* $A = -0.3365$.

provide an analytical expression with high accuracy in the neighborhood of the double Hopf bifurcation point even for the quasi-periodic motions. However, with (A, τ) drifting away from the bifurcation point, the method becomes invalid as shown in Figure 6.

From Figures 3 to 6, one can see that step one of the PIS provides not only a fairly good prediction qualitatively but also an analytical expression with high accuracy for both periodic and quasi-periodic motions when parameters are chosen to be close to the weak resonant point. However, the analytical expression is not accurate enough quantitatively when the

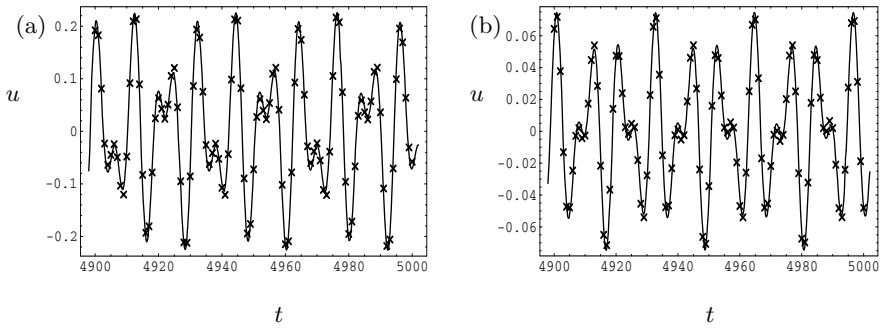


Figure 5. *A comparison between the approximate solution given in* (4.16) *(solid) and the numerical simulation (crossing symbols) in time history for the quasi-periodic solution of system* (4.1)*, where* (a) $\epsilon = 0.1$, (b) $\epsilon = 1$, *and* $A = -0.34$, $\tau = 5.85$ *are chosen in region* (V) *of Figure* 2 *and close to* (A_c, τ_c).

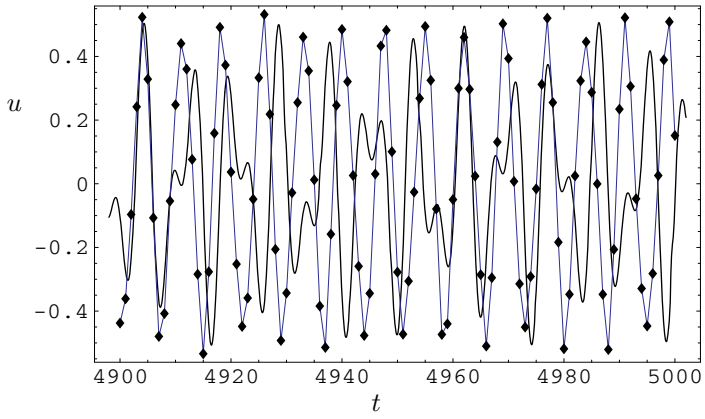


Figure 6. *A comparison between the approximate solution* (4.16) *(solid) and the numerical simulation (line with diamond symbols) in time history for the quasi-periodic solution of system* (4.1) *when* (A, τ) *is in region* (V) *but far away from* (A_c, τ_c)*, where* $\epsilon = 1$, $A = -0.5$, $\tau = 4$.

bifurcation parameters drift away from the bifurcation point. The approximate expression is now considered as an initial guess for step two of the PIS which traces the periodic solutions for bifurcation parameters far away from (A_c, τ_c).

To this end, we choose $\epsilon = 1$ and $(\tau, A) = (5.85, -0.3365)$ which is close to (A_c, τ_c) as the starting point. It follows from Figure 2(b) that this point is located in region (VI) at which the stable periodic solution is given by

$$Z(t) = \begin{Bmatrix} r_{k_1 1} \cos((0.585895 + \epsilon\sigma_1)\, t + \theta_1) \\ -0.585896\, r_{k_1 1} \sin((0.585895 + \epsilon\sigma_1)\, t + \theta_1) \end{Bmatrix}, \tag{4.20}$$

where $r_{k_1 1}$ and σ_1 are determined by (4.18), $\tau_\epsilon = \tau - \tau_c$, $A_\epsilon = A - A_c$, and θ_1 is determined by initial values. Thus, $\theta_1 = 0$ by setting $v(0) = 0$. The solution (4.20) is in good agreement with the numerical one as depicted in Figure 7(a), but such agreement disappears for $\tau = 3$ as shown in Figure 7(b). If the solution (4.20) is considered as an initial guess of step two, then the initial coefficients in the incremental solution given by (3.21), (3.16), and (3.19) can

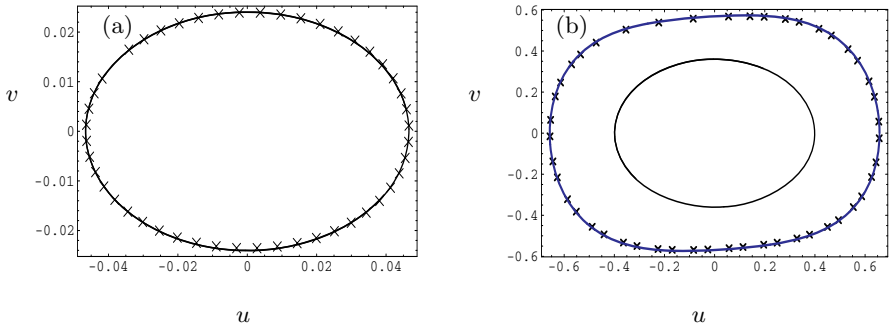


Figure 7. *A comparison between the approximate solution (4.16) (solid), solution (thick blue) from step two of PIS and the numerical simulation (crossing symbols) in phase plane for the periodic solution of system (4.1), where* (a) $\tau = 5.85$, (b) $\tau = 3$, $\epsilon = 1$, *and* $A = -0.3365$ *throughout.*

Table 2

Coefficients of u, v, $\Phi(\phi)$, and ϕ_1 for solution derived from step two of the PIS in Figure 7(b).

$m = 8, \quad A = -0.38, \quad \epsilon = 1.0, \quad \tau = 3.0$								
i	a_{1i}	b_{1i}	a_{2i}	b_{2i}	p_i	q_i	r_i	s_i
0	0	0	0	0	0.10761	0	-3.88243	0
1	0.6575807	0	0.01062	-0.61389	0	0	0	0
2	0	0	0	0	0.166645	-0.04232	0.02121	-0.01261
3	0	0	-0.01085	-0.043191	0	0	0	0
4	0	0	0	0	-0.00416	0.00096	-0.00111	0.00024
5	0	0	0.00021	0.00145	0	0	0	0
6	0	0	0	0	-0.00054	-0.00010	0.00006	-0.00002
7	0	0	-0.00002	-0.00009	0.	0	0	0
8	0	0	0	0	-0.00002	0	0	0

be expressed as (cf. (3.23), (3.24), and (3.25))

$$a_0 = 0, \quad a_1 = \begin{Bmatrix} r_{k_1 1} \\ 0 \end{Bmatrix}, \quad b_1 = \begin{Bmatrix} 0 \\ -0.585895\, r_{k_1 1} \end{Bmatrix}, \qquad a_j = b_j = 0 \quad \text{for } j = 2, \ldots, m, \tag{4.21}$$

$$p_0 = 0.195298 + \frac{\sigma_1}{3}, \quad p_j = q_j = 0 \quad \text{for } j = 1, \ldots, m, \tag{4.22}$$

and

$$r_0 = -p_0\,\tau = -5.85\left(0.195298 + \frac{\sigma_1}{3}\right), \quad r_j = s_j = 0 \quad \text{for } j = 1, \ldots, m. \tag{4.23}$$

With the incremental procedure from $\tau = 5.85$ down to $\tau = 5.85 + n\Delta\tau = 3$ ($n \in Z^+$, $|\Delta\tau| \ll 1$) and step two of the PIS in terms of (3.26)–(3.42), a converged periodic solution at $\tau = 3$ can easily be obtained after a few iterations as shown in Figure 7(b). It follows from Figure 7(b) that although the approximate solution derived from the theorem is far away from the numerical solution, an accurate solution can be obtained through the incremental process of the PIS. The updated solution by step two of the PIS is listed in Table 2.

Similarly, we can use step two to consider the other cases of Figures 3 and 4. However, the incremental step of the PIS has not been presented here for the continuation of quasi-periodic solutions shown in Figures 5 and 6. This will be investigated in future.

Finally, the numerical simulation verifies that there is a stable trivial solution in region (I) but not any stable solution in regions (II), (III), (IV), and (VII).

5. Stuart–Landau system with time delay. As the second example, we consider the Stuart–Landau system with a limit cycle oscillator governed by

$$\begin{aligned} \dot{x} &= \alpha\, x - \omega_0\, y - (x^2 + y^2)x, \\ \dot{y} &= \omega_0\, x + \alpha\, y - (x^2 + y^2)\, y, \end{aligned} \tag{5.1}$$

where ω_0 is the frequency of the oscillator, α a real positive constant. It is easily seen that system (5.1) has a stable limit cycle of amplitude $\sqrt{\alpha}$ with frequency ω_0. The system (5.1) is the normal form of a supercritical Hopf bifurcation for a two-dimensional autonomous system and it can also be obtained by averaging the van der Pol–Mathieu system. Therefore, it is useful for modeling numerous engineering problems. In addition, it is often regarded as a basic element in a large scale system such as the Kuramoto model [29]. The collective dynamics of a large scale system depend on the dynamic behavior of each individual subsystem and the coupling between these subsystems. The time delay often occurs in such a coupling due to finite propagation speeds of signals, finite reaction times, and finite processing times [30, 31]. Thus, the dynamics of each oscillator can be considered as being excited by a driven source which comes from the collective feedback of the other oscillators. Motivated by such a view, we consider

$$\begin{aligned} \dot{x} &= \alpha\, x - \omega_0\, y - \epsilon(x^2 + y^2)x - K\, x_\tau, \\ \dot{y} &= \omega_0\, x + \alpha\, y - \epsilon(x^2 + y^2)\, y - K\, y_\tau \end{aligned} \tag{5.2}$$

to study the resonant double Hopf bifurcation due to the time delay, where $x_\tau = x(t-\tau)$, $y_\tau = y(t-\tau)$, K is a real positive constant and represents a feedback strength, and ϵ is a small or large parameter.

Comparing system (5.2) with the linear part of (2.1), one obtains

$$\begin{aligned} &c_{11} = c_{22} = \alpha, \quad c_{21} = c_{12} = \omega_0, \\ &d_{11} = d_{22} = -K, \quad d_{12} = d_{21} = 0, \quad \det(D) = K^2. \end{aligned} \tag{5.3}$$

For $K \geq \alpha$, substituting (5.3) into (2.18) and (2.19) yields

$$\begin{aligned} \omega_- &= \omega_0 - \sqrt{K^2 - \alpha^2}, \\ \omega_+ &= \omega_0 + \sqrt{K^2 - \alpha^2}, \end{aligned} \tag{5.4}$$

$$\begin{aligned} \tau_-[j] &= \tfrac{1}{\omega_-}\left(2j\pi - \cos^{-1}(\alpha/K)\right), \\ \tau_+[j] &= \tfrac{1}{\omega_+}\left(2j\pi + \cos^{-1}(\alpha/K)\right), \end{aligned} \tag{5.5}$$

and a $k_1 : k_2$ weak resonant double Hopf bifurcation can be expressed as

$$K_c = \frac{\alpha}{\cos\left(\frac{2j\pi(k_2-k_1)}{k_2+k_1}\right)}, \quad \tau_c = \frac{4jk_1\pi}{(k_1+k_2)\omega_-} \tag{5.6}$$

provided that $\omega_0 > \sqrt{K^2-\alpha^2}$, $k_1, k_2 \in Z^+$, $k_1 < k_2$, and $k_1 \neq 1$. Correspondingly, $\omega_\pm$ are expressed as

$$\omega_- = \frac{2\,k_1\,\omega_0}{k_1+k_2}, \quad \omega_+ = \frac{2\,k_2\,\omega_0}{k_1+k_2} \tag{5.7}$$

at the double Hopf point. The explicit expressions in (5.4) and (5.5) are the same as that derived by Reddy and his coauthors [12]. By using (5.6)–(5.7), we obtain the distribution of some weak resonant double Hopf points, as shown in Table 3.

Table 3

Some critical values at possible resonant double Hopf bifurcation for system (5.2).

j	$\omega_- : \omega_+$	α	K_c	ω_-	τ_c
1	2:3	$\frac{-1+\sqrt{5}}{5\sqrt{2\left(5+\sqrt{5}\right)}}$	$\frac{2}{5}\sqrt{\frac{2}{5+\sqrt{5}}}$	$\frac{4}{5}$	2π
	3:4	$\frac{\cos(\frac{2\pi}{7})}{\sqrt{49-49\cos(\frac{2\pi}{7})^2}}$	$\frac{1}{\sqrt{49-49\cos(\frac{2\pi}{7})^2}}$	$\frac{6}{7}$	2π
	3:5	0	$\frac{1}{4}$	$\frac{3}{4}$	2π
	4:5	$\frac{\cos(\frac{2\pi}{9})}{9\sqrt{1-\cos(\frac{2\pi}{9})^2}}$	$\frac{1}{\sqrt{81-81\cos(\frac{2\pi}{9})^2}}$	$\frac{8}{9}$	2π
	5:6	$\frac{\cos(\frac{2\pi}{11})}{11\sqrt{1-\cos(\frac{2\pi}{11})^2}}$	$\frac{1}{\sqrt{121-121\cos(\frac{2\pi}{11})^2}}$	$\frac{10}{11}$	2π
	5:7	$\frac{1}{6\sqrt{3}}$	$\frac{1}{3\sqrt{3}}$	$\frac{5}{6}$	2π
	5:8	$\frac{3\cos(\frac{6\pi}{13})}{13\sqrt{1-\cos(\frac{6\pi}{13})^2}}$	$\frac{3}{13\sqrt{1-\cos(\frac{6\pi}{13})^2}}$	$\frac{10}{13}$	2π
	6:7	$\frac{\cos(\frac{2\pi}{13})}{13\sqrt{1-\cos(\frac{2\pi}{13})^2}}$	$\frac{1}{\sqrt{169-169\cos(\frac{2\pi}{13})^2}}$	$\frac{12}{13}$	2π
2	4:5	$\frac{\cos(\frac{4\pi}{9})}{9\sqrt{1-\cos(\frac{4\pi}{9})^2}}$	$\frac{1}{\sqrt{81-81\cos(\frac{4\pi}{9})^2}}$	$\frac{8}{9}$	4π
	5:6	$\frac{\cos(\frac{4\pi}{11})}{11\sqrt{1-\cos(\frac{4\pi}{11})^2}}$	$\frac{1}{\sqrt{121-121\cos(\frac{4\pi}{11})^2}}$	$\frac{10}{11}$	4π
	6:7	$\frac{\cos(\frac{4\pi}{13})}{13\sqrt{1-\cos(\frac{4\pi}{13})^2}}$	$\frac{1}{\sqrt{169-169\cos(\frac{4\pi}{13})^2}}$	$\frac{12}{13}$	4π
3	6:7	$\frac{\cos(\frac{6\pi}{13})}{13\sqrt{1-\cos(\frac{6\pi}{13})^2}}$	$\frac{1}{\sqrt{169-169\cos(\frac{6\pi}{13})^2}}$	$\frac{12}{13}$	6π

It can be seen from Table 3 that the resonant double points occur periodically with increasing τ_c provided that $\omega_0 > \sqrt{{K_c}^2 - {\alpha_c}^2}$.

Figures 8(a) and (b) show three cases for the frequency ratios 2:3, 3:5, and 7:9, respectively, where the grey regions represent the amplitude death. The parameters for 2:3 and 3:5 are presented in Table 3.

To study the solutions derived from the double Hopf bifurcation, we let $\tau = \tau_c + \epsilon\tau_\epsilon$ and $K = K_c + \epsilon K_\epsilon$ so that system (5.2) becomes

$$\dot{Z}(t) = C\,Z(t) + D_c\,Z(t-\tau_c) + \widetilde{F}(Z(t), Z(t-\tau_c), Z(t-\tau_c-\epsilon\tau_\epsilon), \epsilon), \tag{5.8}$$

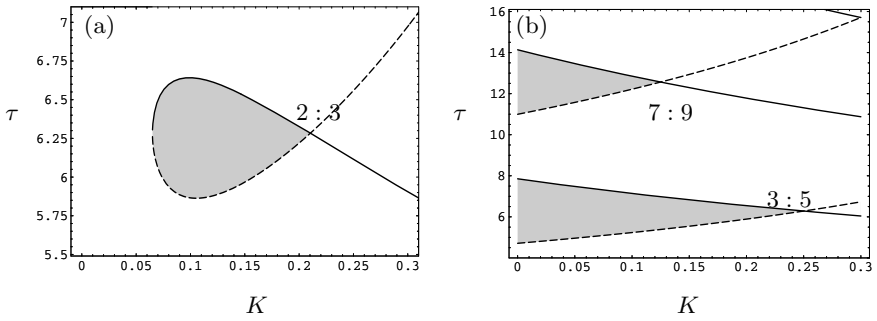


Figure 8. *Resonant double Hopf bifurcation diagrams with frequencies in the ratio of* $\omega_- : \omega_+ = k_1 : k_2$ *for* (a) $\alpha = \frac{-1+\sqrt{5}}{5\sqrt{2\left(5+\sqrt{5}\right)}}$, $\omega_0 = 1.0$, $K_c = \frac{2}{5}\sqrt{\frac{2}{5+\sqrt{5}}}$ *corresponding to frequency ratio* 2:3 *(*$j = 1$*), and* (b) $\alpha = 0$, $\omega_0 = 1.0$ *to ratio* 3:5 *(*$j = 1$, $K_c = \frac{1}{4}$*) and* 7:9 *(*$j = 2$, $K_c = \frac{1}{8}$*), where solid lines represent* τ_+*, dashed lines* τ_-*, and grey regions amplitude death.*

where

$$Z(t) = \left\{ \begin{array}{c} x(t) \\ y(t) \end{array} \right\}, \quad C = \left[\begin{array}{cc} \alpha & -\omega_0 \\ \omega_0 & \alpha \end{array} \right], \quad D_c = \left[\begin{array}{cc} -K_c & 0 \\ 0 & -K_c \end{array} \right],$$

$$D_\epsilon = \left[\begin{array}{cc} -K_\epsilon & 0 \\ 0 & -K_\epsilon \end{array} \right], \quad F(\,Z(t)\,) = \left\{ \begin{array}{c} -x(x^2+y^2) \\ -y(x^2+y^2) \end{array} \right\},$$

$$\widetilde{F}(Z, Z_{\tau_c}, Z_{\tau_c+\epsilon\tau_\epsilon}, \epsilon) = D_c\left[Z_{\tau_c+\epsilon\tau_\epsilon} - Z_{\tau_c}\right] + \epsilon\left[D_\epsilon Z_{(\tau_c+\epsilon\tau_\epsilon)} + F(Z)\right]. \tag{5.9}$$

A double Hopf bifurcation with weak resonance may occur in the system when $\epsilon = 0$.

As an example, we take $\omega_0 = 1$ and $\alpha = \frac{-1+\sqrt{5}}{5\sqrt{2\left(5+\sqrt{5}\right)}}$. From Table 3, a 2:3 resonant double Hopf bifurcation occurs in system (5.2) at $(K, \tau) = (K_c, \tau_c) = (\frac{2}{5}\sqrt{\frac{2}{5+\sqrt{5}}}, 2\pi)$. Thus, the solution of (3.7) can easily be obtained as

$$W(t) = \left\{ \begin{array}{l} p_2 \sin(\frac{4}{5}t) + p_3 \sin(\frac{6}{5}t) + q_2 \cos(\frac{4}{5}t) + q_3 \cos(\frac{6}{5}t) \\ -p_2 \cos(\frac{4}{5}t) - p_3 \cos(\frac{6}{5}t) + q_2 \sin(\frac{4}{5}t) + q_3 \sin(\frac{6}{5}t) \end{array} \right\}. \tag{5.10}$$

It follows from (3.5) and (3.6) that the approximate solution of (4.11) can be expressed as

$$Z(t) = \left\{ \begin{array}{l} a_2 \sin((\frac{4}{5}+\epsilon\sigma_1)t) + a_3 \sin((\frac{6}{5}+\epsilon\sigma_2)t) + b_2 \cos((\frac{4}{5}+\epsilon\sigma_1)t) + b_3 \cos((\frac{6}{5}+\epsilon\sigma_2)t) \\ -a_2 \cos((\frac{4}{5}+\epsilon\sigma_1)t) - a_3 \cos((\frac{6}{5}+\epsilon\sigma_2)t) + b_2 \sin((\frac{4}{5}+\epsilon\sigma_1)t) + b_3 \sin((\frac{6}{5}+\epsilon\sigma_2)t) \end{array} \right\}. \tag{5.11}$$

Substituting (5.9), (5.10), and (5.11) into (3.8) and noting that p_2, p_3, q_2, and q_3 are inde-

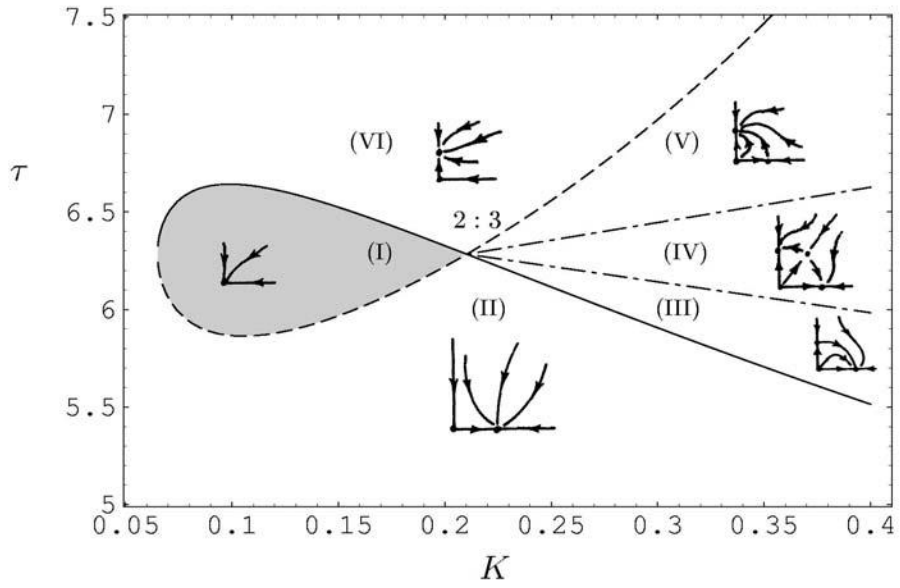


Figure 9. *Classification and bifurcation sets of the solution for the system* (5.2) *due to* 2:3 *resonant double Hopf bifurcation where solid lines, dashed lines, and dot-dashed lines represent boundaries, and amplitude death region is in grey.*

pendent yield

$$
\begin{aligned}
&r_2\left(4\,K_\epsilon+\left(-1+\sqrt{5}\right){r_2}^2+2\left(-1+\sqrt{5}\right){r_3}^2+\sqrt{2\left(5+\sqrt{5}\right)}\sigma_1\right)=0,\\
&r_2\left(25\sqrt{2\left(5+\sqrt{5}\right)}\left({r_2}^2+2{r_3}^2\right)+5\left(5-5\sqrt{5}+16\sqrt{\tfrac{2}{5+\sqrt{5}}}\pi\right)\sigma_1+32\sqrt{\tfrac{2}{5+\sqrt{5}}}\tau_\epsilon\right)=0,\\
&r_3\left(4\,K_\epsilon+2\left(-1+\sqrt{5}\right){r_2}^2+\left(-1+\sqrt{5}\right){r_3}^2-\sqrt{2\left(5+\sqrt{5}\right)}\sigma_2\right)=0,\\
&r_3\left(25\sqrt{2\left(5+\sqrt{5}\right)}\left(2{r_2}^2+{r_3}^2\right)-5\left(5-5\sqrt{5}+16\sqrt{\tfrac{2}{5+\sqrt{5}}}\pi\right)\sigma_2-48\sqrt{\tfrac{2}{5+\sqrt{5}}}\tau_\epsilon\right)=0,
\end{aligned}
\tag{5.12}
$$

where $a_2 = -r_2\sin\theta_1$, $b_2 = r_2\cos\theta_1$, $a_3 = -r_3\sin\theta_2$, $b_3 = r_3\cos\theta_2$, and θ_1 and θ_2 are determined by the initial values. Correspondingly, (5.11) becomes

$$
Z(t)=\begin{Bmatrix} r_2\cos\left(\left(\frac{4}{5}+\epsilon\,\sigma_1\right)t+\theta_1\right)+r_3\cos\left(\left(\frac{6}{5}+\epsilon\,\sigma_2\right)t+\theta_2\right)\\ r_2\sin\left(\left(\frac{4}{5}+\epsilon\,\sigma_1\right)t+\theta_1\right)+r_3\sin\left(\left(\frac{6}{5}+\epsilon\,\sigma_2\right)t+\theta_2\right)\end{Bmatrix},
\tag{5.13}
$$

where r_2, r_3, σ_1, and σ_2 are determined by (5.12), from which the stability can also be analyzed. Thus, all local solutions and their stability in the neighborhood of a double Hopf bifurcation are classified, as shown in Figure 9. The solutions in the same region have the same topological structure. It is easily seen that $(r_{20}, r_{30}) = (0, 0)$ is always a root of (5.12) and the existence of other roots depends on the location of (K, τ). For example, there are four roots in region (IV), given by $(0, 0)$, $(r_{21}, 0)$, $(0, r_{31})$, and (r_{22}, r_{32}), in which the two coexisting

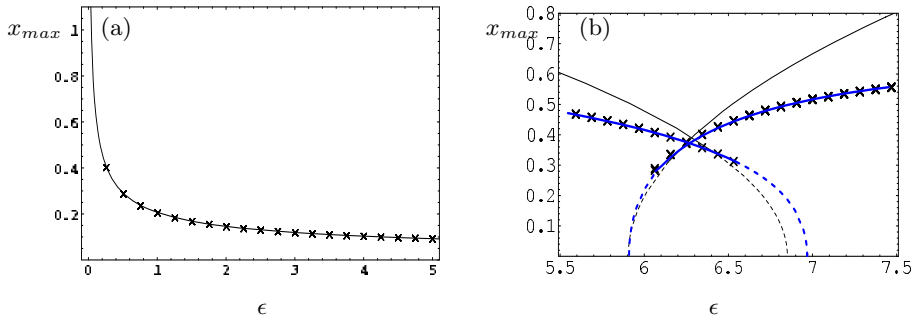


Figure 10. *A comparison among the approximate solution* (5.13) *(solid), the solution from step two of the PIS (thick blue), and the numerical simulation (crossing symbols) in* $\mathrm{Max}(x(t))$ *versus* ϵ *for the periodic solution of system* (4.1) *with* (K,τ) *located in region* (II) *in Figure* 9*:* (a) $K = 0.2$, $\tau = 6.0$, (b) $K = 0.2$, $\tau = 5.0$.

periodic solutions with $(r_{21}, 0)$ and $(0, r_{31})$ are stable, while only one root exists on region (I). All roots can easily be solved from (5.12) for a 2:3 resonant double Hopf bifurcation:

(5.14)

$$
\begin{aligned}
(r_{20}, r_{30}) &= (0,0),\\
(r_{21}, r_{30}) &= (0.0159182\,\sqrt{6751.8\,K_\epsilon - 1067.17\,\tau_\epsilon}, 0),\\
(r_{20}, r_{31}) &= (0, 0.0159182\,\sqrt{6751.8\,K_\epsilon + 1600.76\,\tau_\epsilon}),\\
(r_{22}, r_{32}) &= (0.00919038\sqrt{6751.8\,K_\epsilon + 4268.69\tau_\epsilon}\,,\, 0.00919038\sqrt{6751.8\,K_\epsilon - 3735.11\tau_\epsilon}),\\
\sigma_1 &= 0.1\,(0.878616\,\tau_\epsilon - 16.0735\,K_\epsilon)\,, \quad \sigma_2 = 0.1\,(16.0735\,K_\epsilon + 1.31792\,\tau_\epsilon)\,.
\end{aligned}
$$

It can be obtained from an analysis of the stability that a unique periodic solution exists in all regions except in regions (I) and (IV). The region (I) is a "death island" and there two stable periodic solutions coexist in region (IV). In addition, there is an unstable quasi-periodic solution in region (IV). Similar to the technique described in the previous sections, the stable harmonic solution is investigated to verify the validity and accuracy of step one of the PIS by comparing it with the numerical simulation for system (5.2). Two cases are investigated. The values of the delay and gain are chosen to be, first, close to, and second, far away from the double Hopf bifurcation point at (K_c, τ_c). Figures 10(a) and (b) show the comparisons for the two cases. When (K,τ) is near the double Hopf point, the approximate solution in the analytical form given by (5.11) has good agreement with the numerical simulation even for large ϵ. However, the discrepancy becomes apparent when (K,τ) is away from (K_c, τ_c). Therefore, step one of the PIS is valid and accurate quantitatively if both $|K - K_c|$ and $|\tau - \tau_c|$ are small. It may also provide a fairly accurate qualitative analysis if $|K - K_c|$ or $|\tau - \tau_c|$ is not small. This can be observed in Figures 11(a)–(c), where ϵ is fixed and three different values of K are considered which correspond to the small, medium, and large values of $|K - K_c|$. In Figure 11(a), as τ is increased across the boundaries near the double Hopf point the system enters the amplitude death island from a periodic motion (region (II)) and then goes back to another periodic motion (region (VI)) again. For the two cases of $K = 0.3$ and $K = 0.4$, when

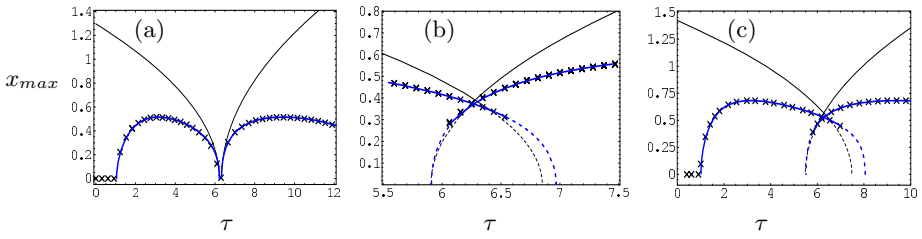


Figure 11. *A comparison between the approximate solution of* (5.13) *(solid), the stable solution from step two of the PIS (thick blue), the unstable solution from step two of the PIS (thick dashed lines), and the numerical simulation (crossing symbols) in* Max(x) *versus* τ *when* τ *transverses from region* (II) *to* (VI) *in Figure* 9 *for* (a) $K = 0.2$, (b) $K = 0.3$, (c) $K = 0.4$, *where* $\epsilon = 1.0$.

τ is in the regions (III)–(V), the system has two coexisting periodic motions with different amplitudes and frequencies. These two periodic solutions are stable in region (IV) (see Figures 11(b) and (c)). Moreover, it follows from Figure 11 that step one of the PIS is able to predict the harmonic motions of the system qualitatively, although it is not accurate enough quantitatively to represent these motions when (K, τ) is far away from the double Hopf bifurcation point. In particular, it is difficult to find the initial guesses for the coexisting solutions if one employs the IHB method. The PIS overcomes this disadvantage as the approximate solution derived from step one can provide such a guess for subsequent continuation.

Next, we use step two of the PIS to trace those periodic solutions at values far away from (K_c, τ_c). To this end, we first choose $\epsilon = 1$ and $(\tau, K) = (6.3, 0.3)$ which is closed to (A_c, τ_c) as a starting point to illustrate the process of the application of the PIS. It follows from Figure 9 that this point is located in region (IV) and there are two coexisting stable periodic solutions given by

$$Z(t) = \begin{Bmatrix} r_{21}\cos((\frac{4}{5}+\sigma_1)\,t+\theta_1) \\ r_{21}\sin((\frac{4}{5}+\sigma_1)\,t+\theta_1) \end{Bmatrix}, \tag{5.15}$$

and

$$Z(t) = \begin{Bmatrix} r_{31}\cos((\frac{6}{5}+\sigma_2)\,t+\theta_2) \\ r_{31}\sin((\frac{6}{5}+\sigma_2)\,t+\theta_2) \end{Bmatrix}, \tag{5.16}$$

respectively, where r_{21}, r_{31}, σ_1, and σ_2 are determined by (5.14), and $\tau_\epsilon = \tau - \tau_c$, $K_\epsilon = K - K_c$. The solutions (5.15) and (5.16) are in good agreement with the numerical one as depicted in Figure 12(a), but such good agreement disappears with τ varying to $\tau = 6.16$ as shown in Figure 12(b). Thus, step two of PIS is used to trace the real solutions. The initial guess of step two is given by

$$a_0 = 0, \quad a_1 = \begin{Bmatrix} r_{21} \\ 0 \end{Bmatrix}, \quad b_1 = \begin{Bmatrix} 0 \\ r_{21} \end{Bmatrix}, \quad a_j = b_j = 0 \quad \text{for } j = 2, \ldots, 5, \tag{5.17}$$

$$p_0 = \frac{4}{5} + \sigma_1, \quad p_j = q_j = 0 \quad \text{for } j = 1, \ldots, 5, \tag{5.18}$$

$$r_0 = -6.16\left(\frac{4}{5} + \sigma_1\right), \quad r_j = s_j = 0 \quad \text{for } j = 1, \ldots, 5, \tag{5.19}$$

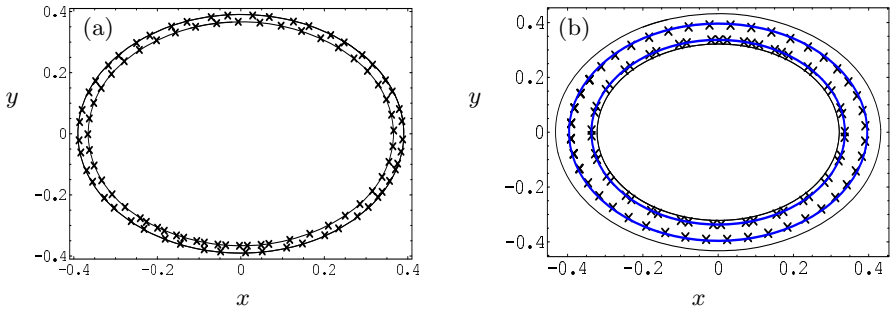


Figure 12. *A comparison between the approximate solution (5.13) (solid), the solution from step two of the PIS (5.23) and (5.24) (thick blue), and the numerical simulation (crossing symbols) in phase plane for the two coexisting periodic solutions of system (4.1), where* (a) $\tau = 6.3$, (b) $\tau = 6.16$, $K = 0.3$, *and* $\epsilon = 1.0$ *throughout.*

and

$$a_0 = 0, \quad a_1 = \begin{Bmatrix} r_{31} \\ 0 \end{Bmatrix}, \quad b_1 = \begin{Bmatrix} 0 \\ r_{31} \end{Bmatrix}, \quad a_j = b_j = 0 \quad \text{for } j = 2, \ldots, 5, \tag{5.20}$$

$$p_0 = \frac{6}{5} + \sigma_2, \quad p_j = q_j = 0 \quad \text{for } j = 1, \ldots, 5, \tag{5.21}$$

$$r_0 = -6.16\left(\frac{6}{5} + \sigma_2\right), \quad r_j = s_j = 0 \quad \text{for } j = 1, \ldots, 5, \tag{5.22}$$

where θ_1 and θ_2 are chosen to be zero. With the incremental procedure from $\tau = 6.3$ to $\tau = 6.3 + n\Delta\tau = 6.16$ ($n \in Z^+$, $|\Delta\tau| \ll 1$) and step two of PIS from (3.26) to (3.42), two converged periodic solutions at $\tau = 6.16$ can easily be obtained after a few iterations, given by

$$\begin{aligned} Z(\phi) &= \begin{Bmatrix} 0.35793 \\ 0 \end{Bmatrix} \cos\phi + \begin{Bmatrix} 0 \\ 0.35793 \end{Bmatrix} \sin\phi, \\ \Phi(\phi) &= 1.29719, \quad \phi_1 = \phi - 7.99074, \end{aligned} \tag{5.23}$$

and

$$\begin{aligned} Z(\phi) &= \begin{Bmatrix} 0.39613 \\ 0 \end{Bmatrix} \cos\phi + \begin{Bmatrix} 0 \\ 0.39613 \end{Bmatrix} \sin\phi, \\ \Phi(\phi) &= 0.71443, \quad \phi_1 = \phi - 4.40092. \end{aligned} \tag{5.24}$$

The stability of the obtained solutions can be analyzed by using (3.41) and (3.42). It follows from Figure 12(b) that the approximate solution derived from the perturbation step of the PIS is far away from the numerical solution but an accurate solution is obtained through step two of the PIS, which shows good agreement with that from the numerical simulation.

Similarly, we can use the PIS to consider the other cases as shown in Figures 10 and 11, where the thick solid lines represent stable solutions and the thick dashed lines instable ones.

6. Discussion and conclusion. We close with a brief discussion of the implication of the results obtained from system (2.1). Many current methods can be employed to classify and describe the dynamics and the bifurcating solutions of systems governed by a set of ODEs [32, 33]. Some of them can be extended to investigate local dynamics of systems governed by a set of DDEs. However, such computation is very tedious [20, 28, 21, 23, 25]. The MMS is valid only for weak nonlinearities and for fixed delay [5, 19]. Therefore, seeking an easy and valid method to investigate delay systems with strong nonlinearity and variable delay has become an open problem. In fact, it is well known that the time delay can lead to various bifurcations and complex dynamics of a delay feedback system. A double Hopf bifurcation is one of these bifurcations. This paper deals with delay-induced resonant double Hopf bifurcations and harmonic solutions in a type of two-dimensional DDE with bifurcation parameters of arbitrary magnitude. It is an extension of the recent investigation [22, 27, 21] of the authors. A new method is proposed to classify those harmonic solutions derived from double Hopf bifurcations, including the periodic, quasi-periodic, and coexisting solutions, both quantitatively and qualitatively. When two parameters vary autonomously in the neighborhood of a double Hopf bifurcation point, the topological structure of the solution and dynamics of the system exhibit genuinely distinct stability types, which correspond to different bifurcation sets. The phase portraits of all cases reduce to the same degenerate one as the two parameters tend to the double Hopf bifurcation point.

For those values of the two parameters close to the double Hopf point, step one of the PIS provides not only a qualitative classification of solutions arising from the double Hopf bifurcation, but also produces an accurate analytical expression for the bifurcating solutions, such as periodic, quasi-periodic, and coexisting periodic solutions.

For those values far away from the double Hopf point, the perturbation step is invalid quantitatively but still valid qualitatively. The PIS developed recently by the authors and their collaborators is extended quantitatively to express the periodic solutions in a closed form. Moreover, the quantitative results obtained by the theorem may be regarded as the initial guess of the incremental step of the PIS to overcome the disadvantage of the IHB method.

The system investigated in this paper consists of two first-order DDEs with variable delay feedback and nonlinearities and has only one limit cycle in the absence of time delay. It is one of the most simple cases. By using this simple system as a model, one can vividly observe the effects of the time delay on it. Physically, the system may be regarded as a basic element integrated into a large scale system. As illustrative examples, two typical systems, namely, the van der Pol–Duffing and the Stuart–Landau systems, are considered to show the advantage of the PIS by comparing the analytical and numerical results. In addition, one can again observe that the time delay can induce a system with one limit cycle to gradually contain quasi-periodic and coexisting periodic motions, as the result of a resonant double Hopf bifurcation derived from the time delay and gain.

In this paper, we treat only weak resonant double Hopf bifurcations. Such bifurcations are codimension-2 phenomena and require two parameters to unfold. It should be noted that a double Hopf bifurcation with 1:2 strong (or low-order) resonance occurs in (4.1) when $\alpha = 0$, $A = -3/8$, and $\tau = 2\pi$. The strong resonant bifurcation requires three parameters to unfold in order to consider the features of solutions near the bifurcation point. It is possible to

find period doubling bifurcations near such a point for a system without the time delay [18]. The possible extension of the PIS to strong resonances and quasi-periodic solutions will be considered in future research.

Acknowledgment. The constructive comments by the anonymous reviewers are gratefully acknowledged.

REFERENCES

[1] F. Giannakopoulos and A. Zapp, *Bifurcations in a planar system of differential delay equations modeling neural activity*, Phys. D, 159 (2001), pp. 215–232.
[2] X. F. Liao, K. W. Wong, and Z. F. Wu, *Bifurcation analysis on a two-neuron system with distributed delays*, Phys. D, 149 (2001), pp. 123–141.
[3] S. Krise and S. R. Choudhury, *Bifurcations and chaos in a predator-prey model with delay and a laser-diode system with self-sustained pulsations*, Chaos Solitons Fractals, 16 (2003), pp. 59–77.
[4] N. MacDonald, *Time Lags in Biological Models*, Lecture Notes in Biomath. 27, Springer-Verlag, Berlin, 1978.
[5] A. H. Nayfeh, C. M. Chin, and J. Partt, *Perturbation methods in nonlinear dynamics—applications to machining dynamics*, J. Manu. Sci. Eng., 119 (1997), pp. 485–493.
[6] A. Raghothams and S. Narayanan, *Periodic response and chaos in nonlinear systems with parametric excitation and time delay*, Nonlinear Dynam., 27 (2002), pp. 341–365.
[7] A. Maccari, *The response of a parametrically excited van der Pol oscillator to a time delay state feedback*, Nonlinear Dynam., 26 (2001), pp. 105–119.
[8] T. Kalár-nagy, G. Stépán, and F. C. Moon, *Subcritical Hopf bifurcation in the delay equation model for machine tool vibrations*, Nonlinear Dynam., 26 (2001), pp. 121–142.
[9] K. Pyragas, *Control of chaos via an unstable delayed feedback controller*, Phys. Rev. Lett., 86 (2002), pp. 2265–2268.
[10] K. Pyragas, *Transmission of signals via synchronization of chaotic time-delay systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 8 (1998), pp. 1839–1842.
[11] D. V. Gregory and R. Rajarshi, *Chaotic communication using time-delayed optical systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 9 (1999), pp. 2129–2156.
[12] D. V. R. Reddy, A. Sen, and G. L. Johnston, *Dynamics of a limit cycle oscillator under time delayed linear and nonlinear feedbacks*, Phys. D, 144 (2000), pp. 335–357.
[13] E. Fridman, *Effects of small delays on stability of singularly perturbed systems*, Automatica J. IFAC, 38 (2002), pp. 897–902.
[14] J. Kalas, and L. Baráková, *Stability and asymptotic behaviour of a two-dimensional differential system with delay*, J. Math. Anal. Appl., 269 (2002), pp. 278–300.
[15] K. Gopalsamy and I. Leung, *Delay induced periodicity in a neural netlet of excitation and inhibition*, Phys. D, 89 (1996), pp. 395–426.
[16] B. F. Redmond, V. G. LeBlanc, and A. Longtin, *Bifurcation analysis of a class of first-order nonlinear delay-differential equations with reflectional symmetry*, Phys. D, 166 (2002), pp. 131–146.
[17] S. A. Campbell and V. B. LeBlanc, *Resonant Hopf-Hopf interaction in delay differential equations*, J. Dynam. Differential Equations, 10 (1998), pp. 327–346.
[18] V. LeBlanc and W. Langford, *Classification and unfoldings of* 1:2 *resonant Hopf bifurcation*, Arch. Rational Mech. Anal., 136 (1996), pp. 305–357.
[19] S. L. Das and A. Chatterjee, *Multiple scales without center manifold reductions for delay differential equations near Hopf bifurcations*, Nonlinear Dynam., 30 (2002), pp. 323–335.
[20] S. A. Campbell, J. Bélair, T. Ohira, and J. Milton, *Limit cycles, tori, and complex dynamics in a second-order differential equations with delayed negative feedback*, J. Dynam. Differential Equations, 7 (1995), pp. 213–236.
[21] J. Xu and K. W. Chung, *Effects of time delayed position feedback on a van der Pol-Duffing oscillator*, Phys. D, 180 (2003), pp. 17–39.

[22] P. Yu, Y. Yuan, and J. Xu, *Study of double Hopf bifurcation and chaos for oscillator with time delay feedback*, Commun. Nonlinear Sci. Numer. Simul., 7 (2002), pp. 69–91.
[23] P. L. Buono and J. Bélair, *Restrictions and unfolding of double Hopf bifurcation in functional differential equations*, J. Differential Equations, 189 (2003), pp. 234–266.
[24] T. Faria and L. T. Magalhães, *Realisation of ordinary differential equations by retarded functional differential equations in neighborhoods of equilibrium points*, Proc. Roy. Soc. Edinburgh Sect. A, 125 (1995), pp. 759–776.
[25] P. L. Buono and V. G. LeBlanc, *Versal unfoldings for linear retarded functional differential equations*, J. Differential Equations, 193 (2003), pp. 307–342.
[26] H. S. Y. Chan, K. W. Chung, and Z. Xu, *A perturbation-incremental method for strongly non-linear oscillators*, Internat. J. Non-Linear Mech., 31 (1996), pp. 59–72.
[27] H. S. Y. Chan, K. W. Chung, and D. W. Qi, *Some bifurcation diagrams for limit cycles of quadratic differential systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 11 (2001), pp. 197–206.
[28] J. Xu and Q. S. Lu, *Hopf bifurcation of time-delay Liénard equations*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 9 (1999), pp. 939–951.
[29] Y. Kuramoto and I. Nishilkawa, *Statistical macrodynamics of large dynamical systems: Case of a phase trasition in oscillator communities*, J. Statist. Phys., 49 (1987), pp. 569–582.
[30] S. Kim, S. H. Park, and C. S. Ryu, *Multistability in coupled oscillator system with time delay*, Phys. Rev. Lett., 79 (1997), pp. 2911–2914.
[31] S. H. Strogatz, *Time delay in the Kuramoto model of coupled oscillator*, Phys. Rev. Lett., 82 (1999), pp. 648–652.
[32] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1993.
[33] A. H. Nayfeh, *Method of Normal Forms*, John Wiley & Sons, New York, 1993.
[34] J. K. Hale and S. M. Verduyn Luncl, *Introduction to Functional Differential Equations*, Springer-Verlag, New York, 1993.
[35] O. Diekmann, S. M. Verduyn Luncl, and H. O. Walther, *Delay Equations: Functional-, Complex-, and Nonlinear Analysis*, Springer-Verlag, New York, 1995.
[36] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag, New York, 1998.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 1, pp. 61–78

A Model for p53 Dynamics Triggered by DNA Damage*

Vijay Chickarmane†, Animesh Ray†, Herbert M. Sauro†, and Ali Nadim†

Abstract. Several recent experiments on DNA-damage-induced signaling networks in mammalian cells have shown interesting dynamics in p53 protein expression during the repair cycle. Pulses of p53 are produced, whose frequency and amplitude are fairly independent of the amount of damage, but the probability of a cell exhibiting this pulsatile behavior increases with damage. This phenomenon has been described as a "digital oscillator." We present here a simple model oscillator comprising two species, p53 and Mdm2, which is activated by the Atm kinase. The Atm kinase exhibits bistable switch-like behavior. The network dynamics essentially consists of the core p53 oscillator, which is turned **ON/OFF** by the Atm switch, which is in turn activated by DNA damage. The complex dynamics are thus explained by the modular nature of the network and are fairly independent of the biological details. A stochastic model of the network dynamics reveals that the pulsatile behavior is robust to intrinsic noise of the protein components and extrinsic noise which arises due to noisy damage signals. The robustness is due to the bistable switch, which makes the system more resilient to stochastic fluctuations in its components. However, the system is more susceptible to noise in the Mdm2 protein production rate.

Key words. mathematical biology, Hopf bifurcation, stochastic, Langevin

AMS subject classifications. 37N25, 92B05

DOI. 10.1137/060653925

The p53 protein is a transcription factor that is present in most higher eukaryotes. It has three major functions: it arrests the cell cycle, thereby giving the cell time to correct any DNA damage, activates transcription of gene indirectly responsible for DNA repair, and can be the cause of apoptosis [1]. The p53 protein is subject to a series of posttranslational modifications that modulate its transcription factor activity [1].

A phosphorylated form of p53, p53-P, is the active transcription factor that binds to a number of target DNA sites that regulate the expression of several genes involved in cell cycle checkpoint arrest and programmed cell death, respectively [2]. Additionally, p53-P induces transcription of the gene p53AIP1, whose product in the cytoplasm leads to the release of cytochrome-c protein from mitochondria, which in turn activates programmed cell death directly by activating the cytoplasmic protease cascade [3]. The p53 protein regulates itself through its interaction with Mdm2. The phosphorylated form of p53 is a transcription factor for the Mdm2 gene, whose protein product targets p53 to degrade by the process of ubiquitination [4]. p53 is a target of the Atm kinase, a protein frequently abnormal in the inherited human disease Ataxia Telangiectasia, which causes the phosphorylation of p53 [5], thus mak-

*Received by the editors March 8, 2006; accepted for publication (in revised form) by J. Collins October 5, 2006; published electronically February 9, 2007. This work was supported by grants from the National Science Foundation (EIA 0205061 and FIBR 0527023 to A. R.; 0432190 and FIBR 0527023 to H. M. S.). V. C. and H. M. S. are grateful for generous support from DARPA/IPTO BioComp program, contract number MIPR 03-M296-01.
http://www.siam.org/journals/siads/6-1/65392.html
†Keck Graduate Institute, 535 Watson Dr., Claremont, CA 91711 (vchickar@kgi.edu, aray@kgi.edu, hsauro@kgi.edu, nadim@kgi.edu).

ing p53 an effective transcription factor. Ionizing radiation generates DNA double strand breaks that trigger the kinase activity of Atm [6, 7, 8]. Hence if DNA damage occurs, Atm is activated, which in turn activates p53 and Mdm2.

Recently, Lahav et al. [9] measured intracellular concentration of total p53 and Mdm2 proteins by fusing their coding sequences to fluorescent reporter domains. Examination of single human cells following treatment with ionizing radiation that produces DNA double stranded breaks (DSB) revealed that p53 and Mdm2 protein concentrations in single cells oscillate in response to DNA damage. The period and the amplitude of these oscillations remained fairly constant over a range of radiation doses. The number of pulses, however, was found to be proportional to the radiation dose; i.e., the larger the dose, the more pulses were observed. Lahav et al. [9] therefore proposed that the system behaved as a "digital oscillator." These observations led to several interesting models and hypotheses [10, 11].

More recently, Alon and colleagues further improved upon these earlier experiments [12], in which the p53 and Mdm2 protein levels were monitored for several days. These experiments led to new insight about the system. The data showed that many cells exhibited the pulsatile behavior which continued over the entire course of the experiment. In general, it was found that the pulses were fairly regular, but with variation in amplitude. Increasing the amount of damage led to more cells being coerced into oscillations. This leads one to believe that the probability of a cell to exhibit oscillations is proportional to the damage. The authors also developed several models that suggest that basal production rates were responsible for increased variation of the amplitude.

In response to the first set of experiments which showed the digital oscillator behavior, Tyson [13] suggested a novel DNA damage control system. Following DNA damage, the steady state p53 concentration was assumed to pass through a Hopf bifurcation (HB) [14, 15] (as a function of the damage parameter) and the p53 and Mdm2 levels begin to oscillate. This would lead to a halt of the cell cycle, giving time for the machinery of the cell to correct the damage. Once the damage is corrected, the system is pulled back from the oscillatory region to its original steady state. This idea was further expanded into an explicit model that took into account the regulation of p53 degradation by Mdm2 and p53 as a transcription factor for Mdm2 RNA synthesis [10]. A crucial aspect of this model was a positive feedback loop by which p53 inhibits the localization of Mdm2 into the nucleus through its interaction with PTEN [16], thereby increasing the p53 levels. This positive feedback leads to bistable behavior which, when coupled to the negative feedback loop (transcription of Mdm2 by p53) results in a relaxation oscillator. This model differed from earlier models which described these oscillations by including a sufficiently long delay using negative feedback, as for example in the case of a hypothetical intermediate step introduced by Bar-Or et al. [17].

Another approach to modeling the p53 dynamics makes explicit use of delays in the system corresponding to the time it takes for transcription and translation of proteins [18, 19, 20]. Using this approach, a recent model [11] described the p53 oscillations arising from the transcriptional delays in a p53-Mdm2 circuit, through its interaction with Atm. In this model the Atm protein acts as a damage sensor, signaling the presence of DNA damage, which effectively turns on the p53 oscillator. The model is described by delay differential equations, where the delay is ascribed to the time required for transcription of Mdm2. DNA repair is modeled as a stochastic process which leads to a variable number of p53 pulses required to correct a given

amount of damage.

Recent experiments [12] that contradict some of the original conclusions on the response of p53 dynamics to the dosage of ionizing radiation have once again opened the doors to ideas regarding the dynamical basis of models which describe the system. Tyson [21] has compared the result of two types of dynamics, i.e., passage of the system through a super-/subcritical Hopf bifurcation, and based upon this has suggested that the variability in the amplitude but steadiness in periodicity may be the signature of the fact that the system traverses a supercritical Hopf bifurcation. In this paper, we would like to argue that the variability in amplitude that arises is due to the stochastic fluctuations in protein numbers, and could be fairly independent of the specific bifurcation if we fix the oscillatory behavior through the occurrence of a bistable switch. The super-/subcritical bifurcations are then equivalent in the sense that the system is rapidly moved from the stable zone into the oscillatory zone independent of the nature of the bifurcation. A strong motivation for a switch to do this task is the Atm sensor. On DNA damage it is known that active Atm molecule concentration rises rapidly and phosphorylates p53. As we will describe in a later section, there is some evidence on the activity of Atm that points to switch-like behavior. Ma et al. [11] use this to suggest that the Atm response is ultrasensitive with respect to the damage.

Our motivation for this paper is twofold: we are interested in describing a generic model that could be used to construct a control system to respond to DNA damage by inducing p53 oscillations, correct the damage, and finally shut off the oscillations; this is much in the spirit of Ciliberto, Novak, and Tyson [10]. The experiments discussed earlier do not necessarily point to a control system to repair the damage, as these cells are derived from cancer cell lines which may already be defective in the repair mechanism. However, we believe it is still important to hypothesize the idea of a general control system, which can be used to describe the correction of damage in normal cells. The second motivation comes from recent experiments which demand that the model must be able to exhibit oscillations with reasonable periodicity, even if it is constantly buffered by stochastic fluctuations. The stochasticity is due to the inherent fluctuations which accompany the birth and death of proteins, and fluctuations due to the damage signal or other noise sources which are external to the core network.

The model we will describe uses a simplified p53-Mdm2 oscillator model which assumes autocatalytic activity in p53. The Atm damage detector is the input into the basic oscillator, which activates the oscillator. We assume that the oscillations correct the damage and ultimately the system is shut off. Our model shares some similarity with that of Ma et al. [11] in that we share the same general structure of a switch determining the onset of oscillations. There are, however, several testable differences. The authors of [11] construct a biologically motivated network that explicitly takes into account delay in the system. Although the idea of using a delay is useful in lumping uncertain details about specific molecular interactions into a few unknowns, which can subsequently be fitted to experimental data, we believe that by exploring novel interactions one can identify testable hypotheses. Our model, therefore, has delays which occur due to the inherent nonlinearity. We also explicitly model the dynamics of the Atm kinase, as a bistable switch, in a different way than in [11]. Bistability turns out to be crucial to our model, and gives rise to pulsatile behavior even in the presence of noise in the damage signaling process. Such behavior would not be possible with ultrasensitivity.

In section 1, we modularize the damage control system and describe the individual basic

components, i.e., the p53-Mdm2 oscillator model, the Atm switch, and the repair module. In section 2, we discuss the integrated model dynamics. The effects of stochasticity are analyzed numerically in section 3, and we finally summarize our conclusions in section 4.

1. Model. Our model network can be modularized into three parts. The first is the p53-Mdm2 oscillator, the second is the Atm activation network, and the third is the repair mechanism which involves both Atm and p53.

1.1. The p53 oscillator. We model the p53-Mdm2 network in a simplistic way by assuming that these are the only two components which interact to give oscillations. This simplification is made to highlight the underlying structure in which the Atm activation module interacts with the p53 oscillator model. In reality it is known that there are several other components, such as the phosphorylated forms of p53, Mdm2, and other genes such as SIAH and ARF [4], which are part of the p53 pathway.

The central biologically known interactions that need to be captured by the model are that p53 activates Mdm2 and that Mdm2, in turn, causes the degradation of p53 [1]. As such, if x and y are used to denote the concentrations of p53 and Mdm2, respectively, the two-variable system can be written in the form

$$\dot{x} = F(x, y), \quad \dot{y} = G(x, y); \tag{1}$$

the functions F and G must have the properties that $G_x > 0$ (for p53 to activate Mdm2) and $F_y < 0$ (for Mdm2 to degrade p53). Here, F_x and G_y refer to partial derivatives. In seeking to obtain a model with possible oscillatory behavior in the form of a limit cycle, a fixed point of the system that resides within the limit cycle needs to be an unstable spiral so that trajectories in the phase plane originating at that fixed point spiral out and asymptote onto the limit cycle. For a two-dimensional system to have an unstable spiral as a fixed point, the Jacobian matrix evaluated at the fixed point must have a trace $T = F_x + G_y$ and a determinant $D = F_xG_y - F_yG_x$ that satisfy [14]

$$T > 0 \quad \text{and} \quad D > T^2/4. \tag{2}$$

Positivity of the trace requires that at least one of the two terms F_x or G_y be positive. This suggests "autocatalysis" by either p53 or Mdm2. In addition, with a positive trace, the second condition in (2) requires that $-4F_yG_x > (F_x - G_y)^2$. For instance, if only p53 autocatalyzes itself ($F_x > 0$, $G_y = 0$), the degree of autocatalysis must not be so large as to exceed $\sqrt{-4F_yG_x}$ and turn the unstable spiral into an unstable node.

In the absence of autocatalysis, one can still get oscillations if p53 also down-regulates Mdm2 or Mdm2 also up-regulates p53, in addition to their normally opposite roles.[1] Consider two types of "motifs," as illustrated in Figure 1, which are discussed in [10]. Figure 1(a) has autocatalysis in p53, whereas in Figure 2(b), in addition to the normal activation of Mdm2 by p53, there is a path by which Mdm2 is down-regulated by p53.

Our model for the p53-Mdm2 oscillator is based upon the former pathway, i.e., via autocatalysis, where some evidence seems to suggest such a process [22]). The p53 oscillator

[1] i.e., $G_x < 0$ and $F_y < 0$, respectively.

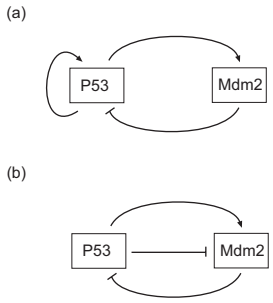


Figure 1. *Two motifs for the* p53-Mdm2 *network.*

model of Ciliberto, Novak, and Tyson [10] is based upon the second motif.[2] This model differs from two previous models [11, 17] in that it involves, in addition to the transcriptional activation of Mdm2 by p53, the negative feedback of Mdm2 on p53 through the process of ubiquitination, a positive feedback loop by which p53 suppresses the nuclear import of phosphorylated Mdm2 [16], which normally ubiquinates p53, thereby causing p53 concentration to rise. This positive feedback leads to the possibility of bistable behavior, which through the transcriptional activity of p53 on the Mdm2 gene results in relaxation oscillations.

With $x(t) = [\mathsf{p53}](t)$ and $y(t) = [\mathsf{Mdm2}](t)$ denoting the dimensionless concentrations, we propose the following biologically motivated system of differential equations to model the oscillator:

$$\begin{aligned} \dot{x} &= \alpha_0 + \alpha_1 x^n/(k_1 + x^n) - \gamma_1 xy - \gamma_2 x, \\ \dot{y} &= \alpha_2 + \alpha_3 x^4/(k_2 + x^4) - \gamma_3 y. \end{aligned} \tag{3}$$

In the first equation for $d[\mathsf{p53}]/dt$, the first term, α_0, represents the production rate of p53. It is known that Atm phosphorylates p53 and its active form is what transcribes Mdm2. This term is therefore ascribed to Atm and also includes a basal production level of p53. The second term represents an autocatalytic process which we assume exists due to positive feedback of p53 on itself [22] and is described with a Hill coefficient n. The third term on the right-hand side represents the active process of ubiquitination of p53 by Mdm2 [1], and the fourth term represents the degradation of p53 independently of Mdm2. Likewise, in the second equation for $d[\mathsf{Mdm2}]/dt$, the first term describes a basal production level of Mdm2, and the second term represents the activation of Mdm2 by p53, described by a Hill equation with exponent 4, since it is known that p53 binds as a tetramer [23]. The last term represents the degradation of Mdm2. In Figure 2 we show nullclines for the system of equations (3), for two different parameter sets. The two cases as we will show correspond to two different bifurcation properties. Panel A shows a stable limit cycle around an unstable state with $[\mathsf{p53}] \simeq 1.7$ and $[\mathsf{Mdm2}] \simeq 2$ for which we have $n = 4$, $\alpha_0 = 1$, $\alpha_1 = 15$, $\alpha_2 = 0.05$, $\alpha_3 = 15$, $k_1 = 80$, $k_2 = 80$, $\gamma_1 = 0.75$, $\gamma_2 = 0.001$, and $\gamma_3 = 0.75$. Panel B shows a stable limit cycle around an unstable state with

[2]In the appendix we describe an oscillator model similar to the motif in Figure 1(b). The model is qualitatively similar to the more refined model of [10] in that it involves a positive feedback loop. The nature of the bifurcations that we obtain are, however, different, since we obtain supercritical Hopf bifurcations, whereas the authors in [10] obtain homoclinic bifurcations.

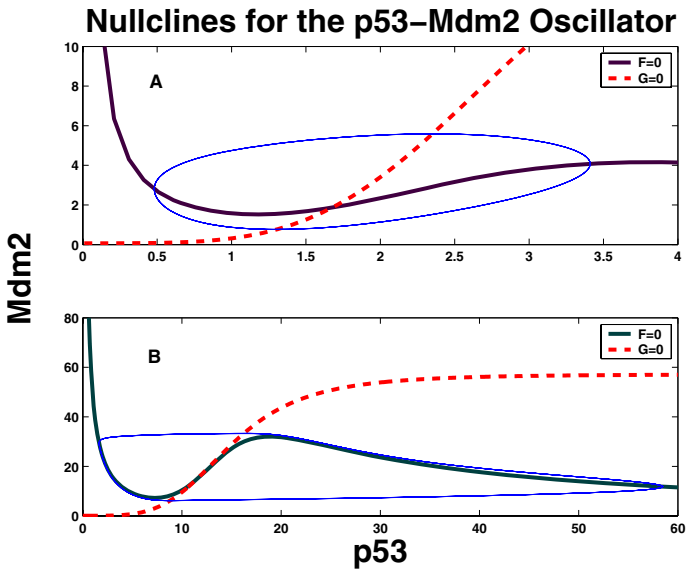


Figure 2. *Nullclines for the two sets of parameter values for the* p53 *oscillator.* $\dot{x} = d[\mathsf{p53}]/dt = 0$*: thick line.* $\dot{y} = d[\mathsf{Mdm2}]/dt = 0$*: dashed line. The thin line in both plots is the stable limit cycle.*

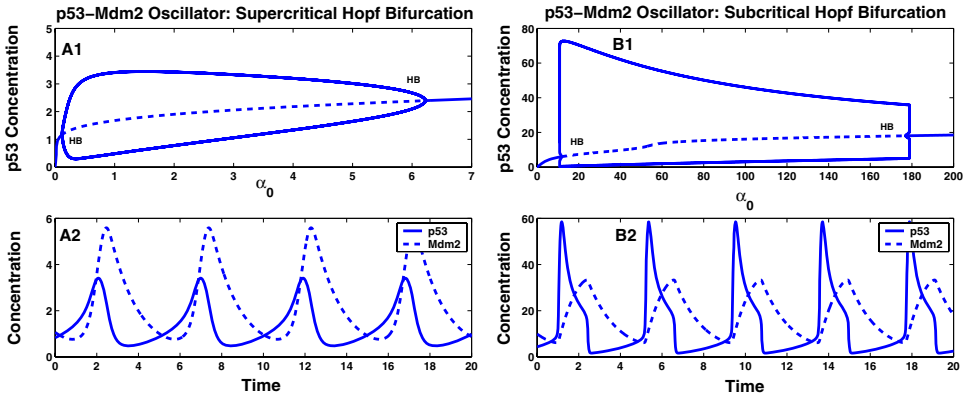


Figure 3. *Bifurcation and time-series plots of the* p53 *oscillator model for the two different parameter sets. Panel* A1 *shows two supercritical Hopf bifurcations occurring at* $\alpha_0 \simeq 0.1, 6.25$*. Panel* B1 *shows two subcritical Hopf bifurcations occurring at* $\alpha_0 \simeq 10, 180$*. The time-series plot for* p53*,* Mdm2 *in* A2 *are obtained by setting* $\alpha_0 = 1$*, and in* B2 *by setting* $\alpha_0 = 50$*.*

$[\mathsf{p53}] \simeq 15$ and $[\mathsf{Mdm2}] \simeq 20$ for which we have $n = 6$, $\alpha_0 = 50$, $\alpha_1 = 700$, $\alpha_2 = 0.05$, $\alpha_3 = 40$, $k_1 = 10^7$, $k_2 = 5\ 10^3$, $\gamma_1 = 1$, $\gamma_2 = 1$, and $\gamma_3 = 0.7$. We should mention at this point that the choice of parameters is not obtained from experimental data but merely serves to describe the dynamics. The nullclines can be manipulated to obtain a specific condition, which is that the slope of the lines which intersect should both be positive for a positive trace and hence lead to oscillatory behavior. In Figure 3 we plot the steady state values of p53 as a function of the bifurcation parameter α_0, and the time-series plots of p53 and Mdm2 for the two parameter sets described. The bifurcation parameter is chosen to be α_0, since this is later related to the

strength of the Atm signal which moves the system from one stable state into another. In both plots, the upper graph shows the bifurcation diagram for p53 levels as a function of the parameter α_0. Panel A1 shows two supercritical Hopf bifurcations. Setting $\alpha_0 = 1$, we obtain p53-Mdm2 oscillations as shown in panel A2. The oscillations show the two protein levels as out of phase, which is expected since the basic network contains an activation-inactivation loop. For the second parameter set, panel B1 shows two supercritical Hopf bifurcations. In panel B2 we plot the p53-Mdm2 time series for $\alpha_0 = 50$.

The crucial observation about both oscillators, from a dynamical point of view, is that if we wish to control the onset of oscillations and "fix" the amplitude and frequency of the oscillations, then we must be able to quickly turn **ON** the value of the bifurcation parameter α_0 such that it is within the oscillatory region, and also maintain it at that value until such time that these oscillations are required. Once this time has elapsed, the bifurcation parameter can be switched **OFF**. This then brings the system out of the oscillatory region back to its original stable state. This means that in the supercritical case α_0 must be rapidly moved from $0 \to 1$, and in the second case from $0 \to 50$, since we are assuming that p53 levels are low to start with, i.e., before damage has occurred. The rapid movement is required to fix the system into the oscillatory region very quickly so that we obtain oscillations with fixed amplitude and frequency. Ciliberto, Novak, and Tyson [10] point out that in the case of a supercritical bifurcation, the slow increase of the amplitude of the oscillations as a function of the bifurcation parameter would result in a variable amplitude of oscillation as a function of time, as the system is slowly dragged into the oscillatory zone. In fact, Ciliberto, Novak, and Tyson use the "sudden birth" of oscillations property of the homoclinic bifurcation [14, 15] in their model to predict the required digital behavior. Therefore, their model has a switch naturally built into the model. However, as we will argue, a switch has one more interesting property which is that it provides stability, keeping the bifurcation point from wandering.

We will therefore couple this oscillator to a bistable switch, which can move the bifurcation parameter in and out of the oscillatory zone quickly. This can be done by modifying the first equation in (3) to read

$$\dot{x} = \alpha_0 z + \alpha_1 x^n/(k_1 + x^n) - \gamma_1 xy - \gamma_2 x,$$

where we now fix $\alpha_0 = 1$, and where $z(t) = [\text{Atm-P}](t)$ is the switch variable that will be described in the next subsection. In all subsequent calculations we choose the second parameter set, since the oscillations have larger amplitudes, and this allows us to set up a simple stochastic model. However, the results we obtain are fairly independent of the nature of the bifurcation, since the switch rapidly moves the system into and out of the oscillatory zone.[3]

1.2. The Atm kinase switch. Atm proteins are normally sequestered in the form of dimers/multimers and are thought to undergo autophosphorylation [8, 24]. A class of phosphatases, PP-2A, are known to dephosphorylate Atm-P [25]. When a double strand break occurs, two events happen very quickly. The first is that the phosphatases dissociate from

[3]The alternative oscillator model described in the appendix could also be used as the p53 oscillator model, since a prudent choice of the bifurcation parameter can once again be made to move the system into and out of the bifurcation zone.

Atm multimers which results in an increased net phosphorylation rate, and the second is that the Atm dimers dissociate to form activated monomers. It has been hypothesized that "turning off" the suppressive effect of the phosphatases on the autophosphorylation of Atm allows rapid signal transduction of the DNA damage signal [25]. Upon DNA damage, Atm-P accumulates rapidly in the nucleus [8]. We assume that the effect of ionizing radiation is to reduce the ability of the phosphatases to bind to Atm. This reduction in the phosphatase activity following DNA damage could be thought to occur due to competition among Atm-P and the signal Dmg (DNA Damage) to bind to the phosphatases or some other proteins that recruit PP-2A. Atm-P phosphorylates Nbs1, which is part of the MRN complex, and that the MRN complex further increases the activity of Atm [26, 27]. We take this as a cue to hypothesizing the interaction between Atm and Nbs1 as a positive feedback system. As such, active Atm (i.e., phosphorylated Atm) will activate Nbs1 in the MRN complex (through phosphorylation), leading to further activity in Atm. This explains why Atm-P levels increase rapidly in the cell upon DNA damage. Thus, Nbs1-P positively regulates activity of Atm by causing its phosphorylation. Alternatively, one can regard this effect to be due to a further activation of Atm-P by Nbs1-P. The preceding discussion motivates a simple model for Atm into which is built a positive feedback loop wherein the phosphorylated form of Atm-P promotes further phosphorylation of Atm (this actually occurs through Nbs1, but for simplifying the model we do not introduce Nbs1). Atm therefore acts as a kinase promoting phosphorylation of itself. Hence once a small amount of Atm-P is produced, it leads to further increase in Atm-P, until all the Atm is converted into its phosphorylated form. This should occur only when significant DNA damage occurs. This process can be modeled by the equation

$$\dot{z} = \frac{\alpha_{1s} z w}{(k_{1s} + w)} - \frac{\alpha_{2s}}{(k_{0d} + \mathsf{Dmg})} \frac{z}{(k_{2s} + z)}, \tag{4}$$

where $z = [\mathsf{Atm\text{-}P}]$ and $w = [\mathsf{Atm}]$ and w can be obtained by using the conservation equation $\mathsf{Atm} + \mathsf{Atm\text{-}P} = 50$ so that $w = 50 - z$. Here, Dmg denotes the DNA damage signal, representing the amount of DNA damage which exists in the cell. The kinetic constants for this equation are chosen to be $\alpha_{1s} = 1$, $k_{1s} = 0.05$, $\alpha_{2s} = 50$, $k_{2s} = 50$, and $k_{0d} = 0.1$. The second term on the right-hand side of the equation for $\dot{z} = d[\mathsf{Atm\text{-}P}]/dt$ describes the dependence of the rate of dephosphorylation as a function of the damage level. We assume that ionizing radiation, which causes DNA damage, also reduces the ability of phosphatases to bind to Atm. Hence when the DNA damage signal is high, the rate of dephosphorylation is low, which allows Atm-P to reach high levels rapidly. When the damage starts to decline, the dephosphorylation rate increases, Atm-P begins to dephosphorylate, and subsequently Atm reaches high levels. In Figure 4 the upper plot shows the bifurcation plot of Atm-P as a function of Dmg. There are two saddle-node bifurcations which make this a bistable switch. For low values of damage, the system's stable condition is a low value of Atm-P, and all of the kinase is in its unphosphorylated form. When the damage crosses the threshold $\mathsf{Dmg} \simeq 0.9$, there is jump in the Atm-P level. If subsequently the damage is slowly reduced, the switch remains **ON** until $\mathsf{Dmg} \simeq 0.42$. At this second threshold Atm-P gets fully unphosphorylated, and Atm levels rise. In the lower panel we show these dynamics in time, assuming that the damage is reduced exponentially in time. Hence in this model for the Atm kinase activity we couple the DNA damage signal by making it the bifurcation parameter.

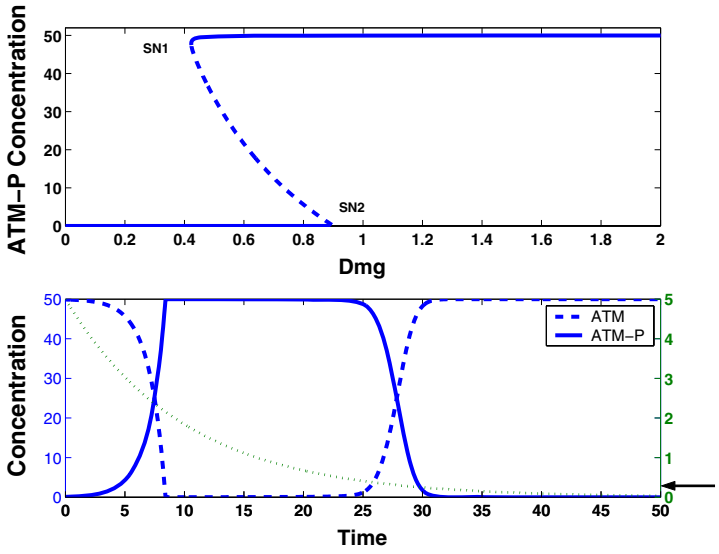


Figure 4. *Bifurcation and time-series plots of the* Atm *kinase activity. The upper panel shows a bifurcation plot of the* Atm-P *kinase concentration as a function of the DNA damage signal* Dmg. *There are two saddle-node bifurcations,* SN1 *and* SN2, *corresponding to the thresholds* $\mathsf{Dmg} \simeq 0.9, 0.42$ *at which the switch is turned* **ON/OFF** *respectively. The lower plot shows the* Atm *and* Atm-P *dynamics as a function of time, assuming that the damage (dotted line) which starts out at a high level decreases exponentially in time:* $\mathsf{Dmg}(t) = 50\exp -0.1t$. *The* Atm-P *level which switches* **ON** *initially is ultimately switched* **OFF** *after the damage signal (dotted line) reduces below the threshold* 0.42 *(indicated by an arrow).*

Finally, we assume that the DNA is repaired by the combined actions of Atm-P and p53, as described by

$$\frac{d[\mathsf{Dmg}]}{dt} = -\alpha_d[\mathsf{Dmg}][\mathsf{Atm\text{-}P}][\mathsf{p53}] = -\alpha_d[\mathsf{Dmg}]zx, \tag{5}$$

where $\alpha_d = 2.510^{-4}$ is the rate of repair in dimensionless units.

2. Integrated model. As shown in Figure 5, there are three functional modules which interact with one another through the following: An initial DNA damage stops the process of dephosphorylation of Atm-P, thereby increasing its levels rapidly; increased levels of Atm-P activate p53, and this turns **ON** the p53-Mdm2 oscillator; the p53 oscillations lead to the start of DNA damage repair, and this gradually reduces the damage signal, Dmg; finally, the rate of dephosphorylation begins to increase, the Atm switch is turned **OFF**, and the p53-Mdm2 oscillator is shut down.

In Figure 6 we plot p53, Mdm2, Dmg, Atm-P, and Atm levels for an initial DNA damage signal $\mathsf{Dmg} = 10$. In the upper plot, after a short time interval, the Atm switch is turned **ON**, and the p53 oscillations occur. As the damage decreases, Atm-P decreases rapidly at a later time, and the system assumes the steady state, i.e., low values of p53, Mdm2 (in all simulations, the initial conditions are chosen to be the final values of all the protein concentrations, after any initial damage is repaired). In the lower panel are displayed the Mdm2 and Atm time-series plots that show that, after the repair, Mdm2 stops oscillating and Atm assumes its steady state

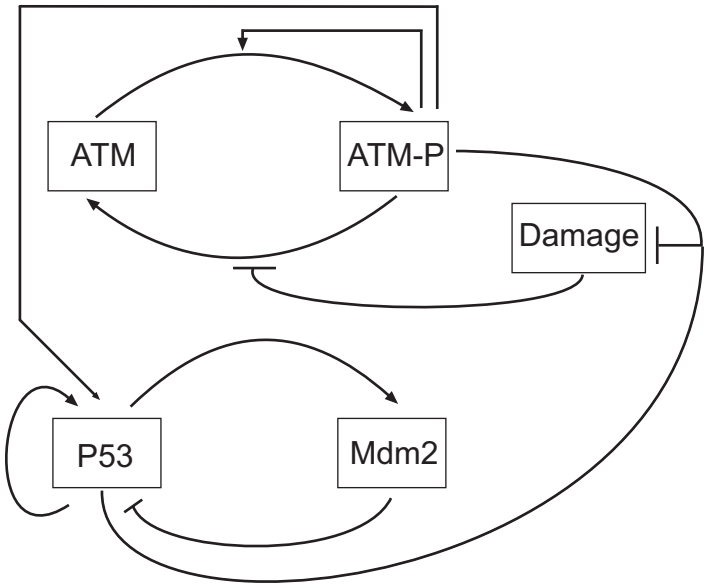


Figure 5. *The core regulatory network. The three modules indicate the* p53 *oscillator, the* Atm *kinase switch, and the DNA damage repair, respectively. These are described in detail in the text.*

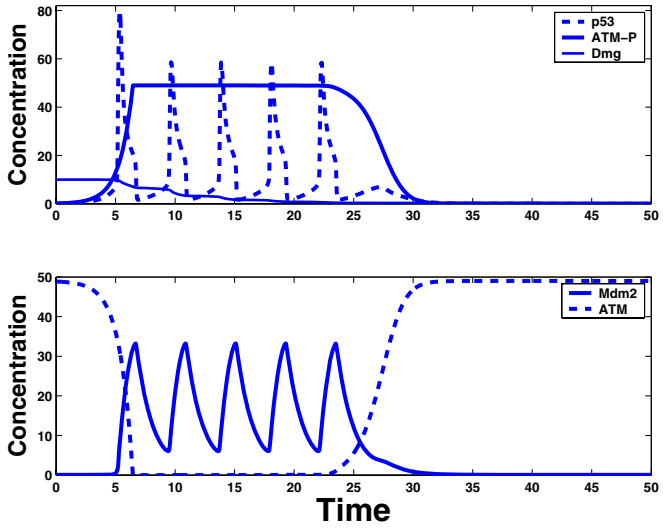


Figure 6. *Time-series plots of the protein concentrations,* p53, Mdm2, Atm, Atm-P, *and* Dmg. *These plots were obtained for an initial concentration of the damage signal* Dmg = 10.

(Atm = 50). The amplitude and frequency of oscillations do not change because the value of Atm-P is fixed at Atm-P = 50 during the repair process. Atm-P increases/decreases rapidly and hence the entry/exit from the oscillatory region for the p53-Mdm2 oscillator is very rapid. This essentially fixes the amplitude and frequency of oscillations.

In Figure 7 we plot p53 for various initial damage levels. The plots show that the number of pulses increases with increasing damage levels. In each case the amplitude and frequency

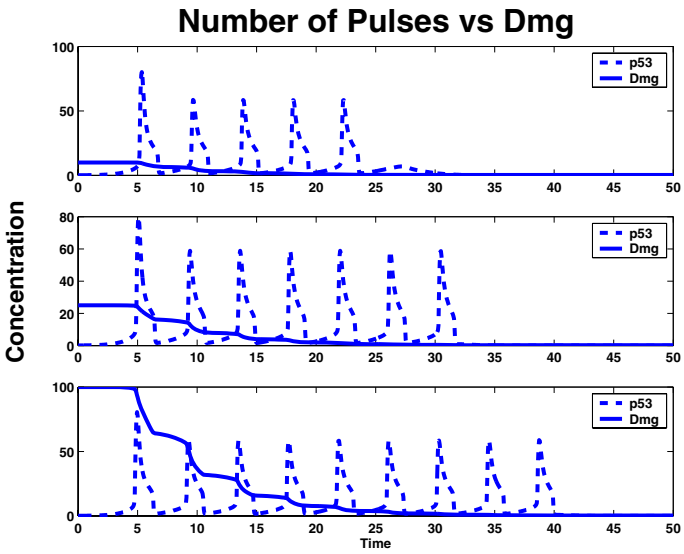


Figure 7. *Number of pulses as a function of initial damage. The plots were obtained for* Dmg $= 10, 25, 100$. *The number of pulses increases with damage levels, though not linearly. A linear relation could be obtained if the rate of repair were made proportional only to the* p53 *levels.*

of oscillations remain fixed. The number of oscillations are monotonic in the damage level. If the rate of repair is made proportional to $-$[p53][Atm] (and independent of Dmg), any initial damage would decrease almost linearly with time, and hence the number of p53 oscillations would become linear in the damage level.

3. The effects of stochasticity. Having explored the damage control system, we now describe the effects of stochasticity on the switch-oscillator system. We assume that there exists a fixed amount of damage which causes the network to exhibit pulsatile behavior. We are not so much interested in the control system in this section but rather on how the network functions for a fixed amount of damage, in the presence of noise. Stochastic effects are naturally inherent in biochemical networks due to the randomness involved in chemical reactions. In gene regulatory networks, the stochasticity is ascribed to slow promoter dynamics and low copy numbers of transcription factors (the randomness is related to the birth and death of protein numbers) [28]. The noise is naturally divided into two components: intrinsic noise due to the fluctuations of species comprising the network and extrinsic noise due to external factors [29] such as temperature fluctuations and noise in signaling events. A complete treatment of this problem would involve deriving a master equation, which in most cases is difficult to solve analytically, and hence we resort to Monte Carlo simulations, i.e., Gillespie runs [30]. However, it is possible to make the approximation that the random effects can be added to deterministic equations to simulate the noisy environment, leading to a Langevin treatment [31, 32, 33]. We supplement the deterministic equations with noise terms, which are Gaussian white noise with variance equal to the mean value of the individual rates contributing to the Langevin equation. For example, the Langevin equation for a simple birth-death process for protein p would be

$$\frac{dp}{dt} = A - \lambda p + \sqrt{A + \lambda p}\ \eta(t), \tag{6}$$

where η is a unit variance, Gaussian white noise which satisfies $\langle \eta(t)\eta(t')\rangle = \delta(t - t')$ and where the individual noise terms due to the birth ($\propto \sqrt{A}$) and death ($\propto \sqrt{\lambda p}$) are lumped together into one noise term. This is because each is an independent Gaussian source of noise, and the sum of independent such sources is Gaussian with variance equal to the sum of the variances. This treatment is formally called the chemical Langevin equation and has been derived for mass action kinetics [31]. Here we assume that under the conditions of rapid promoter dynamics (p53 binds rapidly to Mdm2 relative to their decay rates), the mean field equations can be supplemented by noise terms, whose strength is equal to the sum of the variances of the individual uncorrelated noise terms. We also consider the addition of extrinsic noise sources, particularly in the damage signal, and noise in the basal rate of production of Mdm2. We do not consider explicitly extrinsic noise in p53, since that is already taken care of by assuming noise in the switch that controls the p53 production rate. These assumptions lead to the equations

$$\begin{aligned}
\dot{z} &= \tfrac{\alpha_{1s} z w}{(k_{1s}+w)} - \tfrac{\alpha_{2s}}{(k_{0d}+\mathsf{Dmg})}\tfrac{z}{(k_{2s}+z)} + \sqrt{\tfrac{\alpha_{1s} z w}{(k_{1s}+w)} + \tfrac{\alpha_{2s}}{(k_{0d}+\mathsf{Dmg})}\tfrac{z}{(k_{2s}+z)}}\ \eta_1(t) + D_1\ \eta_2(t),\\
\dot{x} &= \alpha_0 z + \alpha_1 x^n/(k_1 + x^n) - \gamma_1 x y - \gamma_2 x + \sqrt{\alpha_0 z + \tfrac{\alpha_1 x^n}{(k_1+x^n)} + \gamma_1 x y + \gamma_2 x}\ \eta_3(t),\\
\dot{y} &= \alpha_2 + \alpha_3 x^4/(k_2 + x^4) - \gamma_3 y + \sqrt{\alpha_2 + \tfrac{\alpha_3 x^4}{(k_2+x^4)} + \gamma_3 y}\ \eta_4(t) + D_2\ \eta_5(t),
\end{aligned} \tag{7}$$

where $\langle \eta_i(t)\eta_j(t')\rangle = \delta_{ij}\delta(t - t')$, the noise terms η_1, η_3, and η_4 are due to the intrinsic noise of Atm-P, p53, and Mdm2, respectively, and η_2 and η_5 are due to extrinsic noise sources in the damage signal and Mdm2 basal rate of production, respectively. The above equations can be simulated using the Euler or implicit Euler method [33], and a typical realization is shown in panel A, Figure 8. For this plot, Dmg = 10, and $D_1 = D_2 = 0$; i.e., there is no extrinsic noise, and hence the fluctuations we see in Atm-P, Mdm2, and p53 are due to intrinsic noise. The Atm-P levels rapidly turn **ON**, and as soon as the system crosses the Hopf bifurcation, oscillations emerge. Note the fairly regular periodicity of the oscillations. For large damage, Atm-P is switched **ON**, and any intrinsic noise in Atm-P cannot pull it down from its high level, due to the positive feedback. The bistable switch tends to stabilize the input to the system. At the beginning of a cycle, fluctuations in p53 lead to a sudden increase, due to positive p53 autoregulation. This in turn leads to Mdm2 production, which ultimately begins to degrade p53 due to ubiquitination. At this point p53 levels fall, and this cuts off the source for Mdm2, which subsequently decays to small levels. What is crucial is that during the decay, further fluctuations of p53 are ineffective at restarting the process; it must wait for Mdm2 levels to fall below the noise, at which point the next cycle gets started. This is a robust mechanism which shows how oscillations with fairly regular periodicity can be achieved in the presence of noise. We have performed several Monte Carlo simulations, where for each run we collect the pulse delay time between neighboring oscillations of p53 and Mdm2, as well as the p53 pulse amplitude. In the lower panels B and C, we plot histograms of these, which show the coefficient of variation (CV), defined as the ratio of the standard deviation and the mean, to be roughly 30%, for each. While these numbers are different from those reported

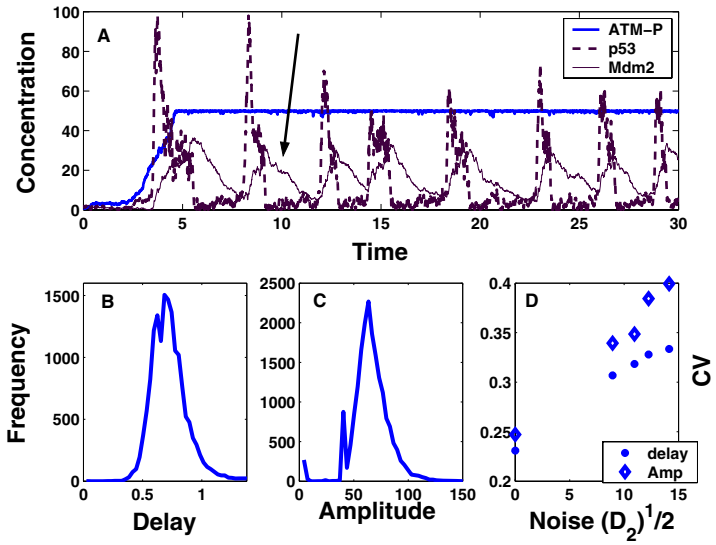


Figure 8. *Panel* A *shows a typical realization of* p53, Mdm2, *and* Atm-P *oscillations. The arrow marks the time interval when* Mdm2 *is in the decay mode, since it has reduced* p53 *levels considerably and hence is not being produced anymore. However, its concentration is so high that subsequent stochastic fluctuations of* p53 *are not sufficient to start the next cycle. Hence the system waits until* Mdm2 *levels are so low that a fluctuation of* p53 *kick-starts the next pulse using its positive autoregulation. Panels* B *and* C *are histograms for the delay between neighboring pulses of* p53 *and* Mdm2 *and amplitude of* p53. *Panel* D *is the result of Monte Carlo simulations for the delay and amplitude CV's for increasing values of the noise in the* Mdm2 *basal production rate.*

in [12], where the authors find a larger CV in amplitude than in the delay, our model is highly simplified and hence we do not expect it to match the experimental results exactly but only to reproduce some common features. We have performed several Monte Carlo runs for different values of D_1, extrinsic noise in the switch due to damage signaling. These show robustness of the oscillations. The robustness can be explained by the stabilizing influence of the positive feedback in the bistable switch. More interesting is the noise in the Mdm2 basal production rate, D_2. In panel D, we plot the CV for the delay and amplitude of oscillations, which show a small increase in the CV for both as a function of the noise strength D_2. Increasing the Mdm2 basal rate delays the onset of p53 oscillations, and only a rare, large fluctuation can get the system to oscillate. This leads to both increase in delay as well as more variation in amplitude. The extrinsic noise in Mdm2 could arise from other protein components which we have not considered in the model.

The regularity in the periodicity of the oscillations is, as we have argued, due to the bistable nature of the switch. This can be seen when we compare the effects of extrinsic noise in the damage signal on the network with a bistable switch as compared to an ultrasensitive switch. We would expect that extrinsic noise in the Dmg signal would move the bifurcation point up and down the steady state curve for an ultrasensitive switch. An example of such a switch is shown in panel A, Figure 9. The arrow marks the point at which the system crosses into the oscillatory zone. To obtain this curve we assume that the deterministic equation for Atm-P follows

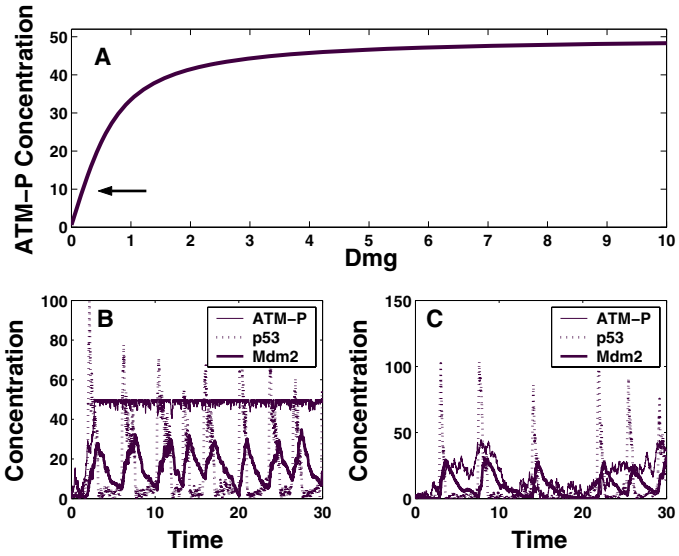


Figure 9. *Panel* A *shows an ultrasensitive switch with respect to the* Dmg. *In panels* B *and* C*, time series of* p53, Mdm2, *and* Atm-P *are shown with extrinsic noise,* $D_1 = 10$, Dmg $= 10$, *for the cases of the bistable switch and ultrasensitive switch, respectively.*

$$\frac{dz}{dt} = \frac{\alpha_{1s} w}{(k_{1s} + w)} - \frac{\alpha_{2s}}{(k_{0d} + \mathsf{Dmg})} \frac{z}{(k_{2s} + z)}, \tag{8}$$

where $\alpha_{1s} = 1$, $k_{1s} = 50$, $k_{0d} = 0.01$, $\alpha_{2s} = 1$, $k_{2s} = 100$. Although there is no positive feedback, ultrasensitivity can be achieved by adjusting the parameters α_{1s}, etc. In panels B and C we plot the result of a simulation for Dmg $= 10$, $D_2 = 0$, and $D_1 = 10$, i.e., large noise fluctuations in the damage signal, but no extrinsic noise in Mdm2. Panel B is for the bistable switch, and panel C is for the ultrasensitive switch. As we can see, the network which employs the bistable switch still exhibits fairly robust oscillations, whereas the network with the ultrasensitive switch has larger differences in periodicity. Hence although a deterministic treatment may show that an ultrasensitive switch is enough to move the system rapidly into the oscillatory zone, the stochastic treatment of the same problem shows that the effect of ultrasensitivity is to transduce the input noise into the system and reduce the periodic structure of the oscillations. In contrast, a bistable switch imparts more stability to the system and periodicity is maintained.

4. Conclusions. In this paper we have discussed a generic scheme for a DNA damage response system involving the key players p53, Mdm2, and the damage sensor Atm. By using a modular approach, we coupled a bistable switch to an oscillator, with the switch controlling the oscillations. The switch itself is turned **ON** when DNA damage occurs. Although knowledge of the molecular details of the components of the oscillator are crucial for making any comparison with experimental results, we have shown that the interaction between the bistable switch and the oscillator is one way to obtain pulsatile behavior. We also studied the effects of stochastic fluctuations in a simplified way by treating both internal and extrinsic noise that were added to deterministic solutions. We found that the system exhibited fairly robust oscillatory behavior

in the presence of noise. From the results of Monte Carlo simulations we argued that the robustness arises from the stability provided by the positive feedback in the bistable switch as well as the actual mechanism by which p53 and Mdm2 interact, i.e., through a combination of autocatalysis and Mdm2 mediated degradation. We speculate that these features will still hold independently of the parameter values of the network, although the actual details of the shapes and amplitudes may vary. The important noise contribution that does seem to degrade the robustness of the oscillations was the extrinsic noise in Mdm2. An experiment in which extrinsic noise can be introduced into Mdm2 would reveal further properties of the pulsatile behavior of the p53-Mdm2 loop. Our oscillator model uses autocatalysis in p53 in addition to known interactions such as transcription of Mdm2 by p53 and degradation of p53 by Mdm2. The oscillator model of Ciliberto, Novak, and Tyson [10] differs fundamentally from this treatment, as it inherently contains a bistable property. As the authors point out in [10], under certain conditions (one example could be inefficient transcriptional activity of p53) cells could exhibit one of the two stable p53 concentration levels. This is one prediction that could be used to differentiate between a model with autocatalysis and one with inherent bistability. Our model for the damage network which uses the Atm damage sensor as a bistable switch leads to a sudden rise in Atm-P following DNA damage, and lasts until the damage is corrected. Moreover there is a certain amount of damage which allows one of two stable states [34] to exist; i.e., the system is at a stable state with low values of p53 concentration, or the system is exhibiting oscillations. An experiment which could slowly ratchet up the amount of damage could possibly search for hysteresis. Such experiments [35, 36] have demonstrated hysteresis in Mammalian gene networks. One major prediction that arises from our model based upon the hysteretic effect is due to the Atm switch.

In Figure 10 we plot the steady state values of Atm-P as a function of Dmg for the same set of parameters as for Figure 4, but for $k_{0d} = 0.6$. The increase in this parameter value can be thought of physically as arising due to a reduction of the ability of the phosphatases to bind to Atm. Hence once the switch is turned **ON**, subsequent decrease of damage is not enough for the switch to turn **OFF**, and this leads to one of the saddle-node points crossing the vertical axis, leading to irreversibility in the switch-like behavior. Hence initially once the damage increases beyond the minimum threshold, Atm-P will permanently remain **ON**, even after the damage is completely repaired. This dynamics indicates that if it were possible to introduce into a cell a constitutively active, gain-of-function, mutant Atm kinase gene ("always **ON**" mutant), then the p53 level would oscillate even without DNA damage.

Appendix. An alternative model for the p53 oscillator can also be used; this model, in addition to the activation-inactivation loop, has a process whereby p53 causes repression of Mdm2. This model takes the form

$$\begin{aligned} \dot{x} &= \alpha_0^* - \gamma_1^* xy - \gamma_2^* x, \\ \dot{y} &= \alpha_1^* x^4/(k_1^* + x^4) - \alpha_2^* xy/(k_2^* + y^2) - \gamma_3^* y \end{aligned}$$

with the parameters given by $0.1 < \alpha_0^* < 0.5$, $\alpha_1^* = 3$, $\alpha_2^* = 2.25$, $k_1^* = 8$, $k_2^* = 1.5$, $\gamma_1^* = 0.1$, $\gamma_2^* = 0.001$, and $\gamma_3^* = 0.005$. The mathematical structure of this model is qualitatively similar to the model in Ciliberto, Novak, and Tyson [10], although the details vary (the nature of the Hopf bifurcation in this paper is subcritical). For this model, in Figure 11 we plot the steady

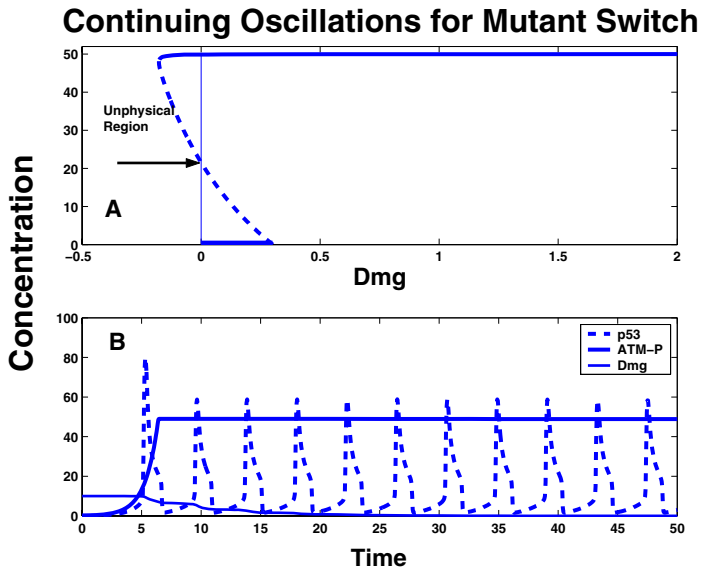


Figure 10. *Steady state values of* Atm-P *as a function of* Dmg. *All parameters are the same as for Figure 4, except* $k_{0d} = 0.6$. *The saddle node on the lower branch crosses the y-axis, and hence this gives rise to irreversibility. Hence, even after any initial DNA damage is repaired,* p53 *can continue to oscillate. This is seen in the lower plot, which shows oscillations even when damage falls to zero.*

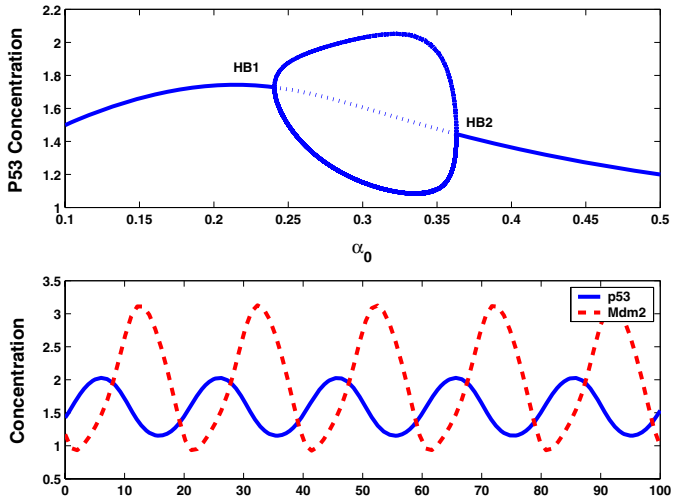


Figure 11. *Bifurcation and time-series plots for* p53, Mdm2 *for the alternative oscillator model. The bifurcation plots of* p53 *as a function of the bifurcation parameter* α_0^* *show two supercritical HB's occurring at* $\alpha_0^* \simeq 0.24, .36$. *The time-series plots for* p53, Mdm2 *are obtained by setting* $\alpha_0^* = 0.3$.

state values of p53 as a function of the bifurcation parameter α_0^*, and the time-series plots for p53 and Mdm2. One could now assume that it is through the term α_0^* that the switch variable Atm-P will enter the equations. The first Hopf bifurcation is supercritical. Hence, this

oscillator can also be coupled to the Atm-P switch in exactly the same way as the oscillator with the autocatalysis to obtain pulsatile behavior by allowing Atm-P to drag the bifurcation parameter into and out of the oscillatory zone. The pulsed behavior is therefore independent of the detailed mechanism used to build the oscillator but is related to the way in which the oscillator is controlled by external means, i.e., through Atm-P.

Acknowledgments. Herbert M. Sauro and Vijay Chickarmane thank Dr. Carsten Peterson and Dr. Michael Kastan for useful discussions.

REFERENCES

[1] V. Vogelstein, D. Lane, and A. J. Levine, *Surfing the p53 network*, Nature, 408 (2000), pp. 307–310.
[2] K. W. Kohn and Y. Pommier, *Molecular interaction map of the p53 and Mdm2 logic elements, which control the off-on switch of p53 in response to DNA damage*, Biochem. and Biophys. Res. Comm., 331 (2005), pp. 816–827.
[3] K. Matsuda, K. Yoshida, Y. Taya, K. Nakamura, Y. Nakamura, and H. Arakawa, *p53AIP1 regulates the mitochondrial apoptotic pathway*, Cancer Research, 62 (2002), pp. 2883–2889.
[4] S. L. Harris and A. J. Levine, *The p53 pathway: Positive and negative feedback loops*, Oncogene, 24 (2005), pp. 2899–2908.
[5] S. Banin, L. Moyal, S. Shieh, Y. Taya, C. W. Anderson, L. Chessa, N. I. Smorodinsky, C. Prives, Y. Reiss, Y. Shiloh, and Y. Ziv, *Enhanced phosphorylation of p53 by ATM in response to DNA damage*, Science, 281 (1998), pp. 1674–1677.
[6] J. H. Lee and T. T. Paull, *ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex*, Science, 308 (2005), pp. 551–554.
[7] R. T. Abraham, *Checkpoint signaling: Epigenetic events sound the DNA strand-breaks alarm to the ATM protein kinase*, BioEssays, 25 (2003), pp. 627–630.
[8] C. J. Bakkenist and M. B. Kastan, *DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation*, Nature, 421 (2003), pp. 499–506.
[9] G. Lahav, N. Rosenfeld, A. Sigal, N. Geva-Zatorsky, A. J. Levine, M. B. Elowitz, and U. Alon, *Dynamics of the p53-Mdm2 feedback loop in individual cells*, Nature Genetics, 36 (2004), pp. 147–150.
[10] A. Ciliberto, B. Novak, and J. J. Tyson, *Steady states and oscillations in the p53-Mdm2 network*, Cell Cycle, 4 (2005), pp. 107–112.
[11] L. Ma, J. Wagner, J. J. Rice, W. Hu, A. J. Levine, and G. A. Stolovitzky, *A plausible model for the digital response of p53 to DNA damage*, Proc. Nat. Acad. Scien., 102 (2005), pp. 14266–14271.
[12] N. Geva-Zatorsky, N. Rosenfeld, S. Itzkovitz, R. Milo, A. Sigal, E. Dekel, T. Yarnitzky, Y. Liron, P. Polak, G. Lahav, and U. Alon, *Oscillations and variability in the p53 system*, Mol. Syst. Biol., 2 (2006), 2006.0033.
[13] J. J. Tyson, *Monitoring p53's pulse*, Nature Genetics, 36 (2004), pp. 113–114.
[14] S. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*, Addison-Wesley, Reading, MA, 1994.
[15] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag, New York, 1995.
[16] L. D. Mayo and D. B. Donner, *The PTEN, Mdm2, p53 tumor suppressor-oncoprotein network*, TRENDS in Biochemical Sciences, 27 (2002), pp. 462–467.
[17] R. L. Bar-Or, R. Maya, L. A. Segel, U. Alon, A. J. Levine, and M. Oren, *Generation of oscillations by the p53-Mdm2 feedback loop: A theoretical and experimental study*, Proc. Natl. Acad. Sci. USA, 97 (2000), pp. 11250–11255.
[18] N. A. M. Monk, *Oscillatory expression of Hes1, p53, and NF-kappaB driven by transcriptional time delays*, Current Biology, 13 (2003), pp. 1409–1413.
[19] G. Tiana, M. H. Jensen, and K. Sneppen, *Time delay as a key to apoptosis induction in the p53 network*, Euro. Phys. J., 29 (2002), pp. 135–140.

[20] J. Wagner, L. Ma, J. J. Rice, W. Hu, A. J. Levine, and G. A. Stolovitzky, p53-Mdm2 *loop controlled by a balance of its feedback strength and effective dampening using* Atm *and delayed feedback*, IEE Proc. Systems Biology, 152 (2005), pp. 109–118.
[21] J. J. Tyson, *Another turn for p*53, Mol. Syst. Biol., 2 (2006), 2006.0032.
[22] V. Benoit, A. C. Hellin, S. Huygen, J. Gielen, V. Bours, and M. P. Merville, *Additive effect between NF-kappaB subunits and p*53 *protein for transcriptional activation of human p*53 *promoter*, Oncogene, 19 (2000), pp. 4787–4794.
[23] K. G. McLure and P. W. K. Lee, *How p*53 *binds DNA as a tetramer*, The EMBO Journal, 17 (1998), pp. 3342–3350.
[24] E. U. Kurz and S. P. Lees-Miller, *DNA damage-induced activation of ATM and ATM-dependent signaling pathways*, DNA Repair, 3 (2004), pp. 889–900.
[25] A. A. Goodarzi, J. C. Jonnalagadda, P. Douglas, D. Young, R. Ye, G. B. Moorhead, S. P. Lees-Miller, and K. K. Khanna, *Autophosphorylation of ataxia-telangiectasia mutated is regulated by protein phosphatase* 2*A*, The EMBO Journal, 23 (2004), pp. 4451–4461.
[26] T. T. Paull and J. H. Lee, *The Mre*11*/Rad*50*/Nbs*1 *complex and its role as a DNA double-strand break sensor for ATM*, Cell Cycle, 6 (2005), pp. 737–740.
[27] J. H. Lee and T. T. Paull, *Direct activation of the ATM protein kinase by the Mre*11*/Rad*50*/Nbs*1 *complex*, Science, 304 (2004), pp. 93–96.
[28] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins, *Stochasticity in gene expression: From theories to phenotypes*, Nat. Rev. Genet., 6 (2005), pp. 451–464.
[29] P. S. Swain, M. B. Elowitz, and E. D. Siggia, *Intrinsic and extrinsic contributions to stochasticity in gene expression*, Proc. Natl. Acad. Sci. USA, 99 (2002), pp. 12795–12800.
[30] D. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, J. Phys. Chem., 81 (1977), pp. 2340–2361.
[31] D. Gillespie, *The chemical Langevin equation*, J. Phys. Chem., 113 (2000), pp. 297–306.
[32] J. Hasty, J. Pradines, M. Dolnik, and J. J. Collins, *Noise-based switches and amplifiers for gene expression*, Proc. Natl. Acad. Sci. USA, 97 (2000), pp. 2075–2080.
[33] D. Adalsteinsson, D. McMillen, and T. C. Elston, *Biochemical Network Stochastic Simulator (BioNetS): Software for stochastic modeling of biochemical networks*, BMC Bioinformatics, 5 (2004), 15113411.
[34] V. Chickarmane, A. Nadim, A. Ray, and H. M. Sauro, *A p*53 *Oscillator Model of DNA Break Repair Control*, http://xxx.lanl.gov/abs/q-bio.MN/0510002.
[35] H. H. Chang, P. Y. Oh, D. E. Ingber, and S. Huang, *Multistable and multistep dynamics in neutrophil differentiation*, BMC Cell Biol., (2006), pp. 7–11.
[36] B. P. Kramer and M. Fussenegger, *Hysteresis in a synthetic mammalian gene network*, Proc. Natl. Acad. Sci. USA, 102 (2005), pp. 9517–9522.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 1, pp. 79–98

Topological Entropy of Braids on the Torus*

Matthew D. Finn† and Jean-Luc Thiffeault†

Abstract. We present a fast method for computing the topological entropy of braids on the torus. This work is motivated by the need to analyze large braids when studying two-dimensional flows via the braiding of a large number of particle trajectories. Our approach is a generalization of Moussafir's technique for braids on the disk. Previous methods for computing topological entropy include the Bestvina–Handel train-track algorithm and matrix representations of the braid group. However, the Bestvina–Handel algorithm is computationally intractable for large braid words, and matrix methods give only lower bounds, which are often poor for large braids. Our method is computationally fast and gives exponential convergence towards the exact entropy. As an illustration we apply our approach to the braiding of both periodic and aperiodic trajectories in the sine flow.

Key words. topological entropy, braid groups

AMS subject classifications. 37B40, 37M25, 20F36

DOI. 10.1137/060659636

1. Introduction.

Investigation of two-dimensional fluid mixing by topological techniques is rapidly gaining popularity [1]. Topological perspectives on mixing involve either studying braiding motion of the stirring apparatus itself [7], or the diagnosis of mixing by analyzing braiding of orbits of the flow [13, 15, 16, 30]. The quantity that is usually of interest is the topological entropy of the braid [6], which serves as a lower bound for the topological entropy of the flow. The topological entropy of the flow is related to the exponential growth rate of material lines [5, 14, 22], which has long been a favorite measure of mixing quality, though it is by no means the only one [12]. In many cases the braid entropy is quite a sharp bound on the flow entropy [11, 13, 16], which is one reason why analyzing braids is useful. Another reason is that experimental particle trajectory data can be found easily by particle image velocimetry, but it is usually very difficult to measure entropies directly from material stretching or by computing Lyapunov exponents.

There are many techniques for calculating braid topological entropies, or lower bounds on them, including train-tracks [2, 17], the Burau representation of the braid group [19], and others [20, 21]. Where braiding of periodic (or aperiodic) orbits is used to analyze a flow, one needs to interpret braids that have both a large number of strands and a large number of generators. In this scenario, exact methods based on train-tracks quickly become prohibitively expensive computationally, and methods based on the Burau matrix representation of the braid group usually give very poor lower bounds. Recently, Moussafir described a fast method

*Received by the editors May 11, 2006; accepted for publication (in revised form) by J. Meiss October 17, 2006; published electronically February 9, 2007. This work was funded by the UK Engineering and Physical Sciences Research Council grant GR/S72931/01.

http://www.siam.org/journals/siads/6-1/65963.html

†Department of Mathematics, South Kensington Campus, Imperial College London, London SW7 2AZ, United Kingdom (matthew.finn@imperial.ac.uk, jeanluc@imperial.ac.uk).

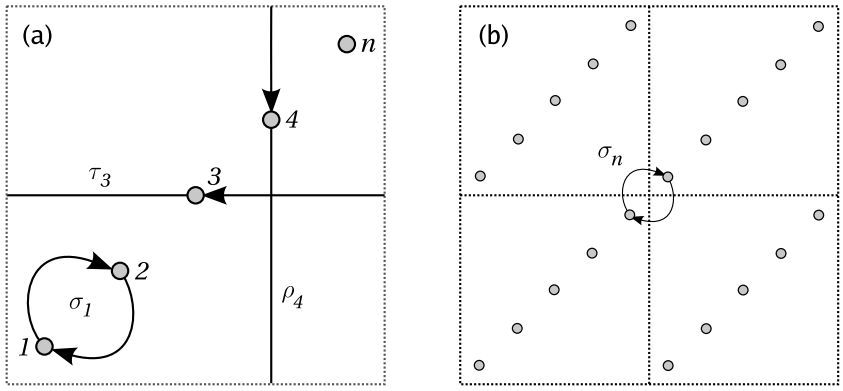


Figure 1. *The spatially periodic domain with n punctures.* (a) *Any motion of the punctures (up to homotopy) can be written as a braid word consisting of a string of σ_i, ρ_i, and τ_i motions and their inverses.* (b) *It is convenient to define the additional operation σ_n, the clockwise exchange of the last and first punctures.*

for calculating the entropy of pseudo-Anosov braids to arbitrary precision [21]. This method is based on a Dynnikov coordinate representation of an integral lamination [9]. There are many other ways of encoding integral laminations, such as triangulation, Dehn–Thurston parametrization [10, 26, 27, 29], or a coding of train-tracks on the hyperbolic disk [18]. (We discuss this further in section 6.1.) In this paper we extend Moussafir's technique to braids on the torus using a triangulation of the surface. We are motivated by the fact that many interesting dynamical systems are defined in periodic (cylindrical or annular) or biperiodic (toroidal) spatial domains. For instance, the alternating sine flow [28] can be analyzed from a topological perspective using this approach [13].

Figure 1 shows the setting for the dynamical system under study. We consider a flow on a torus, so that the domain is periodic in both directions. We identify n distinguished points or punctures. We then consider motions that move the punctures in such a way that they always return to their initial configuration, possibly having been permuted amongst themselves. The three types of motions that we consider are illustrated in Figure 1(a); they are

1. σ_i, the clockwise interchange of the ith and $(i+1)$th puncture,
2. ρ_i, the ith puncture making a full tour around the vertical periodic direction,
3. τ_i, the ith puncture making a full tour around the horizontal periodic direction.

The inverse of any of these motions is obtained by reversing its direction. The elementary motions $\{\sigma_i, \rho_i, \tau_i\}$ are generators of the braid group on n strands on the torus [3, 4]. Recognizing the periodicity of the domain, we also define an additional operation σ_n to be the clockwise interchange of the nth puncture with the first puncture. To be precise, we mean here the first puncture in the "copy" of the domain above and to the right of the nth puncture, as pictured in Figure 1(b), so that both periodic boundaries are crossed in performing σ_n. Defining σ_n in this way keeps both periodic directions on an equal footing and is also convenient in what follows, as it is related to a translational symmetry for the punctures.

A sequence of generators, such as $\rho_3^{-1}\sigma_2\rho_6\sigma_6$, is called a braid word, and we use the convention that the elementary motions in a braid word occur from left to right, so that ρ_3^{-1} occurs first in our example. By *planar braid* we mean a braid word that can be written using only generators from the set $\{\sigma_1, \ldots, \sigma_{n-1}\}$ and their inverses. A planar braid is equivalent

to a braid on the plane with n punctures. In other words, a planar braid does not take advantage of the periodic directions. By *cylinder braid* (or annular braid) we mean a braid word that can be written using only generators from either the set $\{\sigma_1, \ldots, \sigma_{n-1}, \rho_1, \ldots, \rho_n\}$ or $\{\sigma_1, \ldots, \sigma_{n-1}, \tau_1, \ldots, \tau_n\}$ and their inverses. In other words, a cylinder braid takes advantage of one periodic direction but not the other. Finally, a *torus braid* is a braid word that is neither a planar braid nor a cylinder braid. In this paper we will derive a general method for torus braids, which includes planar and cylinder braids as special cases.

The description of our method here is intended to be accessible to dynamicists and requires no specialized understanding of braid groups. The paper is divided as follows. In the next section we describe how integral laminations (equivalence classes of simple closed curves) can be encoded by triangulation of the flow domain. Section 3, the heart of the paper, gives the details of how this encoding evolves under fundamental braiding motions. Several examples are given in section 4 to illustrate and verify the method. In section 5 we show how to compute the braid topological entropy from evolution of integral laminations, and demonstrate the rapid convergence. We also look at braiding in the sine flow as an example application. In section 6 we place our work in the wider context of the topological study of surface homeomorphisms, and we summarize our current work and discuss possibilities for further study.

2. Encoding of integral laminations by triangulation. We wish to calculate a lower bound on how rapidly material lines are stretched in a continuous-time flow based on the motion of a finite number of punctures, or a finite set of periodic orbits. Our approach, inspired by the method of Moussafir [21], is to study the stretching and folding of integral laminations. An integral lamination is an equivalence class (under isotopy) of simple closed curves that are not isotopic to any part of the boundary (treating the punctures as boundaries) and cannot be continuously shrunk to a point. For example, a loop that encloses at least two punctures belongs to an integral lamination. We usually represent the equivalence class by just drawing a typical loop, and when we refer to a loop we will usually mean the isotopy class it belongs to. This will always be clear from the context. An example of a loop on our doubly periodic toroidal domain is shown in Figure 2(a).

In order to calculate how a given loop is transformed under the action of the braid, it is necessary to have a way of encoding the loop, and a method for evolving this encoding under the action of the braid. An elegant way of encoding the isotopy class of a loop is by triangulating the entire domain and counting the number of crossings the loop makes with the edges of the triangulation [29]. The number of crossings for our example is shown in Figure 2(b). Note that there are infinitely many different ways the domain can be triangulated, so in this paper we choose a triangulation that is most convenient for studying the action of braiding motions. The notation we use for counting crossings is illustrated in Figure 3. For n punctures the resulting encoding $\{x_i, y_i, z_i\}$ contains $3n$ crossing numbers.

To ensure that all homotopic curves produce the same set of crossings we insist that the curve is first pulled tight, which means that no loops are allowed where the curve enters and leaves a triangle by the same edge. Under the pulled-tight assumption, the set of crossing numbers then uniquely identifies the integral lamination. The reason for this is that the path of the loop can be determined uniquely in each triangle (up to homotopy), and therefore the global picture of the loop is also unique.

For completeness, we now describe the construction procedure for one triangle. Consider

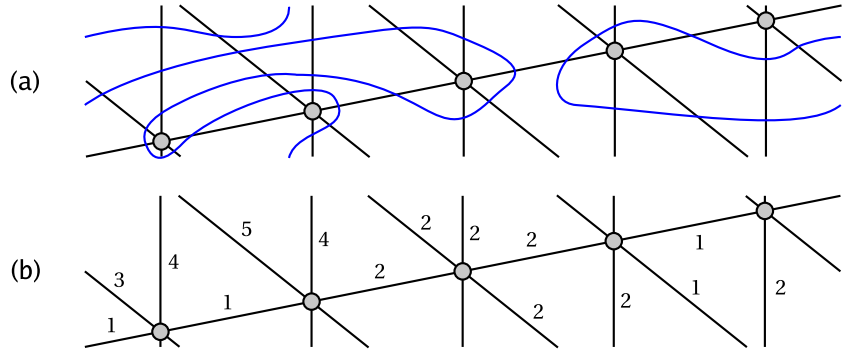


Figure 2. (a) *A representative loop of an integral lamination on a toroidal domain with five punctures. The top and bottom edges, and the left and right edges, are identified so that the picture shows exactly ten triangles.* (b) *The number of crossings with each edge of the triangulation, from which the lamination can be reconstructed.*

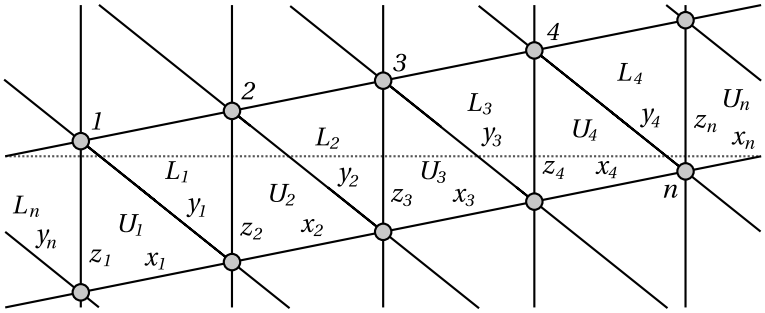


Figure 3. *Notation for our chosen triangulation of the doubly periodic domain with n punctures. The domain is divided into $2n$ triangles U_i and L_i. Two copies of the domain are shown one above the other for clarity.*

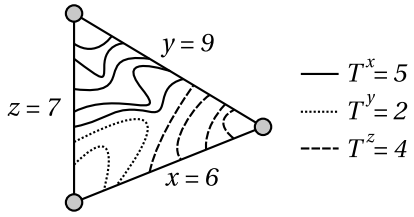


Figure 4. *The path of the lamination inside a triangle T can be completely reconstructed from the set of crossing numbers $\{x_i, y_i, z_i\}$. The solution in a single triangle is illustrated. Equation (1) determines the number of arcs that must pass directly between each pair of edges. Up to homotopy, there is only one way to draw these arcs without crossings.*

the triangle T depicted in Figure 4, where T is U_i or L_i from Figure 3. The crossing numbers with each of the three edges are x, y, and z. In our notation, any part of the curve passing between edges x and y is counted by T^z, and, likewise, T^y counts arcs passing between x and z, and T^x counts arcs between y and z. Since we assume that the curve is pulled tight at all times, there cannot be any arcs that enter and leave by the same edge, so $T^x + T^y + T^z$ counts

all the arcs. We must have $x = T^y + T^z$, $y = T^x + T^z$, and $z = T^x + T^y$, which gives

$$\begin{aligned} T^x &= \tfrac{1}{2}(y + z - x), \\ T^y &= \tfrac{1}{2}(x + z - y), \\ T^z &= \tfrac{1}{2}(x + y - z). \end{aligned} \tag{1}$$

Given T^x, T^y, and T^z, there is only one way to draw the arcs without them crossing each other. The T^x, T^y, and T^z arcs are highlighted using different dash patterns for the example in Figure 4. The numbers given by (1) are crucial to the arguments that follow.

It is worth remarking that although every integral lamination corresponds to a unique set of crossings $\{x_i, y_i, z_i\}$, the converse is not true. Exactly those triplets related to each L_i and U_i that satisfy the triangle inequalities $y + z \geq x$, $x + z \geq y$, and $x + y \geq z$ implied by (1) correspond to a valid closed loop.

3. Deformation of integral laminations under braid operations. In this section we describe the effect of braiding operations on an integral lamination. We start with a representative loop that is pulled tight, and we record the initial set of crossing numbers $\{x_i, y_i, z_i\}$. All we have to do now is to determine how these numbers are updated under the action of each braid operation.

To calculate the new set of crossings we simply determine the number of crossings with the *preimage* of each edge. We denote the preimage of edge e by e^*, which is a curve that becomes e (up to homotopy) after the braid operation. Our argument is that the number of crossings of e^* before the braid operation has to be equal to the number of crossings with e afterwards. This must be the case, as the only way the number of crossings could change is if the end of a loop were to cross through the edge as it is deformed from e^* to e—but by the pulled-tight assumption any such loop would have to be wrapped around a puncture, and by construction no punctures pass through the edge as it deforms from e^* to e.

It will turn out that we have to consider explicitly only the braid group operations ρ_i, ρ_i^{-1} $(i = 1, \ldots, n)$ and σ_i, σ_i^{-1} $(i = 1, \ldots, n-1)$. We then determine the effect of other independent group elements τ_i and τ_i^{-1} by invoking group presentation rules and the operation σ_n (see Figure 1(b)). This is explained in detail in section 3.5 for those not familiar with the braid group presentation. For convenience we assume in what follows that indices are treated "modulo" n, so that the puncture to the right of puncture n is puncture 1.

3.1. Crossing update rules for ρ_i. For the braid operation ρ_i, we need consider only edges incident on the ith puncture, since the number of crossings with other edges will remain unchanged. For example, the relevant preimages are shown for ρ_3 in Figure 5 as dashed lines. The affected crossing numbers here are x_2, y_2, x_3, y_3, and z_3. Many of the preimages are other edges, so the number of crossings is already known. This is not a coincidence: the triangulation was chosen to have this property. In this example, x_2 and y_3 are the exact preimages of y_2 and x_3, respectively. The edge z_3 is the preimage of itself, so in fact its crossing number will not change under ρ_3. The preimages x_2^* and y_3^* are not edges in the triangulation, so more work is required to determine how many crossings are made. We illustrate in Figure 6 how to determine the number of crossings with the x_2^* preimage. All arcs entering x_2 and z_3 must cross x_2^* (by the pulled-tight assumption), unless they loop directly from x_2 to z_3 (shown

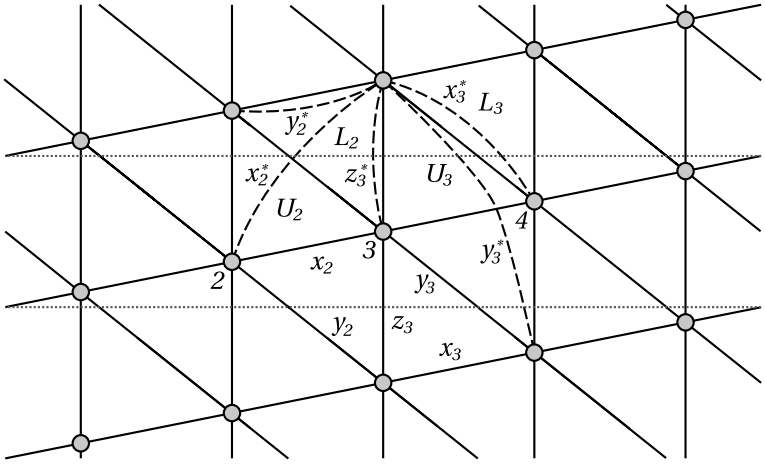


Figure 5. *Preimage illustration for ρ_3.*

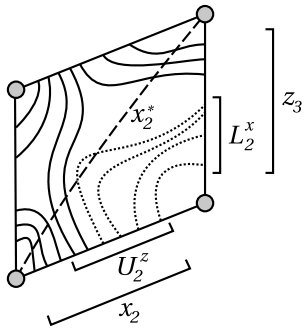


Figure 6. *The number of crossings with a preimage x_2^* that is not part of the triangulation.*

dotted). The number of these loops is exactly $\min(U_2^z, L_2^x)$. Hence the number of preimage crossings is $x_2 + z_3 - 2\min(U_2^z, L_2^x)$. The number of crossings y_3^* can be deduced in a similar way.

A very important feature to note here is that any parts of the curve that cross over the preimage and immediately back again (see Figure 6) are automatically discounted, so that the pulled-tight assumption still holds for the updated set of crossing numbers. This prevents the counting of loops that might later be shortened, and thus the growth rate of the braid from being overestimated.

For a general ρ_i, the update rules for the affected crossing numbers are

$$
\begin{aligned}
x_{i-1}^* &= x_{i-1} + z_i - 2\min(U_{i-1}^z, L_{i-1}^x) = x_{i-1} + z_{i-1} - 2\min(U_{i-1}^x, L_{i-1}^z),\\
y_{i-1}^* &= x_{i-1},\\
x_i^* &= y_i,\\
y_i^* &= z_i + y_i - 2\min(U_i^y, L_i^z) = y_i + z_{i+1} - 2\min(U_i^z, L_i^y),\\
z_i^* &= z_i.
\end{aligned}
\tag{2}
$$

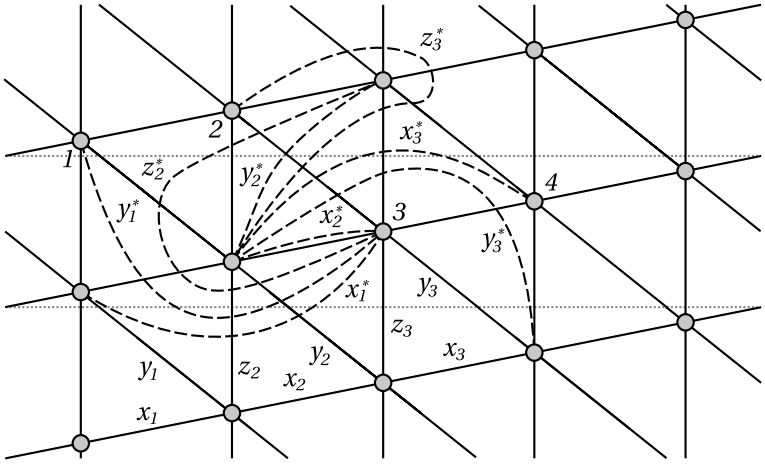


Figure 7. *Preimage illustration for σ_2.*

3.2. Crossing update rules for ρ_i^{-1}. We can work out the update rule for ρ_i^{-1} by noting the π-rotational symmetry of the triangulation about a puncture and relabeling the variables in the rules (2) given above. The update rules for affected crossing numbers are

$$\begin{aligned} x_{i-1}^* &= y_{i-1}, \\ y_{i-1}^* &= y_{i-1} + z_i - 2\min(U_{i-1}^z, L_{i-1}^y) = y_{i-1} + z_{i-1} - 2\min(U_{i-1}^y, L_{i-1}^z), \\ x_i^* &= z_i + x_i - 2\min(U_i^x, L_i^z) = x_i + z_{i+1} - 2\min(U_i^z, L_i^x), \\ y_i^* &= x_i, \\ z_i^* &= z_i. \end{aligned} \tag{3}$$

3.3. Crossing update rules for σ_i. The same ideas can be used to determine the updated crossing numbers following a σ_i operation. However, since there are two moving punctures, more edges are affected, and the preimages are therefore more complicated. An example of a preimage diagram for σ_2 is shown in Figure 7. The dashed lines show the preimages of all edges whose crossing number may change. The edge x_2 is its own preimage, and so the number of crossings remains unchanged. The number of crossings with y_2^* is determined using the quadrilateral trick illustrated in Figure 6. The other preimages require further calculation.

Figure 8 illustrates how to compute the crossings with x_1^* and y_1^*. In a way similar to that in Figure 6, the number of crossings with preimage x_1^* is given by the number of arcs crossing x_1 and x_2, minus twice the number of loops directly between x_1 and x_2. The number of such loops is exactly $\min(L_1^y, U_2^x, L_2^z)$. Hence $x_1^* = x_1 + x_2 - 2\min(L_1^y, U_2^x, L_2^z)$. The preimage problem for y_1^* is similar, but involves four triangles and hence a minimum of four numbers, so that $y_1^* = y_1 + x_2 - 2\min(U_1^z, L_1^y, U_2^x, L_2^z)$. Preimages x_3^* and y_3^* are handled in the same way as x_1^* and y_1^*, by invoking rotational symmetry.

The two remaining crossing numbers to find are z_2^* and z_3^*. Since these preimages pass through a total of seven triangles, it is nontrivial to determine the number of crossings directly. Instead we employ a trick and deduce z_2^* and z_3^* by invoking the quadrilateral solution again with the *updated* crossing numbers. We introduce temporary edges p and q directly between

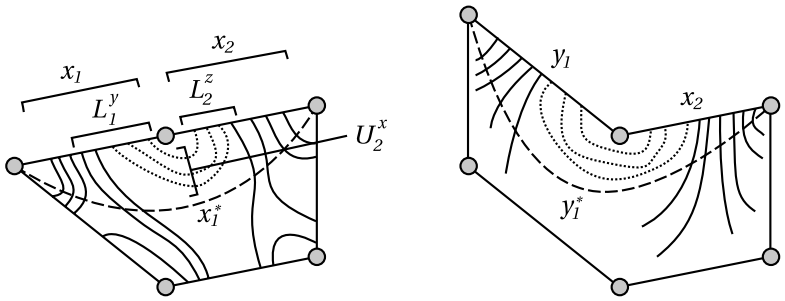


Figure 8. *Two slightly more difficult preimage problems for the operation σ_2.*

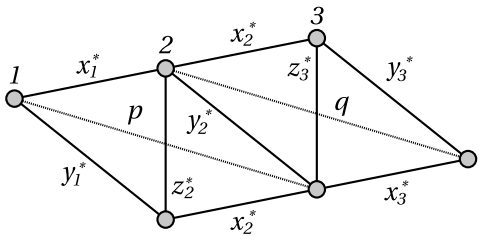


Figure 9. *The number of crossings with z_2^* and z_3^* are difficult to determine directly from the old set of crossing numbers. However, by introducing temporary edges p and q, they may be calculated using the quadrilateral solution shown in Figure 6 from the already updated crossing numbers.*

punctures 1 and 3 and between 2 and 4, as shown in Figure 9. The preimage of p is y_1, and the preimage of q is y_3. Since x_1^*, y_1^*, p, x_2^*, y_2^*, q, x_3^*, and y_3^* are known, z_2^* and z_2^* can be deduced easily.

In general, for σ_i the update rules for the affected crossing numbers are

$$
\begin{aligned}
x_{i-1}^* &= x_{i-1} + x_i - 2\min(L_{i-1}^y, U_i^x, L_i^z),\\
y_{i-1}^* &= y_{i-1} + x_i - 2\min(U_{i-1}^z, L_{i-1}^y, U_i^x, L_i^z),\\
x_i^* &= x_i,\\
y_i^* &= z_i + x_i - 2\min(U_i^x, L_i^z) = x_i + z_{i+1} - 2\min(U_i^z, L_i^x),\\
z_i^* &= x_{i-1}^* + y_{i-1}^* - \min(y_{i-1}^* + y_{i-1} - x_i^*, x_{i-1}^* + y_{i-1} - y_i^*),\\
x_{i+1}^* &= x_i + x_{i+1} - 2\min(U_i^z, L_i^x, U_{i+1}^y),\\
y_{i+1}^* &= x_i + y_{i+1} - 2\min(U_i^z, L_i^x, U_{i+1}^y, L_{i+1}^z),\\
z_{i+1}^* &= x_i^* + y_i^* - \min(x_i^* + y_{i+1} - y_{i+1}^*, y_i^* + y_{i+1} - x_{i+1}^*).
\end{aligned}
\tag{4}
$$

This includes the operation σ_n, which switches the first and last punctures across the periodic boundaries (see Figure 1(b)); we shall require this operation in sections 3.5–3.6 to find the update rules for τ_i and τ_i^{-1}.

3.4. Crossing update rules for σ_i^{-1}. Since the triangulation does not have a reflection symmetry about a vertical line through the midpoint of two punctures, it is not possible to deduce the update rules for σ_i^{-1} by a relabeling in the rules for σ_i. An example of a preimage

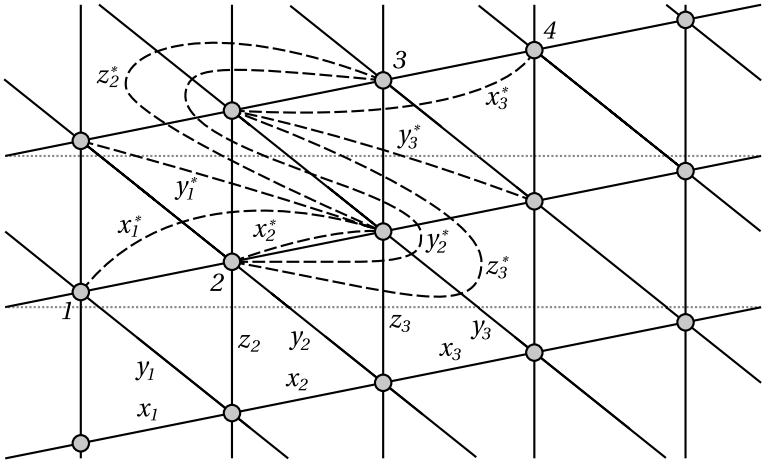


Figure 10. *Preimage illustration for σ_2^{-1}. Note that due to asymmetry the preimage problems are slightly different from those for σ_2.*

diagram for σ_2^{-1} is shown in Figure 10. The preimage curve y_i^* is the most complicated yet, as it passes through ten triangles. However, the crossing number of all these preimages can be calculated as before, using the techniques illustrated in Figures 6, 8, and 9.

The update rules for the affected crossing numbers are

$$
\begin{aligned}
x_{i-1}^* &= x_{i-1} + x_i - 2\min(U_{i-1}^z, L_{i-1}^x, U_i^y),\\
y_{i-1}^* &= y_{i-1} + x_i - 2\min(L_{i-1}^x, U_i^y) = x_{i-1} + y_i - 2\min(L_{i-1}^y, U_i^x),\\
x_i^* &= x_i,\\
y_i^* &= x_i^* + z_i^* - \min(z_i^* + y_i - x_i^*, x_i^* + y_i - z_{i+1}^*),\\
z_i^* &= x_i + y_i - 2\min(U_i^x, L_{i-1}^y, U_{i-1}^z, L_{i-1}^x, U_i^y),\\
x_{i+1}^* &= x_i + x_{i+1} - 2\min(L_i^y, U_{i+1}^x, L_{i+1}^z),\\
y_{i+1}^* &= y_i + x_{i+1} - 2\min(L_i^x, U_{i+1}^y) = x_i + y_{i+1} - 2\min(L_i^y, U_{i+1}^x),\\
z_{i+1}^* &= x_i + y_i - 2\min(L_i^y, U_{i+1}^x, L_{i+1}^z, U_{i+1}^y, L_i^x).
\end{aligned}
\tag{5}
$$

3.5. Crossing update rules for τ_i. In sections 3.1–3.4 we showed how to update the set of crossing numbers $\{x_i, y_i, z_i\}$ for the braid operations ρ_i, ρ_i^{-1}, σ_i, and σ_i^{-1}. To complete the set of update rules for any braid we must give the corresponding rules for τ_i and τ_i^{-1}. In performing τ_i, the ith puncture moves once around the torus in the horizontal direction (see Figure 1). In doing so it passes through many edges in our triangulation, so it is difficult to draw the preimages and to derive the number of crossings directly. However, we can deduce the update rules for τ_i and τ_i^{-1} by appealing to group properties [3] that relate σ_i, ρ_i, and τ_i. In Figure 11 we illustrate how τ_1 is achieved through a sequence of σ_i^{-1} (including σ_n^{-1} defined in Figure 1(b)) followed by one ρ_1^{-1}. The other τ_i and τ_i^{-1} are produced in a similar manner. A computational recipe for each τ_i operation is given in this section, and τ_i^{-1} in the following section.

To calculate the updated set of crossing numbers $\{x_i, y_i, z_i\}$ for τ_i do the following:

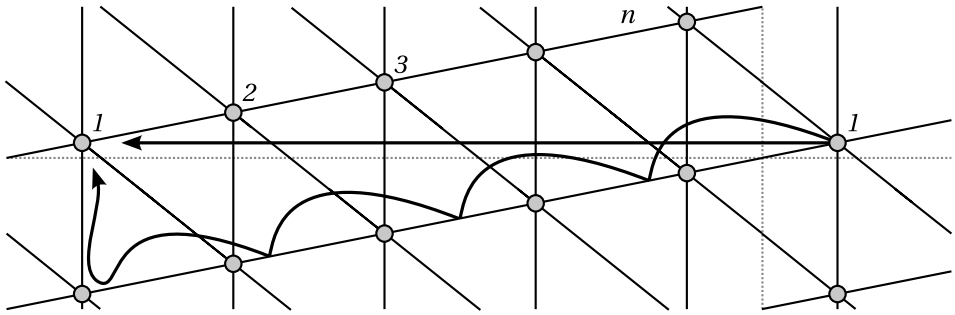


Figure 11. *The operation τ_1 is achieved using a sequence of σ_j^{-1} operations, including the operation σ_n^{-1} (see Figure 1), followed by a single ρ_1^{-1}. This is described in more detail in the text.*

1. Use (5) to perform, in turn, σ_{i-1}^{-1}, σ_{i-2}^{-1}, ..., σ_{i+2}^{-1}, and σ_{i+1}^{-1}. Treat the indices "modulo" n, so that σ_n^{-1} follows σ_1^{-1}.
2. Relabel $x_i \leftarrow x_{i+1}$, $y_i \leftarrow y_{i+1}$, and $z_i \leftarrow z_{i+1}$. This leaves all punctures except the ith one in the correct position.
3. Use (3) to perform ρ_i^{-1}.

3.6. Crossing update rules for τ_i^{-1}. To calculate the updated set of crossing numbers $\{x_i, y_i, z_i\}$ for τ_i^{-1} invert the operation of section 3.5 as follows:

1. Use (2) to perform ρ_i.
2. Relabel $x_i \to x_{i+1}$, $y_i \to y_{i+1}$, and $z_i \to z_{i+1}$. This leaves the punctures in the wrong position, but the next sequence corrects everything ...
3. Use (4) to perform, in turn, σ_{i+1}, σ_{i+2}, ..., σ_{i-2}, and σ_{i-1}. Treat the indices "modulo" n, so that σ_1 follows σ_n.

4. Illustrations using simple braids. To illustrate the use of the update rules given in section 3 we show how a loop is deformed under some simple previously studied braids. We have implemented the update rules in a short C++ program, using the Gnu Multiple Precision library to allow the number of crossings to grow arbitrarily large while maintaining exact arithmetic. We also have a Matlab script to draw the loop, one triangle at a time, using the procedure described in section 2. This was used to produce all the figures in this section. In each case our initial condition is a closed loop that passes between the first two punctures, with $x_1 = y_1 = 1$ and all other x_i, y_i, and z_i set to zero (see the upper-left frame in Figure 12).

Figure 12 shows the roll up of the loop under the repeated action of the planar braid σ_1 with three punctures. In this case it is clear that the third puncture is redundant. This braid has zero entropy and is very poor at stirring as it results in linear growth of material lines. Note that even though coils form around the pair of moving punctures, there is always exactly one crossing of the loop from one copy of the domain to the copy above. This is because there was exactly one crossing in the initial loop and this is only a planar braid, so it cannot create any further crossings under the pulled-tight assumption.

Figure 13 shows the result of repeating τ_2 with three punctures. In this simple braid the second puncture moves in a straight line to the left, but catching the loop on the puncture exactly once. The illustration is provided to validate the method described in sections 3.5–3.6

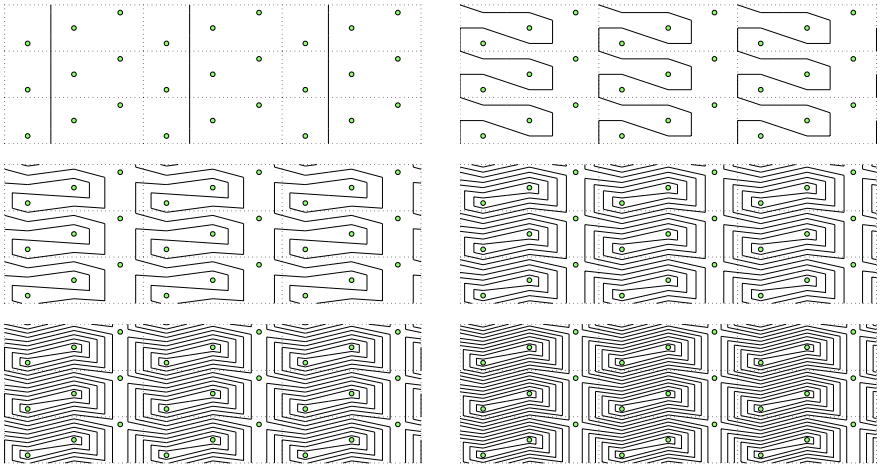


Figure 12. *Lamination evolution under the braid σ_1. This results in linear growth in the number of crossings, and zero entropy.*

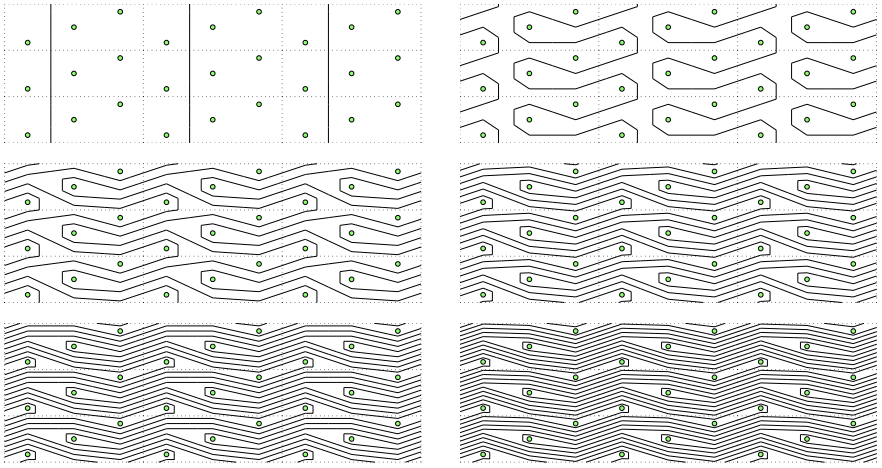


Figure 13. *Lamination evolution under the braid τ_2. This braid also has zero topological entropy, but the illustration is provided to verify the correctness of the procedure for performing τ_i using a combination of σ_i and ρ_i.*

for performing τ_i operations using a combination of σ_i and ρ_i motions.

The planar pigtail braid $\sigma_1\sigma_2^{-1}$ with three punctures is illustrated in Figure 14 [7]. This braid is pseudo-Anosov and has a growth rate per braid letter of $\frac{1}{2}(1+\sqrt{5})$, which is the golden ratio. This "golden braid" has been proved to have the highest topological entropy per braid letter for a planar braid on three strands [8]. It is also conjectured that no braid on more strands has a higher entropy per braid letter than the golden braid [21, 31].

Figure 15 illustrates a related cylinder braid $\sigma_1\sigma_3\sigma_2^{-1}\sigma_4^{-1}$ with four punctures. This is similar to the pigtail braid but wrapped around a cylinder so that the first and last punctures are allowed to exchange places. The cylinder braid $\sigma_1\sigma_2^{-1}$ with *two* punctures has a growth rate per braid letter given by the silver ratio of $1+\sqrt{2}$, and this can be proved to be the optimum

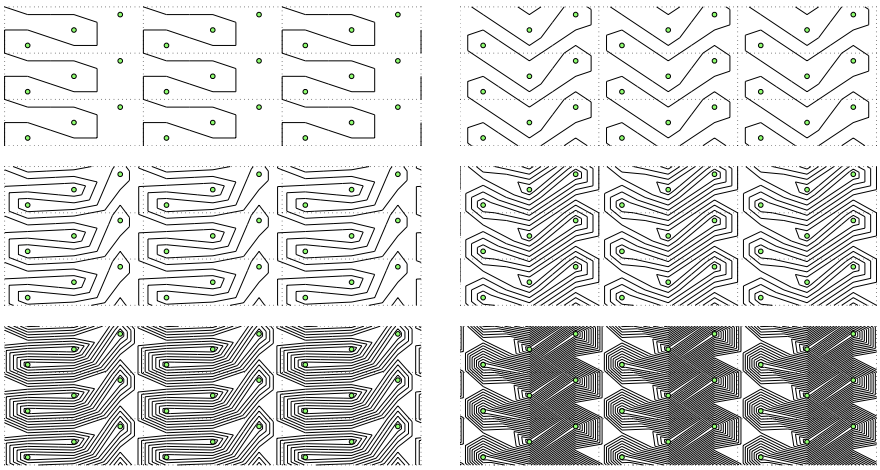


Figure 14. *Lamination evolution under the (planar) golden braid* $\sigma_1\sigma_2^{-1}$.

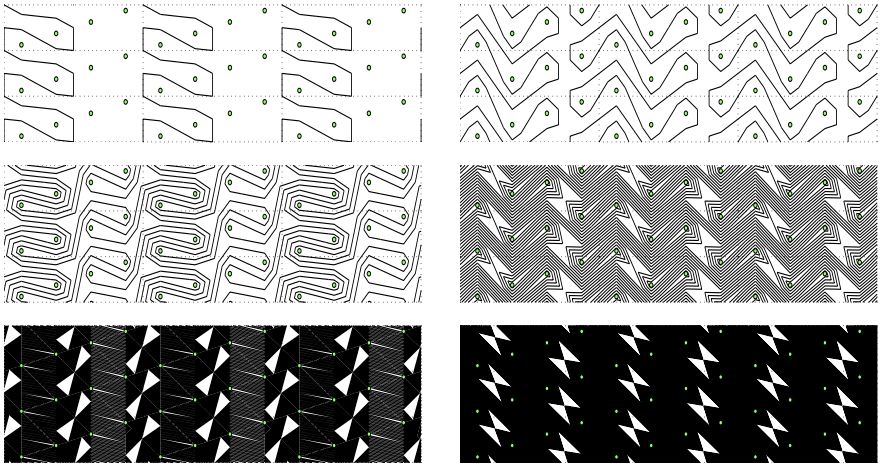


Figure 15. *Lamination evolution under the cylinder braid* $\sigma_1\sigma_3\sigma_2^{-1}\sigma_4^{-1}$. *Note that the stretched curve becomes visibly dense more quickly than for the golden braid in Figure* 14.

per braid letter over all cylinder braids. (The proof follows that of D'Alessandro, Dahleh, and Mezić [8] using a matrix representation for the cylinder braid group.) The silver braid entropy is almost fifty percent higher than the golden braid entropy, showing that periodic boundary conditions can be exploited to enhance chaos. The extra stretching is clearly visible by comparing Figures 14 and 15.

A final validation of our encoding and update rules is given by checking that a loop is unchanged when it is subjected to the identity braid. For the braid group on the torus, one way of writing the identity using σ_i, ρ_i, and τ_i is $\sigma_1^{-2}\rho_1^{-1}\tau_2\rho_1\tau_2^{-1}$, which is taken from the group presentation as written down by Birman [3]. In Figure 16 the evolution of our test loop is shown using this braid, and it is unchanged, as required.

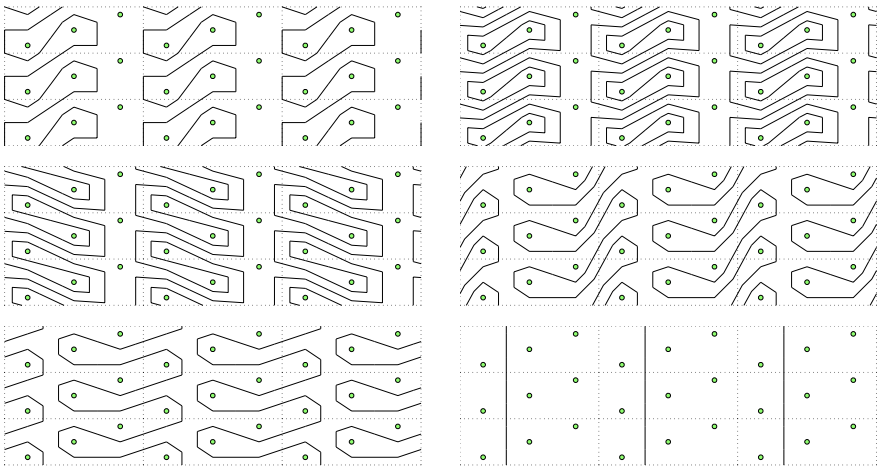


Figure 16. *Lamination evolution under the toroidal identity braid $\sigma_1^{-2}\rho_1^{-1}\tau_2\rho_1\tau_2^{-1}$. The initial loop (not shown) is the vertical line between the first two punctures. The images show, in order, the loop after σ_1^{-1}, σ_1^{-2}, $\sigma_1^{-2}\rho_1^{-1}$, $\sigma_1^{-2}\rho_1^{-1}\tau_2$, $\sigma_1^{-2}\rho_1^{-1}\tau_2\rho_1$, and $\sigma_1^{-2}\rho_1^{-1}\tau_2\rho_1\tau_2^{-1}$. The final loop is the same as the initial condition, as required.*

5. Calculating topological entropies from laminations. Having completed the description of the dynamical system for the deformation of a loop, we now discuss how the topological entropy of a braid is related to evolution of the dynamical system. The results of Moussafir [21] for the punctured sphere apply here: for an appropriate initial loop, the logarithm of the growth rate of the total number of crossings between the loop and triangulation converges to the topological entropy, h, of the braid. If $S(n) = \sum_i x_i(n) + y_i(n) + z_i(n)$ is the total number of crossings after n applications of the braid, then asymptotically $\log S(n)/n \to h$ as $n \to \infty$. An approximation of the entropy is derived by evaluating $\log S(n)/n$ for sufficiently large n. This amounts to estimating the slope of $\log S(n)$ using $S(0)$ and $S(n)$. A more rapidly converging estimate is found by using consecutive iterations of the braid to estimate this slope, as this rapidly becomes independent of transient crossing numbers for small n. The estimate we use is therefore

$$h^\dagger = \log S(n) - \log S(n-1). \tag{6}$$

Convergence is exponential (provided that the braid has a pseudo-Anosov component), and so in practice very few iterations are required to obtain h to double precision. Finite-order braids, which have zero topological entropy, are detected easily in practice by checking for subexponential convergence of $h^\dagger$.

As an example, the exact topological entropy of the four-strand braid $\sigma_1\sigma_3\sigma_2^{-1}\sigma_4^{-1}$ (see the previous section and Figure 15) is known to be twice the logarithm of the silver ratio $h = 2\log(1+\sqrt{2}) = 1.762747174039086\ldots$. In Table 1 we show how the total crossing number and the entropy estimate $h^\dagger$ given by (6) evolve under the first few iterations of the braid (using the initial loop described in section 4). Double precision accuracy of the entropy, which is sufficient for most purposes, is reached after about 20 iterations.

Braids naturally arise when considering periodic trajectories of points in a two-dimensional flow. For a spatially periodic flow the space-time plot of these trajectories is typically a

Table 1

Convergence of the entropy estimate $h^\dagger$ in (6) towards the exact topological entropy $h = 2\log(1+\sqrt{2}) = 1.762747174039086\ldots$ for the silver braid $\sigma_1\sigma_3\sigma_2^{-1}\sigma_4^{-1}$. Convergence is exponential, with approximately one extra digit per iteration [21].

Iteration n	Total crossings $S(n)$	Entropy $h^\dagger$	Error $\lvert h - h^\dagger\rvert$
1	24	2.48490664978800	0.72215947574891
2	154	1.85889877206568	0.09615159802660
3	912	1.77868738766070	0.01594021362162
4	5330	1.76546652708556	0.00271935304647
5	31080	1.76321328732169	0.00046611328261
6	181162	1.76282713309230	0.00007995905321
7	1055904	1.76276089245107	0.00001371841199
8	6154274	1.76274952773491	0.00000235369583
9	35869752	1.76274757786911	0.00000040383002
10	209064250	1.76274724332535	0.00000006928627
11	1218515760	1.76274718592673	0.00000001188765
12	7102030322	1.76274717607868	0.00000000203960
13	41393666184	1.76274717438903	0.00000000034994
14	241259966794	1.76274717409913	0.00000000006004
15	1406166134592	1.76274717404939	0.00000000001030
16	8195736840770	1.76274717404085	0.00000000000177
17	47768254910040	1.76274717403939	0.00000000000030
18	278413792619482	1.76274717403914	0.00000000000005
19	1622714500806864	1.76274717403910	0.00000000000001
20	9457873212221714	1.76274717403908	0.00000000000000

cylinder or a torus braid, and the entropy of this braid provides a rigorous lower bound on the topological entropy of the flow [6, 7]. Finn, Thiffeault, and Gouillart [13] derived such a lower bound for the sine flow using Thurston's "iterate and guess" method to construct the train-track for the braid formed by a set of periodic orbits. We will now describe how the same result can be found using our lamination approach.

The sine flow is a time-periodic alternating shear-flow defined on the unit torus $0 \le x, y < 1$. The period is T, with the velocity field given by $(\sin 2\pi y, 0)$ for $0 \le t < T/2$ and $(0, \sin 2\pi x)$ for $T/2 \le t < T$, with t marking time. This simple flow has been well studied because the parameter range $0 \le T \lesssim 2$ gives rich dynamics that vary from complete integrability to almost global chaos (with few visible islands in a Poincaré section).

Since the flow is piecewise steady, it is easy to construct a map to track the motion of points from one period to the next. Consequently, the entropy of the flow for a given T can be found quickly and simply by direct numerical simulation of line stretching. Once the entropy of the flow is known, it is instructive to see what prediction of the entropy is given by considering the braiding of a finite number of arbitrary particle orbits.

For general T it is difficult to locate unstable periodic orbits due to the highly chaotic nature of the sine mapping; however, for the special parameter $T = 1$ it is quite easy to spot some of the low order orbits. In particular, four period-two orbits are given by $\{(0, \frac{1}{4}), (\frac{1}{2}, \frac{1}{4})\}$, $\{(\frac{1}{2}, \frac{3}{4}), (1, \frac{3}{4})\}$, $\{(\frac{1}{4}, 0), (\frac{1}{4}, \frac{1}{2})\}$, and $\{(\frac{3}{4}, \frac{1}{2}), (\frac{3}{4}, 1)\}$. The first pair of orbits is depicted in Figure 17(a), the second pair in Figure 17(b). The points associated with the first pair of periodic orbits do not move in the second half-period, while those associated with the second

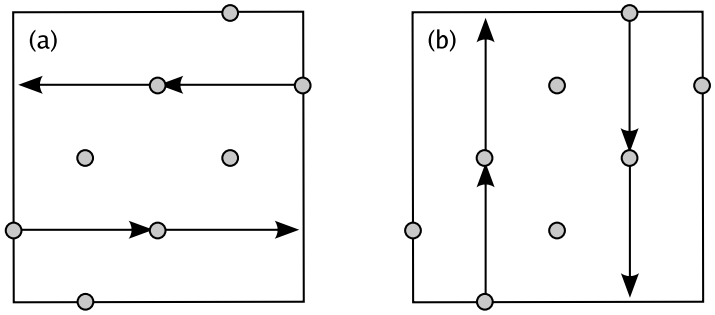


Figure 17. *The four periodic orbits considered for the sine flow with* $T = 1$. (a) $\{(0, \frac{1}{4}), (\frac{1}{2}, \frac{1}{4})\}$ *and* $\{(\frac{1}{2}, \frac{3}{4}), (1, \frac{3}{4})\}$; (b) $\{(\frac{1}{4}, 0), (\frac{1}{4}, \frac{1}{2})\}$ *and* $\{(\frac{3}{4}, \frac{1}{2}), (\frac{3}{4}, 1)\}$.

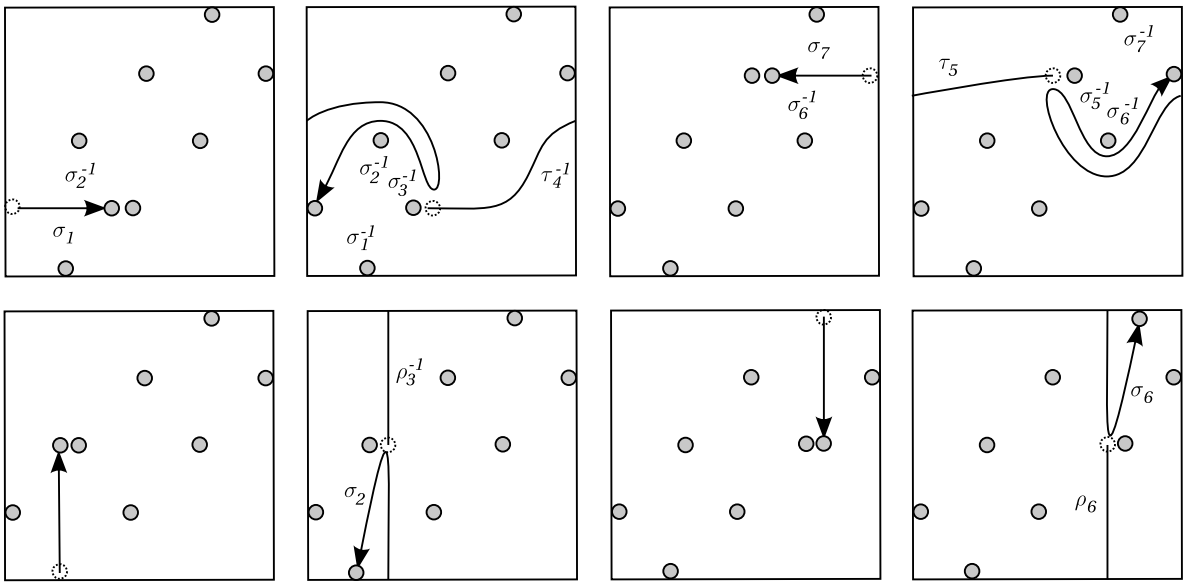


Figure 18. *Illustration of the eight strand braid formed by a set of period-two points in the sine flow with* $T = 1$. *The braid word is* $\sigma_1\sigma_2^{-1}\tau_4^{-1}\sigma_3^{-1}\sigma_2^{-1}\sigma_1^{-1}\sigma_7\sigma_6^{-1}\tau_5\sigma_5^{-1}\sigma_6^{-1}\sigma_7^{-1}\rho_3^{-1}\sigma_2\rho_6\sigma_6$. *This braid has been shown to have an exact entropy of* 1.21875572687... *using train-tracks* [13]. *The entropy estimate using the growth of loops agrees to every decimal place calculated.*

pair do not move in the first half-period.

Now we must determine the braid formed by the eight trajectories forming these four periodic orbits. We first disambiguate the order of the periodic points by displacing them slightly along the x axis, as shown in Figure 18. Then we encode the trajectories in terms of braid group generators, deforming as needed. Deforming is necessary since usually the trajectory does not map directly onto a generator, and some intermediate operations must be inserted. For instance, in the second snapshot in Figure 18 the generators $\sigma_3^{-1}\sigma_2^{-1}\sigma_1^{-1}$ are used to return the point to the leftmost position after a τ_4^{-1} operation. After a full period, the resulting braid word is $\sigma_1\sigma_2^{-1}\tau_4^{-1}\sigma_3^{-1}\sigma_2^{-1}\sigma_1^{-1}\sigma_7\sigma_6^{-1}\tau_5\sigma_5^{-1}\sigma_6^{-1}\sigma_7^{-1}\rho_3^{-1}\sigma_2\rho_6\sigma_6$. Using (6), we find that the entropy of the braid converges to 1.21875572687..., which agrees with an alternative

calculation using train-tracks [13]. The entropy lower bound given by the braid accounts for 82% of the flow entropy of approximately 1.48.

Since it is difficult to find any exact periodic orbits for general T it is natural to ask whether the entropy can be found by considering the braiding of any selection of trajectories. Because point motions in the sine flow are piecewise horizontal and vertical, and, for the purposes of braiding, can be performed sequentially, it is straightforward and computationally fast to calculate all the σ_i, ρ_i, and τ_i operations that occur during each half-period of the flow, for an arbitrary number of points. For the first half-period, where all motions are horizontal, we record a σ_i operation for each change in order of the particle x coordinates. The sign of each crossing is determined by the difference in y coordinates for the two points that cross. Special attention is required when a point crosses the periodic boundary. Crossing leftwards over $x = 0$ is achieved by $\tau_1\sigma_1^{-1}\cdots\sigma_{n-1}^{-1}$; this moves the leftmost point one copy of the domain to the left, and then, through a sequence of σ operations, moves the point to position n and undoes all the undesired crossings with the other points. Similarly, when the rightmost point crosses over $x = 1$ the string $\tau_n^{-1}\sigma_{n-1}^{-1}\cdots\sigma_1^{-1}$ is assigned. The computation is simpler during the second half-period: all point motions are vertical, so no σ_i motions occur. If the ith point in the x direction crosses downwards through $y = 0$, this is labeled ρ_i. Likewise, an upwards crossing of $y = 1$ is assigned ρ_i^{-1}.

To estimate the entropy of the flow we evolve an integral lamination according to the braid that results from the point motions. Since the points are not periodic, the growth per iteration continues to vary, but the average value of the growth always converges. The initial points are chosen at random throughout the domain, and can either live in a chaotic region or in a periodic island. Hence the predicted entropy for a given number of points may depend on the initial positions. To allow for this we use an estimate $\langle h^\dagger \rangle$ obtained by averaging over many realizations with different initial points.

In Figure 19 we show how the braid entropy $\langle h^\dagger \rangle$ converges towards the flow entropy h as the number of trajectories in the braid is increased. For the four different values of T we have considered, convergence appears to be as a power law for large n. Although we are unable to prove that the exact entropy is reached in the limit as $n \to \infty$, we would expect this since in this limit the braid will contain perfect information about the flow. For small values of T (such as $T = 0.5$ in Figure 19) the flow contains large islands of regularity, which do not contribute greatly to line stretching. In this regime n has to become relatively large before there are enough points exploring and encoding the dynamics in the small chaotic region to get a good entropy estimate.

Owing to the efficiency with which the braid is determined and analyzed, we are able to consider much larger times and number of points n than have been considered in previous articles [30]. Another point worth noting is that for $T = 1$, the mean entropy estimate $\langle h^\dagger \rangle = 0.45$ with $n = 8$ points is much worse than the estimate of $h^\dagger = 1.22$ found by using the set of eight judiciously chosen periodic points considered earlier. This highlights the important role of low-order periodic orbits in determining much of the nature of the flow.

We offer the numerical evidence of this section without proof. To make the estimate based on randomly selected trajectories rigorous would require a more precise characterization of the selection. For instance, what measure is used in the selection of the trajectories? (We effectively used Lebesgue.) The compact region that supports the maximum topological

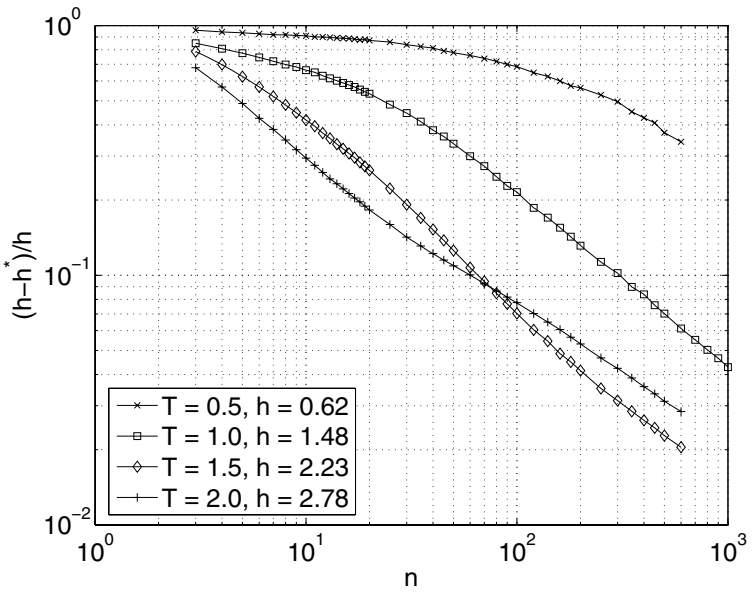


Figure 19. *Convergence of the mean entropy estimate $\langle h^\dagger \rangle$ towards the exact entropy h as the number of trajectories n in the braid is increased. For large n the error appears to decrease according to a power law. For smaller T the convergence is hindered by many of the points being inside periodic islands.*

entropy could be singular with respect to the measure, but for practical applications (such as fluid dynamics) this would render that region irrelevant. Franks and Handel [14] have shown that for sufficiently smooth flows there exists a set of points that give an estimate of the entropy, but in practice this set is not easy to find.

6. Discussion.

6.1. Relation to previous work on surface dynamics. Before closing with a final discussion of our results, we wish to describe our work in the wider context of the topological theory of surface dynamics. This theory seeks a description of equivalence classes, under isotopy, of homeomorphisms of compact surfaces [6, 10]. The key classification theorem is due to Thurston [10, 32], based on earlier work of Nielsen [23, 24, 25], and the collection of results is often called Thurston–Nielsen theory. The classification theorem basically says that all homeomorphisms are isotopic to either a finite-order or pseudo-Anosov representative, or are reducible.

One of Thurston's greatest contributions was to *compactify* Teichmüller space, which is, roughly speaking, the space of equivalence classes of hyperbolic metrics on the surface. He did this by attaching a natural boundary to Teichmüller space, consisting of projective measured foliations [10]. This compactified Teichmüller space is homeomorphic to a finite-dimensional disk, which means that by Brouwer's theorem any continuous mapping from the disk to itself must have at least one fixed point. This holds, in particular, for the induced action of a homeomorphism on the compactified Teichmüller space. For a mapping isotopic to a pseudo-Anosov homeomorphism, there is a fixed point on the boundary to which convergence under repeated application is exponential. This fixed point is a projective measured foliation; the

integral laminations in Figures 14 and 15 are converging to this measured foliation.

Because for a pseudo-Anosov diffeomorphism the length of the integral lamination grows exponentially under iteration, it is necessary to encode it in an efficient manner. There are two dominant methods for this: triangulation (our choice here) and the Dehn–Thurston parametrization [10]. The Dehn–Thurston parametrization involves a pair-of-pants decomposition of a surface. The number of crossings and the "twist" of an integral lamination between each pair of pants is recorded. Penner [26] and Penner and Harer [27] gave the action of Dehn twists on this parametrization for the case of genus g and no boundary (i.e., no punctures). Schaefer, Štefankovič, and Sedgwick [29] showed that the action of Dehn twists on an integral lamination could be computed in polynomial time both for a triangulation or a Dehn–Thurston parametrization. Hamidi-Tehrani and Chen [18] used a similar parametrization for train-tracks.

The pair-of-pants decomposition is more economical, in that it requires fewer coordinates than a triangulation. However, this is not a serious computational limitation, and we found that the action of the braid group (that is, the mapping class group generators) on the punctured torus is more readily expressed in terms of a triangulation. Our main contribution here is to explicitly give this action on crossing numbers, and to use it to calculate topological entropies.

6.2. Summary and discussion of current work. We have derived a dynamical system to compute the evolution of an integral lamination (equivalence class of simple closed curves) under the braiding of an arbitrary number of punctures on the torus. The method is essentially a modification of the Dynnikov coordinate approach employed by Moussafir [21], but we use a triangulation encoding that has favorable properties for studying torus braids. Naturally, our method also works for the special cases of cylinder and planar braids. However, in the planar case our dynamical system still has $3n$ variables, more than the $2n-3$ required by Moussafir, so there is clearly some redundancy in this case. Also, the triangulation of our domain is not unique, so we expect that the details of our method are not unique. However, our triangulation seems to be the best choice for simplifying the arithmetic.

For completeness we point out that our dynamical system does not work directly for $n=2$ punctures because in (4) and (5) this would mean that the indices $i-1$ and $i+1$ refer to the same quantities. There is nothing difficult about the case $n=2$ though, and in principle one could write down the corrected update rules for $\{x_1, y_1, z_1, x_2, y_2, z_2\}$. Alternatively, a lazy but convenient workaround is to include a redundant third puncture, glued to one of the other two, so that it is slaved to its motion. At a little more computational expense this allows one computer code to handle all values of $n \geq 2$. The case $n=1$ is trivial.

By using arguments similar to those of Moussafir [21], it can be shown that as the number of iterations tends to infinity, the logarithm of the total number of crossings grows by the braid topological entropy. Though the number of crossings grows exponentially fast, our numerical implementation of the dynamical system uses a large integer arithmetic library to allow calculation of the entropy to arbitrary precision. As discussed in section 6.1, for pseudo-Anosov braids, convergence is exponentially fast, with approximately one digit of accuracy gained per iteration for the braids we have considered.

In practice, if only a few digits of accuracy are required, then double precision floating-point arithmetic for the crossing numbers is adequate and can speed up code significantly.

A caveat with using floating-point arithmetic is that it destroys reversibility. In general, by performing a long braid followed by its inverse, the initial condition will not be recovered due to exponential growth of small roundoff errors. This is akin to irreversibility due to numerical diffusion in trajectory computations in a chaotic Stokes flow. However, since the dynamical system itself is exact, the only errors are due to roundoff and not to discretization, so even with double precision the dynamical system is surprisingly reversible.

If finite precision is not acceptable, our method is easily adapted to find exact entropies. This can be done by shortcircuiting the minimum functions in the update rules (2)–(5). Under repeated iteration of a braid, we observe numerically that the dynamical system quickly becomes linear. That is, each "min" function always selects the same argument. With this knowledge, each min function can be replaced by a single argument, and a linear system written down for the crossing numbers. The logarithm of the modulus of the largest eigenvalue gives the exact entropy.

This work was motivated by the study of two-dimensional fluid mixing via the braiding motion of fluid particle trajectories [15, 30]. In this setting we have derived an efficient tool that allows practical analysis of large braids. Outside of this particular application it is natural to ask whether the method can be generalized for the braid group on surfaces of higher genus, as considered by Birman [3]. Since any surface can be triangulated, in principle our method could be extended to higher genus. The problem anticipated with generalizing the approach is that it is more difficult to visualize the preimage problems on a surface with many holes, and also it is not clear in general how to exploit group properties to gain a complete set of crossing update rules.

Acknowledgments. We thank Colin Cotter and Jacques-Oliver Moussafir for helpful discussions, as well as Toby Hall, Phil Boyland, and other participants of the *Computation and Topology in Dynamics* workshop in Leiden.

REFERENCES

[1] H. Aref, *The development of chaotic advection*, Phys. Fluids, 14 (2002), pp. 1315–1325.
[2] M. Bestvina and M. Handel, *Train-tracks for surface homeomorphisms*, Topology, 34 (1995), pp. 109–140.
[3] J. S. Birman, *On braid groups*, Comm. Pure Appl. Math., 22 (1969), pp. 41–72.
[4] J. S. Birman, *Braids, Links, and Mapping Class Groups*, Ann. of Math. Stud., Princeton University Press, Princeton, NJ, 1975.
[5] R. Bowen, *Entropy and the fundamental group*, in Structure of Attractors, Lecture Notes in Math. 668, Springer, New York, 1978, pp. 21–29.
[6] P. L. Boyland, *Topological methods in surface dynamics*, Topology Appl., 58 (1994), pp. 223–298.
[7] P. L. Boyland, H. Aref, and M. A. Stremler, *Topological fluid mechanics of stirring*, J. Fluid Mech., 403 (2000), pp. 277–304.
[8] D. D'Alessandro, M. Dahleh, and I. Mezić, *Control of mixing in fluid flow: A maximum entropy approach*, IEEE Trans. Automat. Control, 44 (1999), pp. 1852–1863.
[9] I. A. Dynnikov, *On a Yang–Baxter map and the Dehornoy ordering*, Russian Math. Surveys, 57 (2002), pp. 592–594.
[10] A. Fathi, F. Laundenbach, and V. Poenaru, *Travaux de Thurston sur les surfaces*, Astérisque, 66–67 (1979), pp. 1–284.
[11] M. D. Finn, S. M. Cox, and H. M. Byrne, *Topological chaos in inviscid and viscous mixers*, J. Fluid Mech., 493 (2003), pp. 345–361.

[12] M. D. Finn, S. M. Cox, and H. M. Byrne, *Mixing measures for a two-dimensional chaotic Stokes flow*, J. Engrg. Math., 48 (2004), pp. 129–155.
[13] M. D. Finn, J.-L. Thiffeault, and E. Gouillart, *Topological chaos in spatially periodic mixers*, Phys. D, 221 (2006), pp. 92–100.
[14] J. M. Franks and M. Handel, *Entropy and exponential growth of π_1 in dimension two*, Proc. Amer. Math. Soc., 102 (1988), pp. 753–760.
[15] J.-M. Gambaudo and E. E. Pécou, *Dynamical cocycles with values in the Artin braid group*, Ergodic Theory Dynam. Systems, 19 (1999), pp. 627–641.
[16] E. Gouillart, J.-L. Thiffeault, and M. D. Finn, *Topological mixing with ghost rods*, Phys. Rev. E (3), 73 (2006), paper 036311.
[17] T. Hall, *Train: A C++ Program for Computing Train Tracks of Surface Homeomorphisms*, online at http://www.liv.ac.uk/maths/PURE/MIN_SET/CONTENT/members/T_Hall.html.
[18] H. Hamidi-Tehrani and Z.-H. Chen, *Surface diffeomorphisms via train-tracks*, Topology Appl., 73 (1996), pp. 141–167.
[19] B. Kolev, *Entropie topologique et représentation de Burau*, C. R. Acad. Sci. Paris Sér. I Math., 309 (1989), pp. 835–838; English translation available at http://arxiv.org/abs/math.DS/0304105.
[20] M. Lefranc, *Alternative determinism principle for topological analysis of chaos*, Phys. Rev. E, 64 (2006), paper 035202.
[21] J.-O. Moussafir, *On the Entropy of Braids*, manucript, 2006; available online at http://arxiv.org/abs/math.DS/0603355.
[22] S. Newhouse and T. Pignataro, *On the estimation of topological entropy*, J. Statist. Phys., 72 (1993), pp. 1331–1351.
[23] J. Nielsen, *Untersuchungen zur Topologie des geschlossenen zweiseitigen Flächen.* I, Acta Math., 50 (1927), pp. 189–358.
[24] J. Nielsen, *Untersuchungen zur Topologie des geschlossenen zweiseitigen Flächen.* II, Acta Math., 53 (1929), pp. 1–76.
[25] J. Nielsen, *Untersuchungen zur Topologie des geschlossenen zweiseitigen Flächen.* III, Acta Math., 58 (1932), pp. 87–167.
[26] R. C. Penner, *The action of the mapping class group on curves in surfaces*, Enseign. Math. (2), 30 (1984), pp. 39–55.
[27] R. C. Penner and J. L. Harer, *Combinatorics of Train Tracks*, Ann. of Math. Stud. 125, Princeton University Press, Princeton, NJ, 1991.
[28] R. T. Pierrehumbert, *Tracer microstructure in the large-eddy dominated regime*, Chaos Solitons Fractals, 4 (1994), pp. 1091–1110.
[29] M. Schaefer, D. Štefankovič, and E. Sedgwick, *Computing Dehn Twists and Geometric Intersection Numbers in Polynomial Time*, Technical Report TR 05-009, DePaul University, Chicago, IL, 2005.
[30] J.-L. Thiffeault, *Measuring topological chaos*, Phys. Rev. Lett., 94 (2005), paper 084502.
[31] J.-L. Thiffeault and M. D. Finn, *Topology, braids, and mixing in fluids*, Philos. Trans. Roy. Soc. London Ser. A, 364 (2006), pp. 3251–3266.
[32] W. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. (N.S.), 19 (1988), pp. 417–431.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 1, pp. 99–141

Stability Analysis of π-Kinks in a 0-π Josephson Junction*

G. Derks†, A. Doelman‡, S. A. van Gils§, and H. Susanto¶

Abstract. We consider a spatially nonautonomous discrete sine-Gordon equation with constant forcing and its continuum limit(s) to model a 0-π Josephson junction with an applied bias current. The continuum limits correspond to the strong coupling limit of the discrete system. The nonautonomous character is due to the presence of a discontinuity point, namely, a jump of π in the sine-Gordon phase. The continuum model admits static solitary waves which are called π-kinks and are attached to the discontinuity point. For small forcing, there are three types of π-kinks. We show that one of the kinks is stable and the others are unstable. There is a critical value of the forcing beyond which all static π-kinks fail to exist. Up to this value, the (in)stability of the π-kinks can be established analytically in the strong coupling limits. Applying a forcing above the critical value causes the nucleation of 2π-kinks and -antikinks. Besides a π-kink, the unforced system also admits a static 3π-kink. This state is unstable in the continuum models. By combining analytical and numerical methods in the discrete model, it is shown that the stable π-kink remains stable and that the unstable π-kinks cannot be stabilized by decreasing the coupling. The 3π-kink does become stable in the discrete model when the coupling is sufficiently weak.

Key words. 0-π Josephson junction, 0-π sine-Gordon equation, semifluxon, π-kink

AMS subject classifications. 34D35, 35Q53, 37K50, 39A11

DOI. 10.1137/060657984

1. Introduction.

One important application of the sine-Gordon equation is to describe the propagation of magnetic flux (fluxons) in long Josephson junctions [17, 18, 5]. The flux quanta or fluxons are described by the kinks of the sine-Gordon equation. When many small Josephson junctions are connected through the inductance of the superconductors, they form a discrete Josephson transmission line. The propagation of a fluxon is then described by the discrete sine-Gordon equation. For some materials, Josephson junctions are more easily fabricated in the form of a lattice than as a long continuous Josephson junction. In the strong coupling limit, a discrete Josephson junction lattice becomes a long Josephson junction.

It was proposed in the late 1970's by Bulaevskii that a phase-shift of π may occur in the sine-Gordon equation due to magnetic impurities [7, 8]. Only recently has this predic-

*Received by the editors April 22, 2006; accepted for publication (in revised form) by B. Sandstede September 20, 2006; published electronically March 15, 2007. This work was supported by the Royal Netherlands Academy of Art and Sciences (KNAW) and Netherlands Organization for Scientific Research (NWO).
http://www.siam.org/journals/siads/6-1/65798.html

†Department of Mathematics, University of Surrey, Guildford, Surrey, GU2 7XH, UK (g.derks@surrey.ac.uk).

‡Department of "Modelling, Analysis and Simulation," Center for Mathematics and Computer Science (CWI), Kruislaan 413, 1098 SJ Amsterdam, and Korteweg-de Vries Institute, Faculty of Sciences, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands (a.doelman@cwi.nl).

§Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500AE Enschede, The Netherlands (s.a.vangils@math.utwente.nl).

¶Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003-4515 (susanto@math.umass.edu).

tion been confirmed experimentally [35]. Present technological advances can also impose a π-phase-shift in a long Josephson junction using, e.g., superconductors with unconventional pairing symmetry [34], superconductor-ferromagnet-superconductor (SFS) π-junctions [29], or superconductor-normal metal-superconductor (SNS) junctions in which the charge-carrier population in the conduction channels is controlled [4].

A junction containing a region with a phase jump of π is then called a 0-π Josephson junction and is described by a 0-π sine-Gordon equation. The place where the 0-junction meets the π-junction is called a discontinuity point. A 0-π Josephson junction admits a half magnetic flux (semifluxon), sometimes called a π-*fluxon*, attached to the discontinuity point [13]. A semifluxon is represented by a π-kink in the 0-π sine-Gordon equation [32].

Using the technology described in [13], a 0-π array of Josephson junctions can be created as well. Such a system can be modeled by a discrete 0-π sine-Gordon equation. A short numerical study of a discrete π-kink is given in [31].

The presence of the semifluxon in a 0-π Josephson junction or a 0-π array of Josephson junctions opens a new field where many questions that have been discussed in detail for the 2π-kink (fluxon) in the sine-Gordon equation can be addressed for the π-kink, too. The fact that the π-kink cannot move in space, even in the continuum case, will give a different qualitative behavior such as the disappearance of the zero eigenvalue (Goldstone mode), as will be shown later.

In this paper we will study both the continuous and discrete 0-π sine-Gordon equations, especially the stability of the kinks admitted by the equations. Knowing the eigenvalues of a kink is of interest for experimentalists, since the corresponding eigenfunctions (localized modes) can play an important role in the behavior of the kink [27].

The present work is organized as follows: in section 2 we will describe the mathematical model of the problem and its interpretation as a Josephson junction system. We will discuss the discrete system as well as several continuum approximations. In section 3 we consider the continuous 0-π sine-Gordon equation which describes a continuous long Josephson junction with discontinuity point. It is also the lowest order continuum approximation for the discrete system, not reflecting any lattice spacing (coupling) effects. In [32] it is shown that there exist three types of π-kinks in the 0-π sine-Gordon equation. We will analyze their stability and show that one type is stable and the other two are unstable. A higher order continuum approximation, which includes terms representing a small lattice spacing (strong coupling), is considered in section 4. It is shown that for small values of the lattice spacing parameter, the three types of π-kinks persist and their stability properties do not change. In section 5 the discrete 0-π sine-Gordon with large lattice spacing (small coupling) is analyzed, especially the existence and stability of π-kinks. Numerical calculations connecting the regions of small and large lattice spacing (weak and strong coupling) will be presented in section 6. In this section the analytical results of the previous sections are linked together. Conclusions and plans for future research are presented in section 7.

2. Mathematical models for 0-π junctions.

2.1. The discrete 0-π sine-Gordon equation. The Lagrangian describing the phase of a 0-π array of Josephson junctions is given by

$$L = \int \sum_{n\in\mathbb{Z}} \left[\frac{1}{2}\left(\frac{d\phi_n}{dt}\right)^2 - \frac{1}{2}\left(\frac{\phi_{n+1}-\phi_n}{a}\right)^2 - 1 + \cos(\phi_n + \theta_n) + \gamma\phi_n \right] dt, \tag{2.1}$$

where ϕ_n is the Josephson phase of the nth junction. The phase jump of π in the Josephson phase is described by θ_n, where

$$\theta_n = \begin{cases} 0, & n \le 0, \\ -\pi, & 0 < n. \end{cases} \tag{2.2}$$

The Lagrangian (2.1) is given in nondimensionalized form. The lattice spacing parameter a is normalized to the Josephson length λ_J, the time t is normalized to the inverse plasma frequency ω_0^{-1}, and the applied bias current density $\gamma > 0$ is scaled to the critical current density J_c.

The equation of the phase motion generated by the Lagrangian (2.1) is the discrete 0-π sine-Gordon equation

$$\ddot{\phi}_n - \frac{\phi_{n-1} - 2\phi_n + \phi_{n+1}}{a^2} = -\sin(\phi_n + \theta_n) + \gamma. \tag{2.3}$$

We use $n \in \mathbb{Z}$ for the analytical calculations, but, of course, the fabrication of the junction as well as the numerics are limited to a finite number of sites, say, $2N$. We will take the boundary conditions to represent the way in which the applied magnetic field $h = H/(\lambda_J J_c)$ enters the system, i.e.,

$$\frac{\phi_{-N+1} - \phi_{-N}}{a} = \frac{\phi_N - \phi_{N-1}}{a} = h. \tag{2.4}$$

In what follows we will always consider the case when there is no applied magnetic field; i.e., we will take $h = 0$.

2.2. Approximations to the lattice spacing in the continuum limit. There are various continuum model approximations for (2.3) that can be derived in the continuum limit $a \ll 1$. Writing $\phi_n = \phi(na)$ and expanding the difference terms using a Taylor expansion give

$$\frac{\phi_{n-1} - 2\phi_n + \phi_{n+1}}{a^2} = 2\sum_{k=0}^{\infty} \frac{a^{2k}}{(2k+2)!}\, \partial_{xx}^k \phi_{xx}(na) = L_a \phi_{xx}$$

and

$$\frac{\phi_{n+1} - \phi_n}{a} = \sum_{k=0}^{\infty} \frac{a^k}{(k+1)!}\, \partial_x^k \phi(na) = \widetilde{L}_a \phi_x.$$

Thus the continuum approximation for (2.3) is

$$\phi_{tt} - L_a \phi_{xx} = -\sin(\phi + \theta) + \gamma, \tag{2.5}$$

where $\theta(x)$ is defined similarly to (2.2), i.e.,

$$\theta(x) = \begin{cases} 0, & x < 0, \\ -\pi, & x > 0. \end{cases}$$

The continuum approximation for the Lagrangian is

$$L = \iint_{-\infty}^{\infty} \left[\frac{1}{2} (\phi_t)^2 - \frac{1}{2} \left(\widetilde{L}_a \phi_x \right)^2 - 1 + \cos(\phi + \theta) + \gamma\phi \right] dx\, dt.$$

Note that the normalizations in the discrete system imply that the spatial coordinate x is normalized to the Josephson length λ_J.

There are several ways to derive approximations for the operator L_a when $a \to 0$; see, for example, [28]. The first obvious approximation is

$$\phi_{tt} - \phi_{xx} - \frac{a^2}{12}\phi_{xxxx} = -\sin(\phi + \theta) + \gamma, \quad x \neq 0. \tag{2.6}$$

Another approximation can be found by using that $(1 - \frac{a^2}{12}\partial_{xx})\, L_a = 1 - \frac{a^4}{240}\partial_{xx}^2 + \cdots$. This result reflects the invertibility of L_a up to fourth order. Hence $(1 - \frac{a^2}{12}\partial_{xx})$ acting on (2.5) gives the approximation (up to fourth order terms)

$$\phi_{xx} = \phi_{tt} + \sin(\phi + \theta) - \gamma - \frac{a^2}{12}\partial_{xx}(\phi_{tt} + \sin(\phi + \theta)), \quad x \neq 0. \tag{2.7}$$

Expanding this equation and using the expression for ϕ_{xx} again, we get

$$\begin{aligned} \phi_{xx} = \phi_{tt} &+ \sin(\phi + \theta) - \gamma \\ &- \frac{a^2}{12} \left(\phi_{tttt} + [\sin(\phi + \theta)]_{tt} - \phi_x^2 \sin(\phi + \theta) \right. \\ &\left. + \cos(\phi + \theta)[\phi_{tt} + \sin(\phi + \theta) - \gamma] \right), \quad x \neq 0. \end{aligned} \tag{2.8}$$

The steady state equation for (2.6) is

$$\phi_{xx} + \frac{a^2}{12}\phi_{xxxx} = \sin(\phi + \theta) - \gamma, \quad x \neq 0,$$

while (2.7) yields the equation

$$\phi_{xx} = \left(1 - \frac{a^2}{12}\partial_{xx} \right) \sin(\phi + \theta) - \gamma, \quad x \neq 0,$$

and (2.8) gives

$$\phi_{xx} = \sin(\phi + \theta) - \gamma - \frac{a^2}{12}(-\phi_x^2 \sin(\phi + \theta) + \cos(\phi + \theta)[\sin(\phi + \theta) - \gamma]), \quad x \neq 0.$$

Unfortunately the last two equations are not Hamiltonian, so we have lost the Hamiltonian properties of the original system, while the first equation is singularly perturbed.

Yet another approximation that has a variational structure and is not singularly perturbed can be obtained by combining the two equations that have lost their variational character. Indeed, taking (2.7) twice and subtracting (2.8) give

$$\begin{aligned} \phi_{xx} = \phi_{tt} &+ \sin(\phi + \theta) - \gamma \\ &- \frac{a^2}{12} \left(2\phi_{xxtt} + 2\phi_{xx}\cos(\phi + \theta) - \phi_x^2 \sin(\phi + \theta) - \phi_{tttt} - \phi_{tt}\cos(\phi + \theta) \right. \\ &\left. + \phi_t^2 \sin(\phi + \theta) - \cos(\phi + \theta)(\phi_{tt} + \sin(\phi + \theta) - \gamma) \right), \quad x \neq 0. \end{aligned} \tag{2.9}$$

The Lagrangian for this system is

$$L = \iint \tfrac{1}{2}\phi_t^2 - \tfrac{1}{2}\phi_x^2 - 1 + \cos(\phi+\theta) + \gamma\phi + \frac{a^2}{2}\left[\phi_x\partial_x(\phi_{tt} + \sin(\phi+\theta)) + \frac{1}{2}(\phi_{tt} + \sin(\phi+\theta) - \gamma)^2\right] dx\, dt.$$

The static equation for (2.9) is

$$\phi_{xx} = \sin(\phi+\theta) - \gamma - \frac{a^2}{12}\left(2\phi_{xx}\cos(\phi+\theta) - \phi_x^2\sin(\phi+\theta) - \cos(\phi+\theta)(\sin(\phi+\theta)-\gamma)\right), \quad x \neq 0. \tag{2.10}$$

This equation is a regularly perturbed Hamiltonian system with the Hamiltonian

$$H(\phi,p) = \frac{p^2}{2(1+\frac{a^2}{6}\cos(\phi+\theta))} + \gamma\phi + \cos(\phi+\theta) - \frac{a^2}{24}(\sin(\phi+\theta)-\gamma)^2,$$

which implies $p = \phi_x\left(1 + \frac{a^2\cos(\phi+\theta)}{6}\right)$.

In this paper, we will analyze (2.9) as a continuum strong interaction limit which incorporates some effects of the lattice spacing into the model. The model equation (2.9) is chosen as it is nonsingular and has the same conservative properties as the discrete system, reflecting its physical properties.

3. The π-kinks and their spectra in the continuum limit. In this section, we will consider (2.9) for $a = 0$, which is a model for an ideal long 0-π Josephson junction:

$$\phi_{tt} - \phi_{xx} + \sin(\phi+\theta) = \gamma, \quad x \neq 0. \tag{3.1}$$

For a Josephson junction without an applied bias current or a phase jump, i.e., for $\gamma = 0$ and $\theta(x) \equiv 0$, the model corresponds to the sine-Gordon equation. A stable solution of the sine-Gordon equation is the basic (normalized) stationary, monotonically increasing fluxon, given by

$$\phi_{\text{flux}}(x) = 4\arctan e^x, \quad \phi_{\text{flux}}(0) = \pi \tag{3.2}$$

(see [10]).

In general the discontinuous function $\theta(x)$ in (3.1) will introduce a discontinuity at $x = 0$ for the second derivative ϕ_{xx}. Hence, the natural solution space for (3.1) consists of functions which are spatially continuous and have a continuous spatial derivative. The behavior at infinity is regulated by requiring that the spatial derivative of the solution belongs to $H_1(\mathbb{R})$ (which allows the phase to converge to a nonzero constant at infinity). Therefore, (3.1) is considered as a dynamical system on the function space

$$\mathbb{H} = \{\phi : \mathbb{R} \to \mathbb{R} \mid \phi_x \in H_1(\mathbb{R})\}.$$

It is straightforward to find that for $|\gamma| < 1$ and $x < 0$, the "fixed points" of (3.1) are $\phi_s^- = \arcsin(\gamma)$ and $\phi_c^- = -\arcsin(\gamma) + \pi$. Similarly, for $|\gamma| < 1$ and $x > 0$, they

are $\phi_s^+ = \arcsin(\gamma) + \pi$ and $\phi_c^+ = -\arcsin(\gamma) + 2\pi$. In [32], it is shown that there exist various types of stationary fronts, which connect the equilibria. Most stationary fronts are so-called π-kinks, which are static waves connecting equilibrium states at $x = \pm\infty$ with a phase-difference of π. Such waves are solutions of the static wave equation

$$\phi_{xx} - \sin(\phi + \theta) = -\gamma, \quad x \neq 0. \tag{3.3}$$

In the x-dynamics of (3.3), the points $\phi_s^\pm$ are saddle points and the points with $\phi_c^\pm$ are center points. Thus a π-kink connects ϕ_s^- with ϕ_s^+.

In this section we will consider the stability of those π-kinks. For completeness, we first describe the various types of π-kinks as found in [32]. These π-kinks are constructed by taking suitable combinations of the phase portraits for $\theta = 0$ and $\theta = -\pi$. The phase portraits for $\gamma = 0$ are essentially different from the ones for $0 < \gamma < 1$ (the case $-1 < \gamma < 0$ follows from this one by taking $\phi \mapsto -\phi$ and $\gamma \mapsto -\gamma$). In case $\gamma > 0$ there are homoclinic connections at $k\pi + \arcsin(\gamma)$, $k \in \mathbb{Z}$, k even ($\theta = 0$), or k odd ($\theta = -\pi$). If $\gamma = 0$, then these homoclinic connections break to heteroclinic connections between $k\pi$ and $(k+2)\pi$.

The phase portrait of (3.3) for $\gamma = 0$ is shown in Figure 1(a). Following the notation in [32], in case $\gamma = 0$, there are two types of heteroclinic connections (kinks) in the 0-π junction. The first one, called *type* 1 and denoted by $\phi_\pi^1(x;0)$, connects 0 and π. The point in the phase plane where the junction lies is denoted by $d_1(0)$. The second one, called *type* 2 and denoted by $\phi_{3\pi}^2(x;0)$, connects 0 and 3π. Now the point in the phase plane where the junction lies is denoted by $d_2(0)$. This solution is not a semifluxon, but it will play a role in the analysis of some of the semifluxons for $\gamma \neq 0$.

If $0 < \gamma \ll 1$, then there are three types of π-kinks (heteroclinic connections) in the junction, all connecting $\arcsin(\gamma)$ and $\pi + \arcsin(\gamma)$. A phase portrait of (3.3) for nonzero γ is shown in Figure 1(b). The first semifluxon, called *type* 1 and denoted by $\phi_\pi^1(x;\gamma)$, is a continuation of the connection at $\gamma = 0$. The point in the phase plane where the junction lies is denoted by $d_1(\gamma)$. The π-fluxon $\phi_\pi^1(x;\gamma)$ is monotonically increasing.

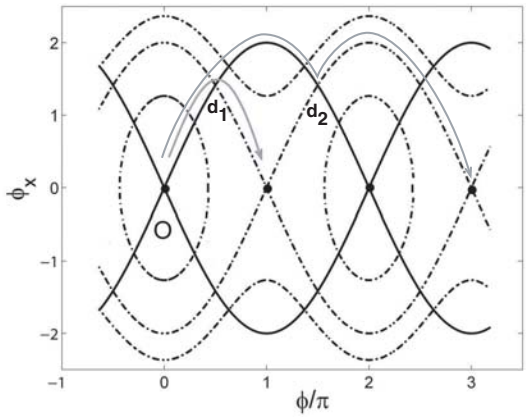
The second one is called *type* 2 and is denoted by $\phi_\pi^2(x;\gamma)$. In the limit for $\gamma \to 0$, it breaks in the 3π-kink and the heteroclinic connection between 3π and π (a -2π-kink or an antifluxon). The point in the phase plane where the junction lies is denoted by $d_2(\gamma)$. The π-fluxon $\phi_\pi^2(x;\gamma)$ is not monotonically increasing but has a hump.

The third one is called *type* 3 and is denoted by $\phi_\pi^3(x;\gamma)$. In the limit for $\gamma \to 0$, it breaks in the heteroclinic connection between 0 and 2π (fluxon) and an antisemifluxon like the type 1 wave but connecting 2π and π. The point in the phase plane where the junction lies is denoted by $d_3(\gamma)$. This π-fluxon has a hump, too, but a lower one than the type 2 wave. Following the first homoclinic orbit, the junction points are ordered such that $d_1(\gamma)$ comes first, followed by $d_2(\gamma)$, followed by $d_3(\gamma)$ (see Figure 1(b)).

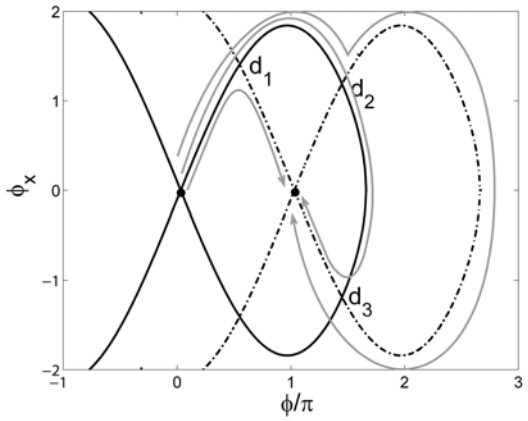
If γ increases, the points $d_2(\gamma)$ and $d_3(\gamma)$ approach each other, until they coincide at [32]

$$\gamma = \gamma^* = \frac{2}{\sqrt{4+\pi^2}} \tag{3.4}$$

in the point $(\pi + \arcsin(\gamma^*), 0)$. At this point, the type 2 wave $\phi_\pi^2(x;\gamma)$ ceases to exist (in the limit it breaks into half the homoclinic connection for $x < 0$ and the full homoclinic connection



(a)



(b)

Figure 1. (a) *The phase portrait of system* (3.3) *for* $\gamma = 0$. *The trajectories for* $x < 0$ *are indicated with bold lines, and the trajectories for* $x > 0$ *are indicated with dashed lines. Any orbit of* (3.1) *switches at* $x = 0$ *from bold to dashed. The type* 1 *semifluxon switches at* d_1 *and corresponds to one of the gray arrow-lines. The* 3π*-fluxon switches at* d_2 *and is denoted by the other gray arrow-line.* (b) *The phase portrait of system* (3.3) *for* $\gamma = 0.1$. *For simplicity, only the stable and unstable manifolds of the fixed points are shown. Apart from* d_1, *there are also the points* d_2 *or* d_3 *which can be used for the switch position of* $x = 0$ *to obtain a solution with a phase difference* π *between the endpoints.*

for $x > 0$). The type 3 kink $\phi_\pi^3(x;\gamma^*)$ consists of half the homoclinic connection for $x < 0$ and the fixed point for $x > 0$, and this wave can be continued for $\gamma > \gamma^*$. For $\gamma > \gamma^*$, the type 3 kink is monotonic.

If γ increases further, the points $d_1(\gamma)$ and $d_3(\gamma)$ approach each other [32] until they coincide at

$$\gamma = \gamma_{\rm cr} = \frac{2}{\pi}. \tag{3.5}$$

When $\gamma = \gamma_{\rm cr}$, the orbit homoclinic to the hyperbolic fixed point for $x < 0$ is tangential at $d_1(\gamma) = d_3(\gamma)$ to the nonhomoclinic stable manifold of the hyperbolic fixed point for $x > 0$.

As soon as $\gamma > \gamma_{\rm cr}$, there is no more intersection of the homoclinic orbit for $x < 0$ with a stable manifold of the hyperbolic fixed point for $x > 0$. This implies that no static π-fluxons can exist. For more details, see [32].

After recalling the description of the π-kinks from [32], we can start the stability analysis. It will be shown that the type 1 π-kink is nonlinearly stable for all $0 \leq \gamma \leq \gamma_{\rm cr}$. The type 2 and type 3 π-kinks are linearly unstable for all values of γ for which they exist. First we consider the linearization about the π-kinks.

Theorem 3.1. *The linearizations about the various π-kinks have the following properties:*

(i) *The eigenvalues of the linearization about the monotonic type 1 π-kink $\phi_\pi^1(x;\gamma)$ are strictly negative for $0 \leq \gamma < \gamma_{\rm cr}$. At $\gamma = \gamma_{\rm cr}$, the largest eigenvalue is zero. These π-kinks are linearly stable.*

(ii) *The largest eigenvalue of the linearization about the type 2 π-kink $\phi_\pi^2(x;\gamma)$ is strictly positive for $0 < \gamma < \gamma^*$. These π-kinks are linearly unstable.*

(iii) *The largest eigenvalue of the linearization about the type 3 π-kink $\phi_\pi^3(x;\gamma)$ is strictly positive for $0 < \gamma < \gamma_{\rm cr}$. These π-kinks are linearly unstable. In the limit for $\gamma \to 0$ and $\gamma \to \gamma_{\rm cr}$, the largest eigenvalue converges to zero.*

Remark 3.2. Note that the instability of the two nonmonotonic π-kinks cannot be established by the classical Sturm–Liouville argument. In the classical, autonomous setting, the derivative of the wave about which the system is linearized is an eigenfunction of the linearized system. This eigenfunction is associated with the translation invariance of the original system and hence corresponds to an eigenvalue $\lambda = 0$. If the wave is nonmonotonic, then its derivative has a zero, which implies that $\lambda = 0$ is not the largest eigenvalue [33] and that the wave must be unstable. Due to the discontinuity at $x = 0$, our system is nonautonomous, and thus not invariant with respect to translations, and $\lambda = 0$ is (in general) not an eigenvalue. Thus, it cannot a priori be concluded that the nonmonotonic π-kinks must be unstable.

To prove Theorem 3.1, it will be shown that the linearization about a π-kink has an eigenvalue of zero if and only if the π-kink takes a value which is a multiple of π at $x = 0$. Since the value at $x = 0$ is related to the point $d_i(\gamma)$, it can be derived that this happens only at $\gamma = \gamma_{\rm cr}$ for the colliding type 1 and type 3 waves. To complete the proof, we will derive expressions for the largest eigenvalue of the linearization about each semikink near $\gamma = 0$ in three separate lemmas and use that the eigenvalues are continuous in γ to derive the sign of the largest eigenvalue on the existence interval of the π-kink.

To linearize about a solution $\phi_\pi^i(x;\gamma)$, write $\phi(x,t) = \phi_\pi^i(x;\gamma) + v(x,t)$, substitute this in the model equation (3.1), and disregard all higher order terms:

$$[D_{xx} - \cos(\phi_\pi^i(x;\gamma) + \theta(x))]\, v = D_{tt}\, v. \tag{3.6}$$

Using the spectral Ansatz $v(x,t) = e^{\lambda t}\widetilde{v}(x)$, where $v(x)$ is a continuously differentiable function, and dropping the tildes, we get the eigenvalue problem

$$\mathcal{L}^i(x;\gamma)\, v = \lambda^2\, v, \tag{3.7}$$

where $\mathcal{L}^i$ is defined as

$$\mathcal{L}^i(x;\gamma) = D_{xx} - \cos(\phi_\pi^i(x;\gamma) + \theta(x)). \tag{3.8}$$

The natural domain for $\mathcal{L}^i$ is $H_2(\mathbb{R})$. We call Λ an eigenvalue of $\mathcal{L}^i$ if there is a function $v \in H_2(\mathbb{R})$, which satisfies $\mathcal{L}^i(x;\gamma)\, v = \Lambda v$. Since $\mathcal{L}^i$ depends smoothly on γ, the eigenvalues of $\mathcal{L}^i$ will depend smoothly on γ, too.

The operator $\mathcal{L}^i$ is symmetric; hence all eigenvalues will be real. A straightforward calculation gives that the continuous spectrum of $\mathcal{L}^i$ is in $(-\infty, -\sqrt{1-\gamma^2})$. Since the eigenfunctions are continuously differentiable functions in $H_2(\mathbb{R})$ by the Sobolev embedding theorem, Sturm's theorem [33] can be applied, leading to the fact that the eigenvalues are bounded from above. Furthermore, if v_1 is an eigenfunction of $\mathcal{L}^i$ with eigenvalue Λ_1 and v_2 is an eigenfunction of $\mathcal{L}^i$ with eigenvalue Λ_2 with $\Lambda_1 > \Lambda_2$, then there is at least one zero of v_2 between any pair of zeros of v_1 (including the zeros at $\pm\infty$). Hence if the eigenfunction v_1 has a fixed sign, then Λ_1 is the largest eigenvalue of $\mathcal{L}^i$.

The following lemma gives a necessary and sufficient condition for $\mathcal{L}^i$ to have an eigenvalue $\Lambda = 0$.

Lemma 3.3. *The eigenvalue problem*

$$\mathcal{L}^i(x;\gamma)v = \Lambda v, \quad x \in \mathbb{R},$$

has an eigenvalue $\Lambda = 0$ if and only if one of the following two conditions holds:

(i) *$D_{xx}\phi^i_\pi(x;\gamma)$ is continuous at $x = 0$; i.e., $\phi^i_\pi(0;\gamma) = k\pi$ for some $k \in \mathbb{Z}$;*

(ii) *$D_x\phi^i_\pi(0;\gamma) = 0$ and there are some $x_\pm$, with $\operatorname{sgn}(x_\pm) = \pm 1$, such that $D_x\phi^i_\pi(x_\pm;\gamma) \neq 0$.*

Proof. Since $\phi^i_\pi(x;\gamma)$ converges to a saddle point for $|x| \to \infty$, this implies that $D_x\phi_\pi(x;\gamma)$ decays exponentially fast to 0 for $|x| \to \infty$. Since $\phi^i_\pi(x;\gamma)$ solves (3.3), differentiating this ODE with respect to x gives

$$\mathcal{L}^i(x;\gamma)\, D_x\phi^i_\pi(x;\gamma) = 0 \quad \text{for} \quad x \neq 0.$$

This implies that for any constant K, the function $w^i_K(x) = K\, D_x\phi^i_\pi(x;\gamma)$ satisfies $\mathcal{L}^i(x;\gamma)\, w^i_K(x) = 0$ for $x \neq 0$. Hence for any K_- and K_+, the solution

$$w^i(x) = \begin{cases} w^i_{K_-}(x), & x < 0, \\ w^i_{K_+}(x), & x > 0, \end{cases}$$

solves $\mathcal{L}^i(x;\gamma)\, w^i(x) = 0$ for $x \neq 0$. The function $w^i(x)$ is continuously differentiable if and only if the following two conditions hold:

1. $w^i_{K_-}(0-) = w^i_{K_+}(0+)$; in other words, $K_-\, D_x\phi^i_\pi(0;\gamma) = K_+\, D_x\phi^i_\pi(0;\gamma)$, since ϕ^i_π is continuously differentiable;
2. $D_x w^i_{K_-}(0-) = D_x w^i_{K_+}(0+)$; thus $K_-\, D_{xx}\phi^i_\pi(0-;\gamma) = K_+\, D_{xx}\phi^i_\pi(0+;\gamma)$.

The first condition is satisfied if $K_- = K_+$ or $D_x\phi^i_\pi(0;\gamma) = 0$. If $D_x\phi^i_\pi(0;\gamma) = 0$, we can choose $K_\pm$ such that the second condition is satisfied and we do not end up with the trivial solution, except when $D_x\phi^i_\pi(x;\gamma)$ is trivial for either $x > 0$ or $x < 0$.

If $D_x\phi^i_\pi(0;\gamma) \neq 0$, we need $D_{xx}\phi^i_\pi$ to be continuous at $x = 0$ in order to satisfy the second condition. Since $D_{xx}\phi^i_\pi(x;\gamma) = \sin(\phi^i_\pi(x;\gamma) + \theta(x)) - \gamma$, $D_{xx}\phi^i_\pi$ is continuous at $x = 0$ if and only if $\sin(\phi^i_\pi(0;\gamma)) = 0$. These arguments prove that if one of the two conditions are satisfied, then $\Lambda = 0$ is an eigenvalue of $\mathcal{L}^i$.

Next we assume that $\Lambda = 0$ is an eigenvalue of $\mathcal{L}^i$; hence there is some continuously differentiable function $v^i(x)$ such that $\mathcal{L}^i(x)v^i(x) = 0$ for $x \neq 0$ and $v^i(x) \to 0$ for $|x| \to \infty$. The only solutions decaying to zero at $+\infty$ are the solutions on the one-dimensional stable manifold, and similarly, the only solutions decaying to zero at $-\infty$ are the solutions on the one-dimensional unstable manifold. The stable and unstable manifolds are formed by multiples of $D_x\phi^i_\pi$. So we can conclude that there exist $K_\pm$ such that

$$v^i(x) = \begin{cases} K_- D_x\phi^i_\pi(x) & \text{for} \quad x < 0, \\ K_+ D_x\phi^i_\pi(x) & \text{for} \quad x > 0. \end{cases}$$

Now we are back in the same situation as above, so we can conclude that either one of the two conditions in the lemma must be satisfied. ∎

The second condition in the lemma does not occur. Indeed, the first part of the second condition, i.e., $D_x\phi^i_\pi(0;\gamma) = 0$, happens only if d_i has its second coordinate zero and hence only at $\gamma = \gamma^*$ with $d_2 = d_3$. At this point, the solution $\phi^2_\pi(x;\gamma^*)$ has ceased to exist and the solution $\phi^3_\pi(x;\gamma^*)$ consists of the fixed point for $x > 0$. Hence this solution does not satisfy the second part of the second condition.

To see for which value of γ the first condition is satisfied, we derive the relation between $\phi^i_\pi(0;\gamma)$ and γ. Multiplying the static equation (3.3) with $D_x\phi^i_\pi$ and rewriting it give

$$D_x[(D_x\phi^i_\pi(x;\gamma))^2] = 2D_x[-\gamma\phi^i_\pi(x;\gamma) - \cos(\phi^i_\pi(x;\gamma) + \theta(x))], \quad x \neq 0.$$

Integration from $\pm\infty$ to 0 and using that $D_x\phi^i_\pi(\pm\infty;\gamma) = 0$ show

$$\begin{aligned}(D_x\phi^i_\pi(0;\gamma))^2 &= 2[-\gamma(\phi^i_\pi(0;\gamma) - \phi^i_\pi(-\infty;\gamma)) - \cos(\phi^i_\pi(0;\gamma)) + \cos(\phi^i_\pi(-\infty;\gamma))], \\ (D_x\phi^i_\pi(0;\gamma))^2 &= 2[-\gamma(\phi^i_\pi(0;\gamma) - \phi^i_\pi(+\infty;\gamma)) + \cos(\phi^i_\pi(0;\gamma)) - \cos(\phi^i_\pi(+\infty;\gamma))].\end{aligned}$$

Subtracting these two equations and using that $\phi^i_\pi(+\infty;\gamma) = \phi^i_\pi(-\infty;\gamma) + \pi$, we get that

$$0 = -\pi\gamma - 2\cos(\phi^i_\pi(0;\gamma)); \quad \text{hence} \quad \cos(\phi^i_\pi(0;\gamma)) = \frac{\pi\gamma}{2}. \tag{3.9}$$

Thus the first condition is only satisfied when $\cos(\phi^i_\pi(0;\gamma)) = \pm 1$; hence $\gamma = \frac{2}{\pi} = \gamma_{\text{cr}}$.

The following step in the analysis of the eigenvalues of the linearization is to consider the behavior of the eigenvalues for γ small. First note that at $\gamma = 0$, we have an explicit expression for the π-fluxon and the 3π-fluxon (see (3.2) for the expression of ϕ_{flux}):

$$\phi^1_\pi(x;0) = \begin{cases} \phi_{\text{flux}}(x - \ln(1+\sqrt{2})) & \text{for} \quad x < 0, \\ \pi - \phi_{\text{flux}}(-x - \ln(1+\sqrt{2})) & \text{for} \quad x > 0, \end{cases} \tag{3.10}$$

$$\phi^2_{3\pi}(x;0) = \begin{cases} \phi_{\text{flux}}(x + \ln(1+\sqrt{2})) & \text{for} \quad x < 0, \\ 3\pi - \phi_{\text{flux}}(-x + \ln(1+\sqrt{2})) & \text{for} \quad x > 0. \end{cases} \tag{3.11}$$

Hence the derivatives of both functions are even and $\cos(\phi^i_\pi(x;0)+\theta)$ is continuous and even, since $\phi^1_\pi(0;0) = \frac{\pi}{2}$ and $\phi^2_{3\pi}(0;0) = \frac{3\pi}{2}$.

For $\gamma \ll 1$, the homoclinic orbit in the system with $\theta = 0$ will be crucial for the approximation of type 2 and type 3 solutions. This orbit is homoclinic to $\arcsin(\gamma)$ and will be

denoted by $\phi_h(x;\gamma)$. It can be approximated up to order γ by using the 2π-fluxon ϕ_{flux} and its linearization.

Lemma 3.4. *For γ small, we have for the even homoclinic connection $\phi_h(x;\gamma)$*

$$\phi_h(x;\gamma) = \phi_{\text{flux}}(x + L_\pi(\gamma)) + \gamma\,\phi_1(x + L_\pi(\gamma)) + \gamma^2 R_2(x + L_\pi(\gamma);\gamma),\ x < 0, \tag{3.12}$$

where the expression for the 2π-fluxon ϕ_{flux} can be found in (3.2),

$$\phi_1(x) = \frac{1}{2}\left[-1 + \cosh x + \int_0^x \frac{\xi}{\cosh\xi} d\xi\right]\frac{1}{\cosh x} - \arctan e^x \left(\frac{x}{\cosh x} + \sinh x\right),$$

and $L_\pi(\gamma)$ is such that $\phi_h(-L_\pi(\gamma);\gamma) = \pi = \phi_{\text{flux}}(0)$, implying

$$L_\pi(\gamma) = \frac{1}{2}|\ln\gamma| + \ln\frac{4}{\sqrt{\pi}} + \mathcal{O}(\sqrt{\gamma}). \tag{3.13}$$

Furthermore, $\gamma^2 R_2(x + L_\pi(\gamma);\gamma) = \mathcal{O}(\gamma)$, uniform for $x < 0$ and $\gamma\phi_1(L_\pi(\gamma);\gamma) = \mathcal{O}(\sqrt{\gamma})$. Thus

$$\phi_h(0) = 2\pi - 2\sqrt{\pi}\sqrt{\gamma} + \mathcal{O}(\gamma). \tag{3.14}$$

Finally, $\phi_1(\tilde{x};\gamma) = \mathcal{O}(1)$ and $R_2(\tilde{x};\gamma) = \mathcal{O}(1)$, uniform for $\tilde{x} < 0$.

Proof. It is more convenient in the following perturbation analysis to follow the normalization of $\phi_{\text{flux}}(x)$; i.e., in this proof we introduce new coordinates $\tilde{x} = x + L_\pi(\gamma)$, where $L_\pi(\gamma)$ is such that $\phi_h(-L_\pi(\gamma);\gamma) = \pi = \phi_{\text{flux}}(0)$. In the following we will drop the tildes and work in those new coordinates. As ϕ_h in the original coordinates was even, we get in the new coordinates $D_x\phi_h(L_\pi(\gamma);\gamma) = 0$. This condition will be used later to determine an asymptotic expression for $L_\pi(\gamma)$.

In the new coordinates, we introduce the expansion

$$\phi_h(x;\gamma) = \phi_{\text{flux}}(x) + \gamma\phi_1(x) + \gamma^2 R_2(x;\gamma), \quad x < L_\pi(\gamma).$$

By linearizing about ϕ_{flux}, it follows that the equation for ϕ_1 is

$$\mathcal{L}(x)\,\phi_1 = -1, \quad \text{where} \quad \mathcal{L}(x) = D_{xx} - \cos(\phi_{\text{flux}}(x)). \tag{3.15}$$

The operator $\mathcal{L}(x)$ is identical to the operator associated with the stability of $\phi_{\text{flux}}(x)$. The homogeneous problem $\mathcal{L}\psi = 0$ has the two independent solutions

$$\psi_b(x) = \frac{1}{\cosh x}, \quad \psi_u(x) = \frac{x}{\cosh x} + \sinh x, \tag{3.16}$$

where $\psi_b(x) = \frac{1}{2}\frac{d}{dx}\phi_{\text{flux}}(x)$ is bounded and $\psi_u(x)$ unbounded as $x \to \pm\infty$. By the variation-of-constants method, we find the general solution to (3.15),

$$\begin{aligned}\phi_1(x;A,B) = &\left[A + \frac{1}{2}\cosh x + \frac{1}{2}\int_0^x \frac{\xi}{\cosh\xi} d\xi\right]\frac{1}{\cosh x}\\ &+ [B - \arctan e^x]\left(\frac{x}{\cosh x} + \sinh x\right),\end{aligned}$$

with $A, B \in \mathbb{R}$. The solution $\phi_1(x)$ of (3.15) must be bounded as $x \to -\infty$ and is normalized by $\phi_1(0) = 0$ (since $\phi_h(0) = \phi_{\text{flux}}(0) = \pi$). Thus, we find that $A = -\frac{1}{2}$ and $B = 0$. Note that $\lim_{x\to-\infty} \phi_1(x) = 1$, which agrees with the fact that $\lim_{x\to-\infty} \phi_h(x) = \arcsin\gamma = \gamma + \mathcal{O}(\gamma^3)$. The solution $\phi_1(x)$ is clearly not bounded as $x \to \infty$, the unbounded parts of $\phi_1(x)$ and $\frac{d}{dx}\phi_1(x)$ are given by

$$\phi_1|_u(x) = -\arctan e^x \sinh x, \quad \frac{d}{dx}\phi_1|_u(x) = -\arctan e^x \cosh x. \tag{3.17}$$

It follows that $\phi_1(x) = \mathcal{O}(\gamma^{-\sigma})$ for some $\sigma > 0$ if $e^x = \mathcal{O}(\gamma^{-\sigma})$, i.e., if $x = \sigma|\ln\gamma|$ at leading order. Using this, it is a straightforward procedure to show that the rest term $\gamma^2 R_2(x;\gamma)$ in (3.12) is of $\mathcal{O}(\gamma^{2-2\sigma})$ for $x = \sigma|\ln\gamma| + \mathcal{O}(1)$ (and $\sigma > 0$). Hence, the approximation of $\phi_h(x)$ by expansion (3.12) breaks down as x becomes of the order $|\ln\gamma|$. On the other hand, it also follows that $\phi^1_{\text{appr}}(x) = \phi_{\text{flux}}(x) + \gamma\phi_1(x)$ is a uniform $\mathcal{O}(\gamma)$-accurate approximation of $\phi_h(x)$ on an interval $(-\infty, L]$ for $L = \frac{1}{2}|\ln\gamma| + \mathcal{O}(1)$. Since $\phi_{\text{flux}}(L) + \gamma\phi_1(L) = \mathcal{O}(\sqrt{\gamma})$ for such L, we can compute $L_\pi = \frac{1}{2}|\ln\gamma| + \mathcal{O}(1)$, as L_π is the value of x at which

$$0 = \frac{d}{dx}\phi_h(x) = \frac{d}{dx}\phi^1_{\text{appr}}(x) + \mathcal{O}(\gamma) = \frac{d}{dx}\phi_{\text{flux}}(x) + \gamma\frac{d}{dx}\phi_1|_u(x) + \mathcal{O}(\gamma).$$

We introduce Y by $e^x = \frac{Y}{\sqrt{\gamma}}$, so that it follows by (3.2) and (3.17) that $Y = \frac{4}{\sqrt{\pi}} + \mathcal{O}(\sqrt{\gamma})$, i.e.,

$$L_\pi(\gamma) = \frac{1}{2}|\ln\gamma| + \ln\frac{4}{\sqrt{\pi}} + \mathcal{O}(\sqrt{\gamma}).$$

A straightforward calculation shows that (in the new coordinates)

$$\phi_h(L_\pi) = 2\pi - 2\sqrt{\pi}\sqrt{\gamma} + \mathcal{O}(\gamma).$$

As $\phi_h(x)$ and $\phi_{\text{flux}}(x)$ both converge exponentially fast to fixed points which are order γ apart for $x \to -\infty$, it follows immediately that $\phi_1(x;\gamma) = \mathcal{O}(1)$ and $R_2(x;\gamma) = \mathcal{O}(1)$, uniform for $x < 0$. ∎

Now we are ready to consider the stability of the various types of π-fluxons individually.

3.1. Stability of the type 1 solution.

Lemma 3.5. *For all $0 \le \gamma < \gamma_{\text{cr}}$, all eigenvalues of $\mathcal{L}^1(x;\gamma)$ are strictly negative. For $\gamma = \gamma_{\text{cr}}$, the operator $\mathcal{L}^1(x;\gamma_{\text{cr}})$ has 0 as its largest eigenvalue. For $\gamma = 0$, the largest eigenvalue is $-\frac{1}{4}(\sqrt{5}+1)$. Furthermore, for all $0 \le \gamma < \gamma_{\text{cr}}$, the type 1 semikinks $\phi^1_\pi(x;\gamma)$ are Lyapunov stable in the following sense. For all $\varepsilon > 0$, there is some $\delta > 0$ such that any solution $\phi(x,t)$ of the semifluxon equation (3.1), which is convergent to 0 at $x \to -\infty$ and to π at $x \to +\infty$ and which satisfies initially $\|\phi(\cdot,0) - \phi^1_\pi(\cdot;\gamma)\|_{H_1} + \|\phi_t(\cdot,0)\|_{L_2} < \delta$, will satisfy $\|\phi(\cdot,t) - \phi^1_\pi(\cdot;\gamma)\|_{L_2} + \|\phi_t(\cdot,t)\|_{L_2} < \varepsilon$ for all $t \in \mathbb{R}$.*

Proof. From Lemma 3.3 it follows that $\mathcal{L}^1$ has an eigenvalue $\Lambda = 0$ at $\gamma = \gamma_{\text{cr}}$. The eigenfunction is $D_x\phi^1_\pi(x;\gamma_{\text{cr}})$ and this function is always positive, since $\phi^1_\pi(x;\gamma_{\text{cr}})$ is monotonically increasing. From Sturm's theorem, it follows that $\Lambda = 0$ is the largest eigenvalue of $\mathcal{L}^1$ at $\gamma = \gamma_{\text{cr}}$. Next we consider $\gamma = 0$. We can explicitly determine all eigenvalues of $\mathcal{L}^1(x;0)$. From the explicit expression for ϕ^1_π it follows that $\mathcal{L}^1(x;0)$ is a continuous even operator. For

fixed Λ, the operator $\mathcal{L}^1(x;0) - \Lambda$ has two linearly independent solutions. Since the fixed point is a saddle point and the decay rate to this fixed point is like e^{-x}, there is one solution that is exponentially decaying at $+\infty$ and there is one solution that is exponentially decaying at $-\infty$, if $\Lambda > -1$. If we denote the exponentially decaying function at $-\infty$ by $v_-(x;\Lambda)$, then the exponentially decaying function at $+\infty$ up to a constant is given by $v_+(x;\Lambda) = v_-(-x;\Lambda)$ (since $\mathcal{L}^1$ is symmetric in x). Obviously, $v_+(0;\Lambda) = v_-(0;\Lambda)$; hence Λ is an eigenvalue if $D_x v_+(0;\Lambda) = D_x v_-(0;\Lambda)$ (i.e., when $D_x v_-(0;\Lambda) = 0$) or if $v_-(0;\Lambda) = 0$.

Using [23], we can derive explicit expressions for the solutions $v_-(x;\Lambda)$ (see also [10]). Using $x_1 = \ln(\sqrt{2}+1)$, we have

$$v_-(x;0) = \operatorname{sech}(x - x_1), \quad v_-(x;\Lambda) = e^{\mu(x-x_1)}\,[\tanh(x-x_1) - \mu], \quad \mu = \sqrt{\Lambda+1}.$$

A straightforward calculation shows that $v_-(0;\Lambda) \neq 0$. The condition $D_x v_-(0;\Lambda) = 0$ gives that

$$\mu^2 - \frac{1}{2}\sqrt{2}\mu - \frac{1}{2} = 0; \quad \text{hence} \quad \sqrt{\Lambda+1} = \frac{1}{4}\sqrt{2}(\sqrt{5}-1) \Rightarrow \Lambda = -\frac{1}{4}(\sqrt{5}+1).$$

Now assume that the operator $\mathcal{L}^1(x;\gamma)$ has a positive eigenvalue $\Lambda^1(\gamma)$ for some $0 \leq \gamma < \gamma_{\mathrm{cr}}$. Since Λ depends continuously on γ, there has to be some $0 < \widehat{\gamma} < \gamma_{\mathrm{cr}}$ such that $\Lambda^1(\widehat{\gamma}) = 0$. However, from Lemma 3.3 it follows that this is not possible.

Nonlinear or Lyapunov stability can be derived by looking at the "temporal Hamiltonian"

$$\mathcal{H}(\phi,p) = \int_{-\infty}^{\infty} \left[\frac{1}{2}p^2 + \frac{1}{2}(\phi_x)^2 - \cos(\phi+\theta) - \gamma(\phi+\theta)\right] dx.$$

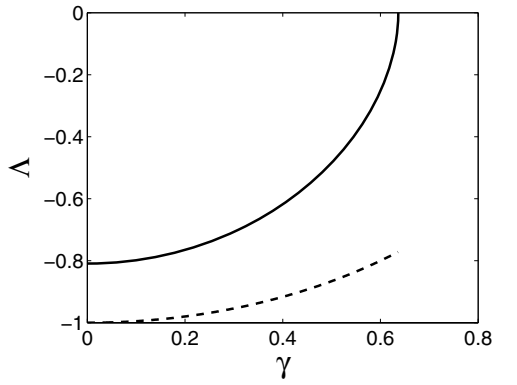
This functional is a Lyapunov function for the system (3.1); i.e., any solution $\phi(x,t) \in H^2(\mathbb{R})$ of (3.1) satisfies $\frac{d}{dt}\mathcal{H}(\phi,\phi_t) = 0$, and hence $\mathcal{H}(\phi(\cdot,t),\phi_t(\cdot,t)) = \mathcal{H}(\phi(\cdot,0),\phi_t(\cdot,0))$ for any $t \in \mathbb{R}$. Furthermore, the linearization $D^2\mathcal{H}$ at $(\phi,p) = (\phi_\pi^1, 0)$ (the point related to the π-fluxon) is given by

$$D^2\mathcal{H}(\phi_\pi^1,0) = \begin{pmatrix} -\mathcal{L}^1(x;\gamma) & 0 \\ 0 & I \end{pmatrix},$$

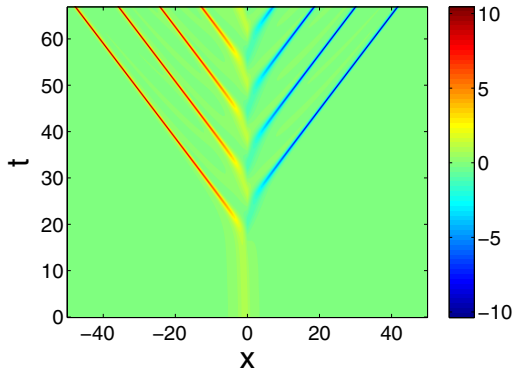
which is a strictly positive definite self-adjoint operator on $L_2(\mathbb{R})\times L_2(\mathbb{R})$ with domain $H_2(\mathbb{R})\times L_2(\mathbb{R})$. So there is some $c > 0$ such that for any $(\phi,p) \in H_2 \times L_2$, we have $H(\phi,p) - H(\phi_\pi^1,0) \geq c(\|\phi-\phi_\pi^1\|_{L_2}^2 + \|p\|_{L_2}^2)$; see, e.g., [14, 36]. Finally, it is straightforward to prove that there is some $C > 0$ such that $H(\phi,p) - H(\phi_\pi^1,0) \leq C(\|\phi-\phi_\pi^1\|_{H_1}^2 + \|p\|_{L_2}^2)$ for any $(\phi,p) \in H_2 \times L_2$. ■

Using standard procedures in MATLAB, the eigenvalues of the type 1 π-fluxon have been calculated numerically as a function of the applied bias current γ and are presented in Figure 2(a). Further details of the computational procedure are presented in section 6. Figure 2(a) shows that the type 1 semifluxon has only one eigenvalue. This eigenvalue tends to zero when the bias current γ approaches the critical value γ_{cr} as has been derived analytically. It was first proposed in [15, 20, 21] that a constant driving force can excite the largest eigenvalue of a semifluxon toward zero.

When we apply a bias current above the critical value γ_{cr}, numerics show that the stationary π-kink bifurcates into a semifluxon that reverses its polarity and releases a fluxon.



(a)



(b)

Figure 2. (a) *The eigenvalue of linear operator associated to the type* 1 *semifluxon as a function of the bias current* γ*. The dashed line is the boundary of the continuous spectrum.* (b) *A simulation of the evolution of a* π*-kink in the presence of a bias current above the critical value* ($\gamma > \gamma_{\rm cr}$)*. The plot presents the magnetic field* ϕ_x*. The numerics show that the instability leads to the release of wave trains of traveling wave fluxons. In this evolution a damping, which is proportional to* ϕ_t*, has been applied to the system.*

This process keeps repeating itself: the semifluxon changes its direction back and forth with releasing a fluxon or antifluxon in every change. A simulation of the release of fluxons from a semifluxon is presented in Figure 2(b). In experiments, the polarity of a semifluxon can also be reversed by applying a magnetic field [13].

When $\gamma = \gamma_{\rm cr}$, the type 1 and type 3 semifluxons coincide. From the numerical analysis of the eigenvalues of the type 3 semifluxon (see section 3.3 for details), it follows that there is an eigenvalue at the edge of the continuous spectrum for $\gamma = \gamma_{\rm cr}$. We conjecture that this eigenvalue bifurcates into the edge of the continuous spectrum at this point as γ increases to $\gamma_{\rm cr}$ (see Figure 4).

3.2. Instability of type 2 solutions.

Lemma 3.6. *For all* $0 < \gamma < \gamma^*$*, the largest eigenvalue of* $\mathcal{L}^2(x;\gamma)$ *is strictly positive. In the limit* $\gamma \to 0$*, the largest eigenvalue of* $\mathcal{L}^2(x;\gamma)$ *converges to* $\frac{1}{4}(\sqrt{5}-1)$.

Proof. Using the approximation for the homoclinic orbit $\phi_h(x;\gamma)$ in Lemma 3.4, we see that, for γ small, an approximation for the π-fluxon of type 2 is given by (as before, $x_1 = \ln(1+\sqrt{2})$)

$$\phi_\pi^2(x;\gamma) = \begin{cases} \phi_{\text{flux}}(x+x_1) + \mathcal{O}(\gamma), & x < 0, \\ \pi + \phi_{\text{flux}}(\widetilde{x}) + \gamma\phi_1(\widetilde{x}) + \gamma^2 R_2(\widetilde{x};\gamma), & 0 < x < L_\pi(\gamma) + x_1, \\ \pi + \phi_{\text{flux}}(-\widehat{x}) + \gamma\phi_1(-\widehat{x}) + \gamma^2 R_2(-\widehat{x};\gamma), & x > L_\pi(\gamma) + x_1, \end{cases} \tag{3.18}$$

with $\widetilde{x} = x - x_1$ and $\widehat{x} = x - 2L_\pi(\gamma) - x_1$.

There is no limit for $\gamma \to 0$, since the semifluxon breaks into two parts, one of them being the 3π-fluxon $\phi_{3\pi}^2(x;0)$. In a similar way as we found the largest eigenvalue for the linearization operator $\mathcal{L}^1(x;0)$ about the π-fluxon $\phi_\pi^1(x;0)$, we can find the largest eigenvalue for the linearization operator $\mathcal{L}^2(x;0)$ about the 3π-fluxon $\phi_{3\pi}^2(x;0)$. The largest eigenvalue is $\Lambda^2(0) = \frac{1}{4}(\sqrt{5}-1)$ and the eigenfunction is

$$\psi^2(x;0) = \begin{cases} e^{\mu_0(x+x_1)}(\mu_0 - \tanh(x+x_1)), & x < 0, \\ e^{\mu_0(-x+x_1)}(\mu_0 - \tanh(-x+x_1)), & x > 0, \end{cases}$$

where $\mu_0 = \sqrt{\Lambda^2(0)+1} = \frac{1}{4}\sqrt{2}(1+\sqrt{5})$. (It can be shown that there is another smaller eigenvalue $\Lambda = -\frac{1}{2}$ and similar eigenfunction if $\mu = \frac{1}{2}\sqrt{2} = \tanh(x_1)$; see Remark 3.8.)

In a similar way, using the approximation (3.18) for $\phi_\pi^2(x;\gamma)$, the eigenfunction of an eigenvalue of ϕ_π^2 for γ small is approximated by

$$\psi^2(x;\gamma) = \begin{cases} e^{\mu(x+x_1)}(\mu - \tanh(x+x_1)) + \mathcal{O}(\sqrt{\gamma}), & x < 0, \\ k_2\, e^{-\mu\widetilde{x}}(\mu - \tanh(-\widetilde{x})) + k_3\, e^{\mu\widetilde{x}}(\mu - \tanh\widetilde{x}) + \mathcal{O}(\sqrt{\gamma}), & 0 < x < L_\pi(\gamma) + x_1, \\ k_4\, e^{\mu(-\widehat{x})}(\mu + \tanh\widehat{x}) + \mathcal{O}(\sqrt{\gamma}), & x > L_\pi(\gamma) + x_1, \end{cases}$$

where k_i and μ have to be determined. The eigenvalue Λ follows from $\mu = \sqrt{\Lambda^2+1}$. Note that the secular term which is growing at infinity with the multiplication factor k_3 is included in this approximation. When $\gamma = 0$ and $k_3 = 0$, the first two lines in the definition of ψ^2 are the eigenfunction of the linearized problem about the heteroclinic connection between 0 and 3π, as presented above. When γ is nonzero, k_3 can be of order $\mathcal{O}(\gamma^\sigma)$ for $\sigma > \frac{\mu}{2}$ as the secular term is of order $\mathcal{O}(\gamma^{-\mu/2})$ at $x = L_\pi(\gamma) + x_1$.

The constants k_2, k_3, and k_4 and the parameter μ have to be chosen such that for $\gamma > 0$ (but small) the function $\psi^2(x,\gamma)$ is continuously differentiable at $x = 0$ and $x = L_\pi(\gamma) + x_1$. From the continuity conditions at $x = 0$, we obtain

$$\begin{aligned} k_2 &= \frac{\sqrt{2}}{4\mu(\mu-1)(\mu+1)} + \mathcal{O}(\sqrt{\gamma}), \\ k_3 &= \frac{(3+2\sqrt{2})^\mu(2\mu^2 - \mu\sqrt{2} - 1)(2\mu - \sqrt{2})}{4\mu(\mu^2-1)} + \mathcal{O}(\sqrt{\gamma}). \end{aligned}$$

From one of the continuity conditions at $x = L_\pi(\gamma) + x_1$, we obtain $k_4 = k_4(k_2, k_3, \mu)$. Now we are left with one more matching condition. Values of μ for which this condition is

satisfied correspond to the eigenvalues of the operator $\mathcal{L}^2(x;\gamma)$ for γ small. More explicitly, the spectral parameter μ has to satisfy the equation

$$\mathcal{F}(\mu) = 16^{\mu} k_3 (\mu-1)^2 (\gamma\pi)^{-\mu}((3\mu+4)\pi\gamma + 16\mu) + \mathcal{O}(\gamma^{-\mu+2}) = 0. \tag{3.19}$$

Note that this expression is not defined at $\gamma = 0$. This corresponds to the singularities in the expression for ϕ^2 as $\gamma \to 0$ due to the fact that $L_\pi(\gamma) \to \infty$ for $\gamma \to 0$. Evaluating $\mathcal{F}(\mu)\,(\gamma\pi)^{\mu}$ at $\gamma = 0$, we see that there are four positive roots for μ, leading to four squared eigenvalues, namely, $\Lambda(0) = \frac{1}{4}(\sqrt{5}-1)$, $-\frac{1}{2}$, and the double eigenvalue $\Lambda(0) = 0$. The first two come from the zeros of k_3 and are related to the eigenvalues of the 3π-fluxon. The double zero eigenvalues are the eigenvalues of the fluxon. One can also notice that there is no term with a multiplication factor k_2 to this leading order. This term appears at most of order $\mathcal{O}(\gamma^{\mu+2})$. Finally, as with the type 1 semifluxon, the root $\mu = 0$ corresponds to the edge of the continuous spectrum and the "eigenfunction" is not in $H_2(\mathbb{R})$.

The proof that the largest eigenvalue is near $\frac{1}{4}(\sqrt{5}-1)$ for γ small will be complete if we can show that $\mathcal{F}_\mu(\sqrt{2}/4(1+\sqrt{5})) \neq 0$, i.e., the nondegeneracy condition that says that the eigenvalue can be continued continuously for γ small.

Simple algebraic calculations give that

$$\mathcal{F}_\mu\left(\frac{\sqrt{2}}{4}(1+\sqrt{5})\right) = c_1 \gamma^{-\frac{\sqrt{2}}{4}(1+\sqrt{5})} + \mathcal{O}(\gamma^{1-\frac{\sqrt{2}}{4}(1+\sqrt{5})}) \tag{3.20}$$

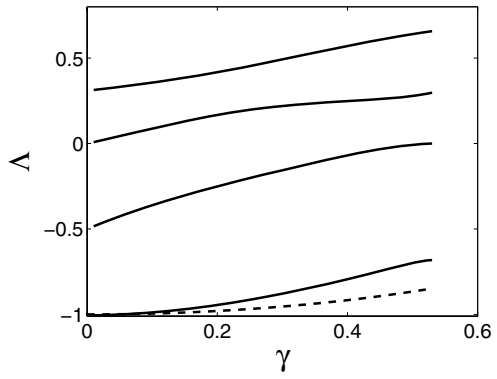
with c_1 a positive constant. Hence, $\mathcal{F}_\mu(\frac{\sqrt{2}}{4}(1+\sqrt{5})) > 0$.

This completes the proof that the largest eigenvalue is near $\frac{1}{4}(\sqrt{5}-1)$ for small but positive γ. Since the largest eigenvalue depends continuously on γ, it can only disappear at a bifurcation point. There are no bifurcation points and it is not possible that the eigenvalue becomes 0 (see Lemma 3.3); hence the largest eigenvalue will be positive as long as fluxon $\phi^2_\pi(x;\gamma)$ exists, i.e., for $0 < \gamma < \gamma^*$. ■

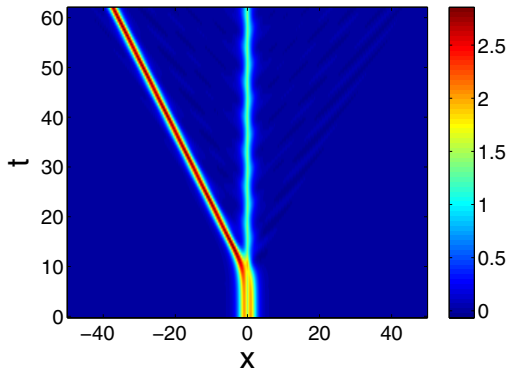
Remark 3.7. We cannot use a comparison theorem, because $\phi^2_\pi < \phi^3_\pi$ for $x < 0$ and $\phi^2_\pi > \phi^3_\pi$ for $x > 0$.

To consider the relation between the eigenvalues of $\mathcal{L}^2(x;\gamma)$ and the stability problem of $\phi^2_\pi(x;\gamma)$, we denote the largest eigenvalue of $\mathcal{L}^2(x;\gamma)$ by $\Lambda^2(\gamma)$. The associated eigenvalues for the linearizations are solutions of the equation $\lambda^2 - \Lambda^2(\gamma) = 0$; hence $\lambda = \pm\sqrt{\Lambda^2(\gamma)}$. Since $\Lambda^2(\gamma) > 0$, this implies that one of the two eigenvalues has positive real part; hence the π-fluxons of type 2 are unstable. The numerically obtained eigenvalues of semifluxons of this type as a function of γ are shown in Figure 3(a). In the proof of Lemma 3.6 we have found three different eigenvalues for γ small and the possibility of a fourth eigenvalue coming out of the continuous spectrum at $\gamma = 0$. In Figure 3(a), we see the continuation of those eigenvalues. In Figure 3(b), we present the evolution of a 3π-kink (3.11) which is the limit of a type 2 semifluxon when $\gamma \to 0$. The separation of a fluxon from the semifluxon is clearly seen and indicates the instability of the state (which confirms the analysis in the proof of Lemma 3.6).

Remark 3.8. A type 2 semifluxon can be seen as a concatenation of a 3π- and a -2π-kink in the limit $\gamma \to 0$. In that limit the other eigenvalues of $\mathcal{L}^2(x;\gamma)$ converge to 0, $-\frac{1}{2}$, and -1.



(a)



(b)

Figure 3. (a) *The eigenvalues of the linear operator associated to the type* 2 *semifluxon as a function of the bias current* γ. *The result that the largest eigenvalue is always positive shows the instability of the type* 2 *semifluxon. When* $\gamma \to 0$, $\Lambda \to \frac{1}{4}(\sqrt{5}-1)$ *which is the largest eigenvalue of a* 3π*-kink. At* $\gamma = 0$, *one eigenvalue comes out of the edge of the continuous spectrum (dashed line).* (b) *The evolution of a* 3π*-kink* (3.11) *for* $\gamma = 0$. *The plot is presented in terms of the magnetic field* ϕ_x. *The separation of a fluxon from the semifluxon can be clearly seen.*

The eigenvalues 0 and -1 are contributions of the -2π-kink. The eigenvalue $-\frac{1}{2}$ corresponds to the first excited state of the 3π-kink with eigenfunction

$$\psi^2(x;0) = \begin{cases} e^{\mu(x+x_1)}(\mu - \tanh(x+x_1)), & x < 0, \\ e^{\mu(-x+x_1)}(\tanh(-x+x_1) - \mu), & x > 0, \end{cases}$$

where $\mu = \sqrt{\Lambda + 1} = \frac{1}{\sqrt{2}}$.

3.3. Instability of type 3 solutions.

Lemma 3.9. *For all* $0 < \gamma < \gamma_{\mathrm{cr}}$, *the largest eigenvalue of* $\mathcal{L}^3(x;\gamma)$ *is strictly positive. For* $\gamma = \gamma_{\mathrm{cr}}$, *the operator* $\mathcal{L}^3(x;\gamma_{\mathrm{cr}})$ *has* 0 *as its largest eigenvalue.*

Proof. The solution $\phi_\pi^3(x,\gamma_{\rm cr}) = \phi_\pi^1(x,\gamma_{\rm cr})$; hence from Lemma 3.5 it follows that the largest eigenvalue is $\Lambda = 0$.

For γ near zero, we will use the approximation for the homoclinic orbit $\phi_h(x;\gamma)$ in Lemma 3.4 to get an approximation for the type 3 fluxon

$$\phi_\pi^3(x;\gamma) = \begin{cases} \phi^1_{\rm appr}(\widehat{x}) = \phi_{\rm flux}(\widehat{x}) + \gamma\phi_1(\widehat{x}) + \gamma^2 R_2(\widehat{x};\gamma), & x < -L_\pi(\gamma) + x_1, \\ \phi^2_{\rm appr}(\widetilde{x}) = \phi_{\rm flux}(-\widetilde{x}) + \gamma\phi_1(-\widetilde{x}) + \gamma^2 R_2(-\widetilde{x};\gamma), & -L_\pi(\gamma) + x_1 < x < 0, \\ \phi^3_{\rm appr}(-x - x_1) = \pi + \phi_{\rm flux}(-x - x_1) + \mathcal{O}(\gamma), & x > 0, \end{cases}$$

where $\widetilde{x} = x - x_1$ and $\widehat{x} = x - x_1 + 2L_\pi(\gamma)$.

In the limit $\gamma \to 0$, the type 3 semifluxon breaks into a type 1 semifluxon and a fluxon. Both are stable and the largest eigenvalue of the fluxon is zero, while the largest eigenvalue of the type 1 semifluxon is negative. Hence to approximate the largest eigenvalue of the type 3 semifluxon for γ small, we set

$$\Lambda(\gamma) = \gamma\Lambda_1(\gamma).$$

To construct the first part of the approximation of the eigenfunction, we consider $x < -L_\pi(\gamma) + x_1$; i.e., $\widehat{x} < L_\pi(\gamma)$. In this part of the argument, we will drop the hat in $\widehat{x}$. On $(-\infty, L_\pi)$, we expand $\psi^1_{\rm approx} = \psi_0 + \gamma\psi_1$; this yields the following equations for $\psi_{0,1}(x)$:

$$\mathcal{L}\psi_0 = 0, \quad \mathcal{L}\psi_1 = [\Lambda_1(0) - \phi_1(x)\sin\phi_{\rm flux}(x)]\psi_0. \tag{3.21}$$

As $\psi^1_{\rm approx}$ has to be an eigenfunction, we have $\psi^1_{\rm approx}(x) \to 0$ as $x \to -\infty$. Furthermore, we remove the scaling invariance by assuming that $\psi^1_{\rm approx}(0) = 1$. This implies that $\psi_0(x)$ is given by

$$\psi_0(x) = \frac{1}{\cosh x} \tag{3.22}$$

(see (3.16)). To solve the ψ_1-equation, we note that $\frac{d}{dx}\phi_1(x)$ is a solution of

$$\mathcal{L}\psi = -\phi_1 \sin\phi_{\rm flux}\frac{d}{dx}\phi_{\rm flux} = -2\phi_1\sin\phi_{\rm flux}\psi_0$$

(see (3.15) and (3.2)) so that we find as a general solution

$$\begin{aligned}\psi_1(x) = &\left[A - \frac{1}{2}\Lambda_1\left(\ln(\cosh x) + \int_0^x \frac{\xi}{\cosh^2\xi}d\xi\right)\right]\frac{1}{\cosh x} \\ &+ \left[B + \frac{1}{2}\Lambda_1\tanh x\right]\left(\frac{x}{\cosh x} + \sinh x\right) + \frac{1}{2}\frac{d}{dx}\phi_1.\end{aligned}$$

Using $\lim_{x\to-\infty}\psi_1(x) = 0$ and $\psi_1(0) = 0$ we find that $A = \frac{\pi}{4}$, $B = \frac{1}{2}\Lambda_1(0)$. As in the case of $\phi_1(x)$, we are especially interested in the unbounded parts of $\psi_1(x)$ and $\frac{d}{dx}\psi_1(x)$,

$$\begin{aligned}\psi_1|_u(x) &= \tfrac{1}{2}\Lambda_1(1+\tanh x)\sinh x - \tfrac{1}{2}\arctan e^x\cosh x, \\ \tfrac{d}{dx}\psi_1|_u(x) &= \tfrac{1}{2}\Lambda_1(1+\tanh x)\cosh x - \tfrac{1}{2}\arctan e^x\sinh x.\end{aligned} \tag{3.23}$$

We note that the error term $|\psi(x) - \psi^1_{\text{appr}}(x)| = \gamma^2|S_2(x;\gamma)|$ is at most $\mathcal{O}(\gamma)$ on $(-\infty, L_\pi)$ (the analysis is similar to that for $\gamma^2|R_2(x;\gamma)|$).

Next consider the second part of the approximation, i.e., x between $-L_\pi(\gamma) + x_1$ and 0. Here we define the translated coordinate $\widetilde{x} = x - x_1$, which is on the interval $(-L_\pi, -x_1)$, and again we drop the tildes. Since we have to match $\psi^1_{\text{appr}}(x)$ to the approximation $\psi^2_{\text{appr}}(x)$ of $\psi(x)$, along $\phi^2_{\text{appr}}(x)$ and thus defined on the interval $(-L_\pi, -x_1)$, we need to compute $\psi^1_{\text{appr}}(L_\pi)$ and $\frac{d}{dx}\psi^1_{\text{appr}}(L_\pi)$ which to the leading order are calculated from (3.23); i.e.,

$$\psi^1_{\text{appr}}(L_\pi) = \frac{2\Lambda_1(0)}{\sqrt{\pi}}\sqrt{\gamma} + \mathcal{O}(\gamma), \quad \frac{d}{dx}\psi^1_{\text{appr}}(L_\pi) = \frac{2\Lambda_1(0) - \pi}{\sqrt{\pi}}\sqrt{\gamma} + \mathcal{O}(\gamma). \tag{3.24}$$

Thus, both $\psi^1_{\text{appr}}(L_\pi)$ and $\frac{d}{dx}\psi^1_{\text{appr}}(L_\pi)$ are $\mathcal{O}(\sqrt{\gamma})$.

Now, we choose a special form for $\psi^2_{\text{appr}}(x)$, the continuation of $\psi(x)$, i.e., the part linearized along $\phi^2_{\text{appr}}(x)$. It is our aim to determine the value of Λ_1, for which there exists a positive integrable C^1 solution ψ of $\mathcal{L}^3(x;\gamma)\psi = \gamma\Lambda_1(0)\psi$. By general Sturm–Liouville theory [33] we know that this value of Λ_1 must be the largest eigenvalue. Our strategy is to try to continue $\psi(x)$ beyond $(-\infty, L_\pi)$ by a function that remains at most $\mathcal{O}(\sqrt{\gamma})$; i.e., we do not follow the approach of the existence analysis and thus do not reflect and translate $\psi^1_{\text{appr}}(x)$ to construct $\psi^2_{\text{appr}}(x)$ (since this solution becomes in general $\mathcal{O}(1)$ for $x = \mathcal{O}(1)$). Instead, we scale $\psi^2_{\text{appr}}(x)$ as $\gamma\widetilde{\psi}(x)$. The linearization $\widetilde{\psi}(x)$ along $\phi^2_{\text{appr}}(x)$ on the interval $(-L_\pi, x_1)$ must solve $\mathcal{L}\widetilde{\psi} = O(\gamma)$; thus, at leading order

$$\widetilde{\psi}(x) = \frac{\tilde{A}}{\cosh x} + \tilde{B}\left(\frac{x}{\cosh x} + \sinh x\right). \tag{3.25}$$

The approximation $\psi^2_{\text{appr}}(x) = \gamma\widetilde{\psi}(x)$ must be matched to $\psi^1_{\text{appr}}(L_\pi)$ and $\frac{d}{dx}\psi^1_{\text{appr}}(L_\pi)$ at $x = -L_\pi$; i.e.,

$$\frac{2\Lambda_1(0)}{\sqrt{\pi}} = -\frac{2\tilde{B}}{\sqrt{\pi}} + \mathcal{O}(\sqrt{\gamma}), \quad \frac{2\Lambda_1(0) - \pi}{\sqrt{\pi}} = \frac{2\tilde{B}}{\sqrt{\pi}} + \mathcal{O}(\sqrt{\gamma}).$$

Note that $\tilde{A}$ does not appear in these equations; as a consequence, $\psi^1_{\text{appr}}(x)$ and $\psi^2_{\text{appr}}(x)$ can only be matched for a special value of Λ_1, $\Lambda_1(0) = \frac{1}{4}\pi$, with $\tilde{B} = -\Lambda_1(0) < 0$. Thus for this special value of Λ_1 and for $\tilde{A} > 0$, we have found a positive C^1-continuation of the solution $\psi(x)$ of the eigenvalue problem for $\mathcal{L}^3(x;\gamma)$—recall that $x < 0$ in the domain of $\widetilde{\psi}(x)$. At the point of discontinuity ($-x_1$ for $\widetilde{\psi}(x)$, or at $x = 0$ in the original coordinates of (3.1)), we have

$$\begin{aligned}
&\psi^2_{\text{appr}}(-x_1) = \gamma\widetilde{\psi}(-x_1) = \gamma[\tfrac{1}{2}\sqrt{2}\tilde{A} - \tfrac{\pi}{8}\sqrt{2}(\ln(\sqrt{2}-1) - \sqrt{2})] + \mathcal{O}(\gamma^2),\\
&\tfrac{d}{dx}\psi^2_{\text{appr}}(-x_1) = \gamma\tfrac{d}{dx}\widetilde{\psi}(-x_1) = \gamma[\tfrac{1}{2}\tilde{A} - \tfrac{\pi}{8}(\ln(\sqrt{2}-1) + 3\sqrt{2})] + \mathcal{O}(\gamma^2).
\end{aligned} \tag{3.26}$$

Hence, we have constructed for a special choice of Λ, $\Lambda = \Lambda_* = \frac{\pi}{4}\gamma + \mathcal{O}(\gamma\sqrt{\gamma}) > 0$, an approximation of a family of positive solutions of the eigenvalue problem for $\mathcal{L}^3(x;\gamma)$ on $x < 0$—in the coordinates of (3.1)—that attain the values given by (3.26) at $x = 0$, and that decay to 0 as $x \to -\infty$. The question is now whether we can "glue" an element of this family in a C^1-fashion to a solution of the eigenvalue problem for $\mathcal{L}^3(x;\gamma)$ on $x > 0$—with

$\Lambda = \Lambda_*$—that decays (exponentially) as $x \to \infty$. If that is possible, we have constructed a positive integrable solution to the eigenvalue problem for $\mathcal{L}^3(x;\gamma)$, which implies that $\Lambda_* > 0$ is the critical eigenvalue and that $\phi_\pi^3(x)$ is unstable.

An approximation of $\psi(x)$ on $x > 0$, $\psi^3_{\text{appr}}(x)$, is obtained by linearizing along $\phi^3_{\text{appr}}(x)$ and by translating x so that $x \in (x_1, \infty)$. Since $\psi^3_{\text{appr}}(x)$ has to match to expressions of $\mathcal{O}(\gamma)$ (3.26) at x_1, we also scale $\psi^3_{\text{appr}}(x)$, $\psi^3_{\text{appr}}(x) = \gamma\hat{\psi}(x)$. We find that $\mathcal{L}\hat{\psi} = \mathcal{O}(\gamma)$ so that $\hat{\psi}(x)$ again has to be (at leading order) a linear combination of $\psi_b(x)$ and $\psi_u(x)$ (3.16). However, $\hat{\psi}$ must be bounded as $x \to \infty$, which yields that $\hat{\psi}(x) = \hat{A}/\cosh x + \mathcal{O}(\gamma)$ for some $\hat{A} \in \mathbb{R}$. At the point of discontinuity we thus have

$$\begin{aligned} \psi^3_{\text{appr}}(x_1) &= \gamma\hat{\psi}(x_1) &&= \tfrac{1}{2}\sqrt{2}\hat{A}\gamma + \mathcal{O}(\gamma^2), \\ \tfrac{d}{dx}\psi^3_{\text{appr}}(x_1) &= \gamma\tfrac{d}{dx}\widetilde{\psi}(x_1) &&= -\tfrac{1}{2}\hat{A}\gamma + \mathcal{O}(\gamma^2). \end{aligned} \tag{3.27}$$

A positive C^1-solution of the eigenvalue problem for $\mathcal{L}^3(x;\gamma)$ exists (for $\Lambda = \Lambda_*$) if there exist $\tilde{A}, \hat{A} > 0$ such that (see (3.26) and (3.27))

$$\begin{aligned} \tfrac{1}{2}\sqrt{2}\tilde{A} \;-\; \tfrac{\pi}{8}\sqrt{2}(\ln(\sqrt{2}-1) - \sqrt{2}) &= \tfrac{1}{2}\sqrt{2}\hat{A}, \\ \tfrac{1}{2}\tilde{A} \;-\; \tfrac{\pi}{8}(\ln(\sqrt{2}-1) + 3\sqrt{2}) &= -\tfrac{1}{2}\hat{A}. \end{aligned} \tag{3.28}$$

Since the solution of this system is given by $\tilde{A} = \frac{1}{4}\pi[\sqrt{2} + \ln(\sqrt{2}-1)] > 0$ and $\hat{A} = \frac{1}{2}\pi\sqrt{2} > 0$, we conclude that the eigenvalue problem for the π-fluxon $\phi_\pi^3(x;\gamma)$ has a positive largest eigenvalue

$$\Lambda_* = \frac{\pi}{4}\gamma + \mathcal{O}(\gamma\sqrt{\gamma}). \tag{3.29}$$

Hence the eigenvalue for γ small is positive. From Lemma 3.3 it follows that there are no zero eigenvalues between 0 and γ_{cr}; hence the largest eigenvalue of $\mathcal{L}^3(\gamma)$ is positive for all values of γ. ∎

Remark 3.10. For any $\lambda = \mathcal{O}(\sqrt{\gamma})$ or, equivalently, any $\Lambda_1 = \mathcal{O}(1)$, there exists a (normalized) solution to the eigenvalue problem for $\mathcal{L}^3(x;\gamma)$ on $x < 0$ that decays as $x \to -\infty$, and that is approximated by $\psi^1_{\text{appr}}(x)$ and $\psi^2_{\text{appr}}(x)$ (matched in a C^1-fashion at $\pm L_\pi$). If Λ_1 is not $\mathcal{O}(\sqrt{\gamma})$ close to $\frac{1}{4}\pi$, however, $\psi^2_{\text{appr}}(x)$ cannot be scaled as $\gamma\widetilde{\psi}(x)$ and the solution is not $\mathcal{O}(\gamma)$ at the point of discontinuity—in general it is $\mathcal{O}(1)$. Moreover, for any $\Lambda_1 = \mathcal{O}(1)$, there also exists on $x > 0$ a 1-parameter family of (nonnormalized) eigenfunctions for the eigenvalue problem for $\mathcal{L}^3(x;\gamma)$ that decay as $x \to \infty$. In this family there is one unique solution that connects continuously to the (normalized) solution at $x < 0$. In fact, one could define the jump in the derivative at $x = 0$, $\mathcal{J}(\lambda;\gamma)$, as an Evans function expression (note that $\mathcal{J}(\lambda;\gamma)$ can be computed explicitly at $\gamma = 0$; see [10]). By definition, λ^2 is an eigenfunction of $\mathcal{L}^3(x;\gamma)$ if and only if $\mathcal{J}(\lambda;\gamma) = 0$. In the above analysis we have shown that $\mathcal{J}(\lambda_*;\gamma) = 0$ for $\lambda_* = \frac{1}{2}\sqrt{\pi\gamma} + \mathcal{O}(\gamma)$.

Remark 3.11. The classical, driven, sine-Gordon equation, i.e., $\theta \equiv 0$ and $\gamma \neq 0$ in (3.1), has a standing pulse solution that can be seen, especially for $0 < \gamma \ll 1$, as a fluxon/antifluxon pair. This solution is approximated for $\frac{d}{dx}\phi > 0$ (the fluxon) by $\phi^1_{\text{appr}}(x)$ and for $\frac{d}{dx}\phi < 0$ (the antifluxon) by $\phi^1_{\text{appr}}(-x)$. It is (of course) unstable; the (approximation of the) critical

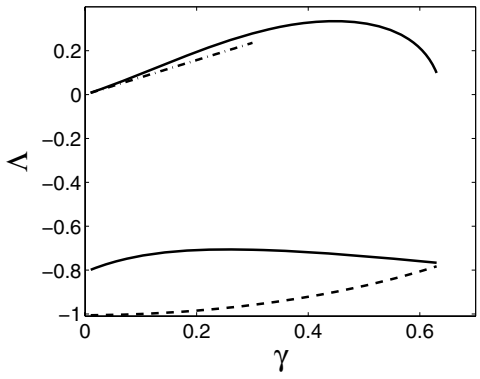


Figure 4. *The eigenvalues of the linear operator associated to the type* 3 *semifluxon as a function of the bias current* γ. *The result that the largest eigenvalue is always positive shows the instability of the type* 3 *semifluxon. When* $\gamma \ll 1$, *according to* (3.29) *the largest eigenvalue is approximated by* $\Lambda = \frac{\pi}{4}\gamma$, *shown in dash-dotted line. The dashed line is the boundary of the continuous spectrum.*

unstable eigenvalue can be obtained from (3.24). The corresponding eigenfunction is approximated by $\psi^1_{\text{appr}}(x)$ on $(-\infty, L_\pi)$, and we conclude from (3.24) that $\frac{d}{dx}\psi^1_{\text{appr}}(L_\pi) = 0$ for $\lambda^2 = \gamma\Lambda_1 = \gamma\frac{\pi}{2} + \mathcal{O}(\gamma\sqrt{\gamma})$ (while $\psi^1_{\text{appr}}(L_\pi) > 0$). Hence, for this value of Λ_1, we can match $\psi^1_{\text{appr}}(x)$ to $\psi^2_{\text{appr}}(x) = \psi^1_{\text{appr}}(-x)$ in a C^1-fashion; it gives a uniform $\mathcal{O}(\gamma)$-approximation of the critical, positive (even, "two-hump") eigenfunction of the fluxon/antifluxon pair at the eigenvalue $\lambda_+ = \frac{1}{2}\sqrt{2\pi}\sqrt{\gamma} + \mathcal{O}(\gamma) > 0$.

To consider the relation between the eigenvalues of $\mathcal{L}^3(x;\gamma)$ and the stability problem of $\phi^3_\pi(x;\gamma)$, we denote the largest eigenvalue of $\mathcal{L}^3(x;\gamma)$ by $\Lambda^3(\gamma)$. The associated eigenvalues for the linearizations are a solution of the equation $\lambda^2 - \Lambda^3(\gamma) = 0$; hence $\lambda = \pm\sqrt{\Lambda^3(\gamma)}$. Since $\Lambda^3(\gamma) > 0$, this implies that one of the two eigenvalues has positive real part; hence the fluxons of type 3 are unstable. In Figure 4, we present numerical calculations of the eigenvalues of the type 3 semifluxon as a function of the bias current γ.

Remark 3.12. A type-3 semifluxon can be seen as a concatenation of a 2π- and a $-\pi$-kink in the limit $\gamma \to 0$. In that limit the other eigenvalue of $\mathcal{L}^3(x;\gamma)$ converges to $-\frac{1}{4}(\sqrt{5}+1)$ (Figure 4) which is a contribution of the $-\pi$-kink.

4. Lattice π-kinks and their spectra in the continuum limit. In this section, we consider (2.9) for a small lattice spacing a, i.e., the driven 0-π sine-Gordon equation with a small perturbation due to lattice spacing effects. For $a = 0$, the semifluxons of all types are constructed as heteroclinic connections with transversal intersections at $x = 0$ in the two-dimensional phase space of the static equation (2.10). Therefore, all three types of semifluxons will still exist in the perturbed system with $0 < a \ll 1$; see [11]. The three types of semifluxons are denoted as $\phi^i_\pi(x;a;\gamma)$ for $i = 1$, 2, and 3. In Figure 5, we present the phase portraits of the sine-Gordon equation both with and without the effect of a perturbation due to lattice spacing.

The lattice spacing a does not affect the stationary points of the phase portraits, as can be easily checked. The existence parameters γ^* and γ_{cr} will be influenced by the lattice spacing a. For a small, they are

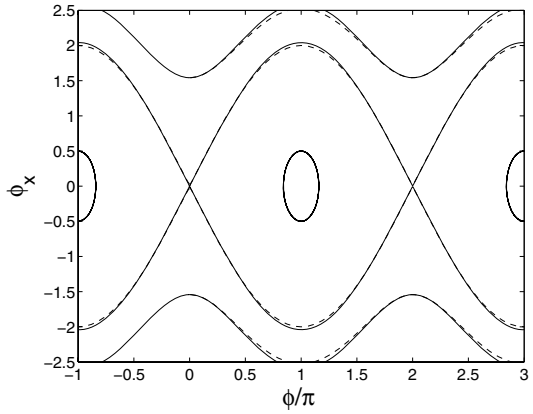


Figure 5. *The phase portrait of the stationary system (2.10) for $\gamma = 0$ and some values of the lattice spacing a. The dashed lines are the unperturbed phase portrait for $a = 0$ and the other lines correspond to $a = 0.5$.*

$$\gamma^*(a) = \frac{2}{\sqrt{4+\pi^2}} + \underbrace{\frac{2\pi}{3(\pi^2+4)^2}}_{\approx 0.0109} a^2 + \mathcal{O}(a^4), \tag{4.1}$$

$$\gamma_{\rm cr}(a) = \frac{2}{\pi} + \underbrace{\frac{\sqrt{\pi^2-4} - \pi + 2\arcsin(\frac{2}{\pi})}{3\pi^2}}_{\approx 0.0223} a^2 + \mathcal{O}(a^4); \tag{4.2}$$

see [31] for details. For $\gamma > \gamma_{\rm cr}(a)$ no static semifluxon exists.

As we have seen in the last section, for $a = 0$, the type 3 semifluxon is marginally unstable at $\gamma = \gamma_{\rm cr}$ and γ near zero. So there is a possibility that lattice spacing effects stabilize the type 3 semifluxon near those values of γ. However, it turns out that this is not the case and the stability of the semifluxons is similar to the case $a = 0$.

Theorem 4.1. *For a small, the linearizations about the π-kinks have the following properties:*

(i) *The eigenvalues of the linearization about the monotonic type 1 π-kink $\phi^1_\pi(x;a;\gamma)$ are strictly negative for $0 \le \gamma < \gamma_{\rm cr}(a)$. At $\gamma = \gamma_{\rm cr}(a)$, the largest eigenvalue is zero. These π-kinks are linearly stable.*

(ii) *The largest eigenvalue of the linearization about the monotonic type 2 π-kink $\phi^2_\pi(x;a;\gamma)$ is strictly positive for $0 < \gamma < \gamma^*(a)$. These π-kinks are linearly unstable.*

(iii) *The largest eigenvalue of the linearization about the monotonic type 3 π-kink $\phi^3_\pi(x;a;\gamma)$ is strictly positive for $0 < \gamma < \gamma_{\rm cr}(a)$. These π-kinks are linearly unstable. In the limit for $\gamma \to 0$ and $\gamma \to \gamma_{\rm cr}(a)$, the largest eigenvalue converges to zero.*

The proof of this theorem will proceed along similar lines as the proof in the previous section. First we consider the eigenvalue problem of a solution $\phi^i_\pi(x;a;\gamma)$, which can be written as

$$\mathcal{L}^i(x;a;\gamma)\, v = \lambda^2\, v,$$

where $\mathcal{L}^i(x;a;\gamma)$ is now defined as the linearization associated with (2.10), i.e.,

$$\begin{aligned}\mathcal{L}^i(x;a;\gamma) &= D_{xx} - \cos\left(\phi_\pi^i(x;a;\gamma)+\theta(x)\right)\\ &\quad - \tfrac{a^2}{12}\Big[2\cos\widetilde{\phi}\,D_{xx} - 2(\phi_\pi^i(x;0;\gamma))_x \sin\widetilde{\phi}\,D_x\\ &\quad - 1 + 2\gamma\sin\widetilde{\phi} - ((\phi_\pi^i(x;0;\gamma))_x)^2\cos\widetilde{\phi}\Big] + \mathcal{O}(a^4),\end{aligned}$$

where $\widetilde{\phi} = \phi_\pi^i(x;0;\gamma)+\theta(x)$.

Lemma 3.3 can be extended to $a \neq 0$ and give a necessary and sufficient condition for $\mathcal{L}^i(x;a;\gamma)$ to have an eigenvalue $\Lambda = 0$.

Lemma 4.2. *The eigenvalue problem*

$$\mathcal{L}^i(x;\gamma)v = \Lambda v, \quad x \in \mathbb{R},$$

has an eigenvalue $\Lambda = 0$ if and only if one of the following two conditions holds:

(i) *$D_{xx}\phi_\pi^i(x;a;\gamma)$ is continuous at $x = 0$; i.e., $\phi_\pi^i(0;a;\gamma) = k\pi - a^2\frac{\gamma}{12} + \mathcal{O}(a^4)$ for some $k \in \mathbb{Z}$;*

(ii) *$D_x\phi_\pi^i(0;a;\gamma) = 0$ and there are some $x_\pm$, with $\operatorname{sgn}(x_\pm) = \pm 1$, such that $D_x\phi_\pi^i(x_\pm;a;\gamma) \neq 0$.*

Proof. As the proof of Lemma 3.3 is based on the fact that the derivative of the semifluxon is a solution of the linearized system for $x \neq 0$, we can follow the same arguments to prove this lemma. Again this leads to two conditions: that either ϕ_{xx}^i is continuous at $x = 0$ or the second condition as stated above.

In order to determine when ϕ_{xx}^i is continuous, we use the static equation (2.10) and expand near $a = 0$,

$$\begin{aligned}D_{xx}\phi_\pi^i(x;a;\gamma) &= \left(\sin(\phi_\pi^i(x;a;\gamma)+\theta(x)) - \gamma\right)\left(1 - \tfrac{a^2}{12}\cos\widetilde{\phi}\right)\\ &\quad + \tfrac{a^2}{6}\sin\widetilde{\phi}\left(\gamma\arcsin\gamma + \sqrt{1-\gamma^2} - \gamma\widetilde{\phi} - \cos\widetilde{\phi}\right) + \mathcal{O}(a^4),\end{aligned}$$

again with $\widetilde{\phi} = \phi_\pi^i(x;0;\gamma)+\theta(x)$. The continuity of $D_{xx}\phi_\pi^i$ at $x = 0$ leads to the expression for $\phi_\pi^i(x;a;\gamma)$ as given above. ∎

At $\gamma = \gamma_{\rm cr}(a)$, the stable manifold of the $\pi + \arcsin\gamma$ and the homoclinic connection at $\arcsin\gamma$ are tangent, implying that $D_{xx}\phi_\pi^i(x;a;\gamma)$ is continuous at $x = 0$. Thus the first condition of the lemma is satisfied at $\gamma = \gamma_{\rm cr}(a)$ for $i = 1, 3$. For the same reasons as before, the second condition is never satisfied.

Since $\Lambda = 0$ is an eigenvalue of the linearized operator $\mathcal{L}^i(x;a;\gamma)$ if and only if $\gamma = \gamma_{\rm cr}(a)$, the sign of the eigenvalues of $\mathcal{L}^i(x;a;\gamma)$ will not change. Thus the behavior of the eigenvalues near $\gamma = 0$ will again determine the stability of the semifluxons.

For $\gamma = 0$ and $\theta = 0$, the sine-Gordon equation with a perturbation due to the lattice spacing has a heteroclinic orbit connecting 0 and 2π. As before, the heteroclinic orbit will play an important role in determining the stability of the semifluxons for small values of γ. For small values of the lattice spacing a, we can approximate this heteroclinic orbit up to order a^2 by using the 2π-fluxon $\phi_{\rm flux}$ and its linearization.

Lemma 4.3. *Let $\phi^a_{\text{flux}}(x)$ denote the heteroclinic orbit of the sine-Gordon equation with a perturbation due to the lattice spacing (i.e., (2.9) with $\theta \equiv 0$ and $\gamma = 0$). For the lattice spacing a small, we have for the symmetric (i.e., $\phi^a_{\text{flux}}(0) = \pi$) heteroclinic connection $\phi^a_{\text{flux}}(x)$*

$$\phi^a_{\text{flux}}(x) = \phi_{\text{flux}}(x) + a^2\phi_a(x) + \mathcal{O}(a^4), \tag{4.3}$$

where

$$\phi_a(x) = -\frac{1}{12}\frac{-3\sinh x + x\cosh x}{\cosh^2 x}. \tag{4.4}$$

This approximation is valid, uniform in $x \in \mathbb{R}$.

Proof. The spatially localized correction to the kink shape $\phi_{\text{flux}}(x)$ due to the perturbation term representing lattice spacing is sought in the form of perturbation series:

$$\phi^a_{\text{flux}}(x) = \phi_{\text{flux}}(x) + a^2\phi_a(x) + \mathcal{O}(a^4).$$

It is a direct consequence that $\phi_a(x)$ satisfies

$$\begin{aligned}\mathcal{L}^1(x;0)\phi_a(x) = f(x) = -\tfrac{1}{12}\big[&2\cos\phi_{\text{flux}}(x)\partial_{xx}\phi_{\text{flux}}(x)\\ &-\sin\phi_{\text{flux}}(x)(\partial_x\phi_{\text{flux}}(x))^2 - \cos\phi_{\text{flux}}(x)\sin\phi_{\text{flux}}(x)\big],\end{aligned} \tag{4.5}$$

where $\mathcal{L}^1(x;0)$ is the linearized operator associated to the fluxon, i.e., $\mathcal{L}^1(x;0) = D_{xx} - \cos\phi_{\text{flux}}(x)$.

Using the variation-of-constants method, we obtain the general solution of (4.5), i.e.,

$$\phi_a(x) = A(x)\operatorname{sech} x + B(x)\,(x\operatorname{sech} x + \sinh x), \tag{4.6}$$

where

$$\begin{aligned}A(x) &= A_0 + \frac{1}{24}\left[2\ln\left(\frac{1-\cosh x-\sinh x}{\cosh x - 1 - \sinh x}\right) + \frac{6\sinh x}{\cosh x} - \frac{4\sinh x}{\cosh^3 x} + \int_0^x \frac{\xi f(\xi)}{\cosh\xi}d\xi\right],\\ B(x) &= B_0 - \frac{1}{24}\left[2 + \frac{1}{\cosh^2 x} - \frac{3}{\cosh^4 x}\right].\end{aligned}$$

The integration constant B_0 is determined by the condition that $\phi_a(x)$ is bounded, leading to $B_0 = \frac{1}{12}$. The integration constant A_0 is determined by the requirement that $\phi^a_{\text{flux}}(0) = \pi$; hence $\phi_a(0) = 0$, giving that $A_0 = 0$. ∎

For $\gamma = 0$, the static model (2.10) for a 0-π Josephson junction with lattice spacing effects has both a π- and a 3π-kink solution. The 2π-heteroclinic orbit found above can be used to derive approximations for those kinks.

Lemma 4.4. *For a small and $\gamma = 0$, we have an explicit expression for the π- and 3π-fluxon up to order $\mathcal{O}(a^2)$, respectively:*

$$\begin{aligned}\phi^1_\pi(x;a;0) &= \phi^1_\pi(x;0) + a^2\begin{cases} -u^1_\pi(x - \ln(1+\sqrt{2})) & \text{for} \quad x < 0,\\ u^1_\pi(-x - \ln(1+\sqrt{2})) & \text{for} \quad x > 0,\end{cases}\\ \phi^2_{3\pi}(x;a;0) &= \phi^2_{3\pi}(x;0) + a^2\begin{cases} -u^1_{3\pi}(x + \ln(1+\sqrt{2})) & \text{for} \quad x < 0,\\ u^1_{3\pi}(-x + \ln(1+\sqrt{2})) & \text{for} \quad x > 0,\end{cases}\end{aligned} \tag{4.7}$$

where $\phi^1_\pi(x;0)$ *and* $\phi^2_{3\pi}(x;0)$ *are the* π*-, respectively, the* 3π*-fluxons as defined in* (3.11) *and*

$$u^1_\pi(x) = \tfrac{1}{12\cosh x}\left(\tfrac{3\sqrt{2}}{2} - \tfrac{1}{2}\ln(3-\sqrt{2}) + 3\tanh x - x\right),$$
$$u^1_{3\pi}(x) = \tfrac{1}{12\cosh x}\left(-\tfrac{3\sqrt{2}}{2} + \tfrac{1}{2}\ln(3-\sqrt{2}) + 3\tanh x - x\right).$$

4.1. Stability of type 1 semifluxon. We will show that the type 1 wave $\phi^1_\pi(x;a;\gamma)$ is linearly stable for small a and $0 \le \gamma \le \gamma_{\rm cr}$ by analyzing the largest eigenvalue of $\mathcal{L}^1(x;a;\gamma)$ for $0 \le \gamma \le \gamma_{\rm cr}(a)$.

Lemma 4.5. *For the lattice spacing parameter* a *sufficiently small and* $0 \le \gamma < \gamma_{\rm cr}(a)$*, the largest eigenvalue of* $\mathcal{L}^1(x;a;\gamma)$ *is strictly negative. For* $\gamma = \gamma_{\rm cr}(a)$*, the operator* $\mathcal{L}^1(x;a;\gamma_{\rm cr}(a))$ *has* 0 *as its largest eigenvalue. For* $\gamma = 0$*, the largest eigenvalue decreases as* a *increases and is proportional to* $-\frac{1}{4}(\sqrt{5}+1) - 0.0652a^2 + \mathcal{O}(a^4)$.

Proof. First we look at the stability of the π-kink at $\gamma = 0$. Writing $v(x) = v^0(x) + a^2 v^1(x) + \mathcal{O}(a^4)$ and $\Lambda = \Lambda_0 + a^2\Lambda_1 + \mathcal{O}(a^4)$ and expanding the eigenvalue problem for the stability of the π-kink $\phi^1_\pi(x;a;0)$ in a Taylor series result in the equations

$$\begin{aligned} \left(\mathcal{L}^1(x;0;0) - \Lambda_0\right) v^0(x) &= 0, \\ \left(\mathcal{L}^1(x;0;0) - \Lambda_0\right) v^1(x) &= \left(\Lambda_1 - u^1_\pi(x)\sin(\phi^1_\pi(x;0)+\theta)\right) v^0(x) - g(x), \end{aligned} \tag{4.8}$$

where $\mu = \sqrt{\Lambda_0 + 1}$, $\Lambda_0 = -\frac{1}{4}(\sqrt{5}+1)$,

$$v^0(x) = \begin{cases} e^{\mu(x-\ln(1+\sqrt{2}))}\left[\tanh(x - \ln(1+\sqrt{2})) - \mu\right] & \text{for} \quad x < 0, \\ e^{\mu(-x-\ln(1+\sqrt{2}))}\left[\tanh(-x - \ln(1+\sqrt{2})) - \mu\right] & \text{for} \quad x > 0, \end{cases}$$

$$\begin{aligned} g(x) = \tfrac{1}{12}\Big[&2v^0_{xx}\Lambda_0 + v^0 + 2v^0_{xx}\cos\widetilde{\phi}(x) - 2\cos^2\widetilde{\phi}(x)v^0 - 2\partial_{xx}(\phi^1_\pi(x;0))\sin\widetilde{\phi}(x)v^0 \\ &- 2\partial_x\phi^1_\pi(x;0)\sin\widetilde{\phi}(x)v^0_x - (\partial_x\phi^1_\pi(x;0))^2\cos\widetilde{\phi}(x)v^0 - v^0\Lambda_0^2 - 2v^0\Lambda_0\cos\widetilde{\phi}(x)\Big], \end{aligned}$$

with again $\widetilde{\phi}(x) = \phi^1_\pi(x;0) + \theta(x)$ (see Lemma 3.5).

The parameter value of Λ_1 is calculated by solving (4.8) for a bounded and decaying solution $v^1(x)$. The general solution can be derived by using the variation-of-constants method because we have the homogeneous solutions of the equation. One can also use the Fredholm theorem (see, e.g., [30]); i.e., the sufficient and necessary condition for (4.8) to have a solution $v^1 \in H_2(\mathbb{R})$ is that the inhomogeneity is perpendicular to the null space of the self-adjoint operator of $\mathcal{L}^1(x;0;0)$. If $\langle\ ,\ \rangle$ denotes an inner product in $H_2(\mathbb{R})$, then this condition gives

$$0 = \langle(\mathcal{L}^1(x;0;0) - \Lambda_0)v^1, v^0\rangle = \langle\Lambda_1 v^0 - u^1_\pi v^0\sin(\phi^1_\pi(x;0)+\theta) - g, v^0\rangle,$$

which implies that

$$\Lambda_1 = \frac{3584(70\sqrt{2}(1+\sqrt{5}) - 99(1+\sqrt{5}))}{24576(-70\sqrt{10} - 350\sqrt{2} + 495 + 99\sqrt{5})} \approx -0.0652. \tag{4.9}$$

Now assume that the operator $\mathcal{L}^1(x;\gamma)$ has a positive eigenvalue $\Lambda^1(\gamma)$ for some $0 \le \gamma < \gamma_{\rm cr}(a)$. Since Λ depends continuously on γ, there has to be some $0 < \widehat{\gamma} < \gamma_{\rm cr}(a)$ such that $\Lambda^1(\widehat{\gamma}) = 0$. However, from Lemma 4.2 it follows that this is not possible. ∎

4.2. Instability of type 2 semifluxon. In Lemma 3.6 we have seen that for $a = 0$, the linearization about the type 2 semifluxon has a strictly positive largest eigenvalue. Also the limits of this eigenvalue for $\gamma \to 0$ and $\gamma \to \gamma^*$ are still strictly positive. Thus a small perturbation associated with the lattice spacing cannot stabilize the type 2 semifluxons.

For completeness, we will consider the case $\gamma = 0$. In this limit, the type 2 semifluxon can be seen as a concatenation of a 3π-kink and a -2π-kink. As before, the limit of the largest eigenvalue for $\gamma \to 0$ will be equal to the largest eigenvalue of the 3π-kink. We have seen that the largest eigenvalue of the 3π-kink at $\gamma = 0$ and $a = 0$ is strictly positive and the following lemma shows that small lattice spacing effects increase this eigenvalue.

Lemma 4.6. *For the lattice spacing parameter a sufficiently small, the largest eigenvalue of the linearization $\mathcal{L}^2(x;a;0)$ about the 3π-kink $\phi^2_{3\pi}(x;a;0)$ is strictly positive. Moreover, it increases as a increases and is proportional to $\frac{1}{4}(\sqrt{5}-1) + 0.0652a^2 + \mathcal{O}(a^4)$.*

Proof. Note that the lowest order analytic expressions for the π- and the 3π-kinks differ only in the sign of the "kink-shift" (see (4.7)). Because of this, we can follow the same steps as the proof of Lemma 4.5. Writing the largest eigenvalue of a 3π-kink as $\Lambda = \Lambda_0 + a^2\Lambda_1 + \mathcal{O}(a^4)$, with $\Lambda_0 = (\sqrt{5}-1)/4$ as has been calculated in Lemma 3.6, we compute Λ_1 to be

$$\Lambda_1 = \frac{3584(665857(\sqrt{5}-1) - 470832\sqrt{2}(\sqrt{5}+1))}{24576(3329285 - 2354160\sqrt{2} - 665857\sqrt{5} + 470832\sqrt{10})} \approx 0.0652. \tag{4.10}$$

■

Thus up to order $\mathcal{O}(a^4)$ the lattice spacing effects destabilize the 3π-kink.

Because a 2π-fluxon in the "ordinary" sine-Gordon equation can be pinned by lattice spacing effects, one might expect to have a stable 3π-kink in the 0-π sine-Gordon equation with larger lattice spacing effects. This is confirmed by numerical calculations in section 6; see Figure 12. If the 3π-kink is stable for $\gamma = 0$, a stable type 2 semikink might exist for $\gamma > 0$ when the repelling force between the 3π-kink and the antifluxon is smaller than the energy to move a fluxon along lattices. However, in section 6 it will be shown numerically that the type 2 semikink is unstable for all values of the lattice spacing; see Figure 13(b).

4.3. Instability of type 3 semifluxon. For γ small or close to $\gamma_{\rm cr}$, it has been shown in Lemma 3.9 that the type 3 semifluxons are weakly unstable. This opens the possibility that the perturbation term representing the lattice spacing stabilizes the semifluxon. This is not the case, however.

Lemma 4.7. *For small lattice spacing a and bias current $0 < \gamma < \gamma_{\rm cr}(a)$, the largest eigenvalue of the linearization $\mathcal{L}^3(x;a;\gamma)$ about the type 3 semifluxon $\phi^3_\pi(x;a;\gamma)$ is strictly positive. For $\gamma = \gamma_{\rm cr}(a)$, the operator $\mathcal{L}^3(x;a;\gamma_{\rm cr})$ has 0 as its largest eigenvalue. For γ near zero and $a^2 = \gamma\hat{a}^2$, the largest eigenvalue of $\mathcal{L}^3(x;a;\gamma)$ is $\Lambda_* = \left(\frac{\pi}{4} + \frac{7}{180}\hat{a}^2\right)\gamma + \mathcal{O}(\gamma\sqrt{\gamma})$.*

Proof. At $\gamma = \gamma_{\rm cr}$, the solution $\phi^3_\pi(x;a;\gamma_{\rm cr}(a)) = \phi^1_\pi(x;a;\gamma_{\rm cr}(a))$. Hence from Lemma 4.5 it follows that the largest eigenvalue of the linearization about $\phi^3_\pi(x;a;\gamma_{\rm cr}(a))$ vanishes.

From Lemma 3.9 it follows that the largest eigenvalue of the linearization about $\phi^3_\pi(x;a;\gamma)$ is positive for $a = 0$ and $0 < \gamma < \gamma_{\rm cr}$. Thus a small perturbation cannot change the positive sign of the largest eigenvalue if γ is not near 0 or $\gamma_{\rm cr}$. Now assume that a small perturbation would lead to a negative largest eigenvalue near $\gamma = 0$ or $\gamma = \gamma_{\rm cr}$. Then there has to be a zero eigenvalue near $\gamma = 0$ or $\gamma = \gamma_{\rm cr}$, but this is not possible according to Lemma 4.5. Thus we can conclude that the largest eigenvalue is always positive.

To complete the proof, we will derive the asymptotic expression of the eigenvalue near $\gamma = 0$. Since both a and γ are small, we relate those two parameters by writing $a^2 = \gamma\hat{a}^2$. Now the approximation for the type 3 semifluxon can be written as

$$\phi_\pi^3(x;\hat{a}\sqrt{\gamma};\gamma) = \begin{cases} \phi_{\text{flux}}(\widehat{x}) + \gamma\phi_1(\widehat{x}) + \gamma\hat{a}^2\phi_a(\widehat{x}) + \gamma^2 R_2(\widehat{x};\gamma), & x < -L_\pi(\gamma) + x_1, \\ \phi_{\text{flux}}(-\widetilde{x}) + \gamma\phi_1(-\widetilde{x}) + \gamma\hat{a}^2\phi_a(-\widetilde{x}) + \gamma^2 R_2(-\widetilde{x};\gamma), & -L_\pi(\gamma) + x_1 < x < 0, \\ \pi + \phi_{\text{flux}}(-x - x_1) + \mathcal{O}(\gamma), & x > 0, \end{cases}$$

where $\widehat{x} = x - x_1 + 2L_\pi(\gamma)$ and $\widetilde{x} = x - x_1$. It can be shown that the shift $L_\pi(\gamma)$ does not depend on $\hat{a}^2$ in lowest order; i.e., $L_\pi(\gamma) = \frac{1}{2}|\ln\gamma| + \ln\frac{4}{\sqrt{\pi}} + \mathcal{O}(\sqrt{\gamma})$.

To find the largest eigenvalue, we set again $\Lambda^3(\gamma) = \gamma\Lambda_1(0)$ and follow the steps in the proof of Lemma 3.9 with some additional terms added to some expressions.

First, we consider the part of the approximation with $x < -L_\pi(\gamma) + x_1$ or $\widehat{x} < L_\pi(\gamma)$. As before, we drop the hat in $\widehat{x}$ in this part of the argument. On $(-\infty, L_\pi)$, the general solution of the eigenvalue problem of the order $\mathcal{O}(\gamma)$ after expanding $\psi^1_{\text{approx}} = \psi_0 + \gamma\psi_1$ is

$$\begin{aligned} \psi_1(x) = &\left[\frac{\pi}{4} - \frac{1}{2}\Lambda_1\left(\ln\cosh x + \int_0^x \frac{\xi}{\cosh^2\xi}d\xi\right)\right]\frac{1}{\cosh x} \\ &+ \left[\frac{1}{2}\Lambda_1(0) + \frac{1}{2}\Lambda_1\tanh x\right]\left(\frac{x}{\cosh x} + \sinh x\right) + \frac{1}{2}\left(\frac{d}{dx}\phi_1 + \hat{a}^2\frac{d}{dx}\phi_a\right) \\ &- \frac{e^x}{360(e^{2x}+1)^3}\left[16\ln 2 + e^{2x}(32\ln 2 - 295 + 60x) + 30x + 137 + 7e^{6x}\right. \\ &\qquad \left. - 16\ln(e^{2x}+1)(e^{2x}+1)^2 + e^{4x}(151 + 30x + 16\ln 2)\right]. \end{aligned}$$

We note that the error term $|\psi(x) - \psi^1_{\text{appr}}(x)| = \gamma^2|S_2(x;\gamma)|$ is still at most $\mathcal{O}(\gamma)$ on $(-\infty, L_\pi)$.

Next consider the second part of the approximation, i.e., x between $-L_\pi(\gamma) + x_1$ and 0 or $\widetilde{x} < -L_\pi(\gamma)$. Again, we drop the tilde in $\widetilde{x}$. We scale $\psi(x)$ as $\gamma\widetilde{\psi}(x)$. The linearization $\widetilde{\psi}(x)$ along $\phi^2_{\text{appr}}(x)$ on the interval $(-L_\pi, -x_1)$ must solve $\mathcal{L}\widetilde{\psi} = O(\gamma)$. Thus, at leading order

$$\widetilde{\psi}(x) = \frac{\tilde{A}}{\cosh x} + \tilde{B}\left(\frac{x}{\cosh x} + \sinh x\right).$$

The last part of the approximation of $\psi(x)$ on $x > 0$, $\psi^3_{\text{appr}}(x)$, is obtained by linearizing along $\phi^3_{\text{appr}}(x)$ and by translating x so that $x \in (x_1, \infty)$. We also scale $\psi^3_{\text{appr}}(x) = \gamma\hat{\psi}(x)$. As $\hat{\psi}$ must be bounded for $x \to \infty$, it follows that $\hat{\psi}(x) = \hat{A}/\cosh x + \mathcal{O}(\gamma)$ for some $\hat{A} \in \mathbb{R}$.

Finally, we have to connect all parts of the eigenfunction in a C^1-fashion. This determines the values of $\Lambda_1(0)$, $\tilde{A}$, $\tilde{B}$, and $\hat{A}$ as

$$\Lambda_1(0) = \frac{1}{4}\pi + \frac{7}{180}\hat{a}^2, \quad \tilde{B} = -\Lambda_1(0), \quad \tilde{A} = \frac{1}{4}\pi[\sqrt{2} + \log(\sqrt{2}-1)], \quad \text{and} \quad \hat{A} = \frac{1}{2}\pi\sqrt{2};$$

thus $\Lambda_1(0) > 0$, $\tilde{B} < 0$, $\tilde{A} > 0$, and $\hat{A} > 0$. And we can conclude that the eigenvalue problem for the π-fluxon $\phi^3_\pi(x;\hat{a}\sqrt{\gamma};\gamma)$ has a positive largest eigenvalue

$$\Lambda_* = \left(\frac{\pi}{4} + \frac{7}{180}\hat{a}^2\right)\gamma + \mathcal{O}(\gamma\sqrt{\gamma}). \qquad \blacksquare$$

5. Semikinks in the weak-coupling limit. In this section we will consider the discrete 0-π sine-Gordon equation (2.3) when the lattice parameter a is large. The time independent version of (2.3) is well known: when $\gamma = 0$, it corresponds to the so-called standard or Taylor–Greene–Chirikov map [9], and when $\gamma \neq 0$, it is called the Josephson map [24]. Since we are interested in the case in which the lattice spacing a is large, we introduce the coupling parameter ε as $\varepsilon = \frac{1}{a^2}$ and the equation becomes

$$\ddot{\phi}_n - \varepsilon\left[\phi_{n-1} - 2\phi_n + \phi_{n+1}\right] = -\sin(\phi_n + \theta_n) + \gamma. \tag{5.1}$$

When there is no coupling, i.e., $\varepsilon = 0$, it can be seen immediately that there are infinitely many steady state solutions:

$$\phi_n = \begin{cases} \cos(k_n\pi)\arcsin\gamma + k_n\pi, & n = 0, -1, -2, \ldots, \\ \cos(k_n\pi)\arcsin\gamma + (k_n+1)\pi, & n = 1, 2, 3, \ldots, \end{cases}$$

where k_n is an integer. The only monotone semikink is the solution with $k_n = 0$ for $n \in \mathbb{Z}$; thus it is natural to identify this semikink with the type 1 semikink. However, it is less clear which solution would correspond to the type 2 and type 3 semikinks. Possible candidates for the type 2 wave are solutions for which there is some $N \in \mathbb{N}$ such that $k_n = 0$ for $n \leq 0$ and $n \geq N$ and $k_n = 1$ for $0 < n < N$. Similarly, candidates for the type 3 wave are solutions for which there is some $N \in \mathbb{N}$ such that $k_n = 0$ for $n \leq -N$ and $n \geq 0$ and $k_n = 1$ for $-N < n < 0$. But there are many other candidates involving combinations of $k_n = 0$ or $k_n = 1$ as well. If one starts with such a wave in the uncoupled limit, i.e., with $\varepsilon \ll 1$ or $a \to \infty$, and uses continuation to follow this wave in the discrete system (5.1) toward $a = 0$ or $\varepsilon \to \infty$, then it turns out that most waves end in a saddle-node bifurcation [3]. More details about the continuation can be found in section 6.

In this section we will focus on the analytical study of the type 1 semikink for the coupling parameter ε small (thus the lattice spacing a large). We will denote this wave by $\Phi^1_\pi(n;\varepsilon;\gamma)$, and for $\varepsilon = 0$, we have

$$\Phi^1_\pi(n;0;\gamma) = \begin{cases} \arcsin\gamma, & n = 0, -1, -2, \ldots, \\ \pi + \arcsin\gamma, & n = 1, 2, 3, \ldots. \end{cases}$$

The existence of the continuation of (5.2) for small coupling ε is guaranteed by the following lemma.

Lemma 5.1. *The steady state solution $\Phi^1_\pi(n;0;\gamma)$, representing the semifluxon of type 1 in the uncoupled limit $\varepsilon = 0$, can be continued for ε small and $\gamma < 1$. It is given by*

$$\Phi^1_\pi(n;\varepsilon;\gamma) = \begin{cases} \arcsin\gamma + \mathcal{O}(\varepsilon^2), & n \leq -1; \\ \arcsin\gamma + \varepsilon\frac{\pi}{\sqrt{1-\gamma^2}} + \mathcal{O}(\varepsilon^2), & n = 0; \\ \pi + \arcsin\gamma - \varepsilon\frac{\pi}{\sqrt{1-\gamma^2}} + \mathcal{O}(\varepsilon^2), & n = 1; \\ \pi + \arcsin\gamma + \mathcal{O}(\varepsilon^2), & n \geq 2. \end{cases} \tag{5.2}$$

For γ close to one, we write $\gamma = 1 - \varepsilon\widetilde{\gamma}$. If $\widetilde{\gamma} > \pi$, then the type 1 solution is

$$\Phi^1_\pi(n;\varepsilon;1-\varepsilon\widetilde{\gamma}) = \begin{cases} \frac{\pi}{2} - \sqrt{\varepsilon}\sqrt{2\widetilde{\gamma}} + \mathcal{O}(\varepsilon), & n \le -1; \\ \frac{\pi}{2} - \sqrt{\varepsilon}\sqrt{2(\widetilde{\gamma}-\pi)} + \mathcal{O}(\varepsilon), & n = 0; \\ \frac{3\pi}{2} - \sqrt{\varepsilon}\sqrt{2(\widetilde{\gamma}+\pi)} + \mathcal{O}(\varepsilon), & n = 1; \\ \frac{3\pi}{2} - \sqrt{\varepsilon}\sqrt{2\widetilde{\gamma}} + \mathcal{O}(\varepsilon), & n \ge 2. \end{cases} \tag{5.3}$$

From (5.3) *we obtain the critical bias current for the existence of static semifluxon as*

$$\gamma_{\rm cr} = 1 - \varepsilon\pi + \mathcal{O}(\varepsilon^2). \tag{5.4}$$

Proof. The existence proof for $\gamma < 1$ follows from the implicit function theorem as given in [22, Theorem 2.1] or [25, Lemma 2.2].

For the case $\gamma = 1-\varepsilon\widetilde{\gamma}$, the implicit function theorem as presented in the references above cannot be applied immediately. However, after some manipulations, the implicit function theorem can be applied again. First we substitute into the steady state equation $\gamma = 1-\varepsilon\widetilde{\gamma}$ and $\Phi = \Phi_0 + \sqrt{\varepsilon}\widetilde{\Phi}$, where $\Phi_0(n) = \frac{\pi}{2}$ for $n \le 0$ and $\Phi_0(n) = \frac{3\pi}{2}$ for $n \ge 1$. This gives the following equations:

$$\begin{aligned} 0 &= \tfrac{\cos(\sqrt{\varepsilon}\widetilde{\Phi}(n))-1}{\varepsilon} + \widetilde{\gamma} - \sqrt{\varepsilon}[\widetilde{\Phi}(n-1) - 2\widetilde{\Phi}(n) + \widetilde{\Phi}(n+1)] =: \widetilde{F}_n(\widetilde{\Phi},\varepsilon), \quad n \ne 0,1, \\ 0 &= \tfrac{\cos(\sqrt{\varepsilon}\widetilde{\Phi}(0))-1}{\varepsilon} + \widetilde{\gamma} - \sqrt{\varepsilon}[\widetilde{\Phi}(-1) - 2\widetilde{\Phi}(0) + \widetilde{\Phi}(1)] - \pi =: \widetilde{F}_0(\widetilde{\Phi},\varepsilon), \quad n = 0, \\ 0 &= \tfrac{\cos(\sqrt{\varepsilon}\widetilde{\Phi}(1))-1}{\varepsilon} + \widetilde{\gamma} - \sqrt{\varepsilon}[\widetilde{\Phi}(0) - 2\widetilde{\Phi}(1) + \widetilde{\Phi}(2)] + \pi =: \widetilde{F}_1(\widetilde{\Phi},\varepsilon), \quad n = 1. \end{aligned}$$

Using that $\lim_{\varepsilon\to 0} \frac{\cos(\sqrt{\varepsilon}\widetilde{\Phi}(n))-1}{\varepsilon} = -\frac{1}{2}(\widetilde{\Phi}(n))^2$, the definitions for $\widetilde{F}$ can be smoothly extended to $\varepsilon = 0$, too. The equations for $\varepsilon = 0$ become

$$\widetilde{\Phi}^2(n) = 2\widetilde{\gamma},\ n \ne 0,1; \quad \widetilde{\Phi}^2(0) = 2(\widetilde{\gamma}-\pi); \quad \text{and} \quad \widetilde{\Phi}^2(1) = 2(\widetilde{\gamma}+\pi).$$

For $|n|$ large, the wave should be asymptotic to the center point of the temporal dynamics; hence $\widetilde{\Phi}(n) = -\sqrt{2\widetilde{\gamma}}$ for $|n|$ large. So for $\widetilde{\gamma} \ge \pi$, there are two monotone semikinks (recall that the full semikink is given by $\Phi_0 + \sqrt{\varepsilon}\widetilde{\Phi}$):

$$\widetilde{\Phi}^{\pm}(n;0;\widetilde{\gamma}) = \begin{cases} -\sqrt{2\widetilde{\gamma}}, & n \le -1; \\ \pm\sqrt{2(\widetilde{\gamma}-\pi)}, & n = 0; \\ -\sqrt{2(\widetilde{\gamma}+\pi)}, & n = 1; \\ -\sqrt{2\widetilde{\gamma}}, & n \ge 2. \end{cases}$$

Note that the $\pm$-solutions collide for $\widetilde{\gamma} = \pi$. The linearization $D\widetilde{F}(\widetilde{\Phi}^{\pm},0)$ is invertible for $\widetilde{\gamma} > \pi$; hence the implicit function theorem can be applied again and we have the existence of monotone semikinks $\Phi_0(n) + \sqrt{\varepsilon}\widetilde{\Phi}^{\pm}(n,\varepsilon,\widetilde{\gamma})$. In analogue with the continuum case, the type 1 wave is the one that has the discontinuity at the lowest value of the phase. The critical bias current for the existence of a static lattice semifluxon follows immediately from the arguments above. ∎

The two $\pm$-solutions near $\gamma_{\rm cr}$ as derived above in the proof are like the type 1 and type 3 semifluxons near $\gamma_{\rm cr}$ in the PDEs studied in the previous two sections. So in analogue to those PDEs, we can define for $\widetilde{\gamma} > \pi$

$$\Phi^1_\pi(n;\varepsilon;1-\varepsilon\widetilde{\gamma}) = \Phi_0(n) + \sqrt{\varepsilon}\widetilde{\Phi}^-(n;\varepsilon;\widetilde{\gamma}) \quad \text{and} \quad \Phi^3_\pi(n;\varepsilon;1-\varepsilon\widetilde{\gamma}) = \Phi_0(n) + \sqrt{\varepsilon}\widetilde{\Phi}^+(n;\varepsilon;\widetilde{\gamma}),$$

where Φ_0 and $\widetilde{\Phi}^\pm$ are as in the proof above. Thus we get

$$\Phi^{1/3}_\pi(n;\varepsilon;1-\varepsilon\widetilde{\gamma}) = \begin{cases} \frac{\pi}{2} - \sqrt{\varepsilon}\sqrt{2\widetilde{\gamma}} + \mathcal{O}(\varepsilon), & n \geq -1; \\ \frac{\pi}{2} \mp \sqrt{\varepsilon}\sqrt{2(\widetilde{\gamma}-\pi)} + \mathcal{O}(\varepsilon), & n = 0; \\ \frac{3\pi}{2} - \sqrt{\varepsilon}\sqrt{2(\widetilde{\gamma}+\pi)} + \mathcal{O}(\varepsilon), & n = 1; \\ \frac{3\pi}{2} - \sqrt{\varepsilon}\sqrt{2\widetilde{\gamma}} + \mathcal{O}(\varepsilon), & n \geq 2. \end{cases} \tag{5.5}$$

The spectral stability of $\Phi^i_\pi(n;\varepsilon;\gamma)$ is obtained by substituting $\phi_n = \Phi^i_\pi(n;\varepsilon;\gamma) + v_n e^{\lambda t}$ in the model equation (2.3). Disregarding the higher order terms in v_n gives the eigenvalue problem

$$L^i(\varepsilon;\gamma)\nu = \Lambda\nu, \tag{5.6}$$

where $\Lambda = \lambda^2$, $\nu = (\dots, v_{-1}, v_0, v_1, \dots)^T$, and $L^i(\varepsilon;\gamma)$ is the linear discrete operator

$$L^i(\varepsilon;\gamma) = \begin{pmatrix} \ddots & \ddots & \ddots & & & & 0 \\ & \varepsilon & -2\varepsilon - A_{-1} & \varepsilon & & & \\ & & \varepsilon & -2\varepsilon - A_0 & \varepsilon & & \\ & & & \varepsilon & -2\varepsilon - A_1 & \varepsilon & \\ 0 & & & & \ddots & \ddots & \ddots \end{pmatrix},$$

$$A_n = \cos\left(\Phi^i_\pi(n;\varepsilon;\gamma) + \theta_n\right), \quad n \in \mathbb{Z}.$$

This operator plays a similar role as the differential operator $\mathcal{L}^i(x;\gamma) = D_{xx} - \cos(\phi^i_\pi(x;\gamma) + \theta(x))$ in section 3. The eigenvalue problem is an infinite dimensional matrix problem for a real and symmetric matrix. Thus the eigenvalues must be real.

In the discrete case, the continuous spectrum of semikinks is bounded. The spectrum is obtained by substituting $v_n = e^{-ikn}$ in (5.6) with $J^i_n = -2\varepsilon - \sqrt{1-\gamma^2}$ from which one obtains the following dispersion relation for such linear waves:

$$\Lambda = -\left(\sqrt{1-\gamma^2} + 4\varepsilon\sin^2(\tfrac{k}{2})\right). \tag{5.7}$$

Thus the continuous spectrum consists of the intervals $\pm i[\sqrt[4]{1-\gamma^2},\ \sqrt{\sqrt{1-\gamma^2}+4\varepsilon}]$ (recall that $\Lambda = \lambda^2$).

In the following two lemmas we will show that all eigenvalues of the linearization $L^1(\varepsilon;\gamma)$ are negative for ε small. Thus for ε small, the type 1 wave is always stable. For $\gamma = 1 - \varepsilon\widetilde{\gamma}$ and $\widetilde{\gamma} > \pi$, it can be shown that the linearization $L^3(\varepsilon;1-\varepsilon\widetilde{\gamma})$ has a positive eigenvalue for ε small. Hence the type 3 wave is unstable for γ near $\gamma_{\rm cr}$ and with ε small.

Lemma 5.2. *For $0 < \gamma < 1$ and ε small, the largest eigenvalue of the operator $L^1(\varepsilon;\gamma)$ is negative up to $\mathcal{O}(\varepsilon^2)$.*

Proof. The eigenvalue problem to calculate the stability of the monotone discrete π-kink $\phi^1_\pi(n;\varepsilon;\gamma)$, $n \in \mathbb{Z}$, is given by (5.6) with $i = 1$. Slightly modifying Baesens, Kim, and MacKay [2], the spatially decaying solution that corresponds to an eigenvalue of the above eigenvalue problem can be approximated by

$$v_n = \begin{cases} c\ell^{-n}, & n \le 0, \\ \hat{c}\, c\ell^{n-1}, & n \ge 1, \end{cases} \tag{5.8}$$

for some c, $\hat{c}$, and $|\ell| < 1$. The diagonal elements in $L^1(\varepsilon;\gamma)$ are $A_n = \sqrt{1-\gamma^2} + \mathcal{O}(\varepsilon^2)$, if $n \neq 0, 1$, $A_0 = \sqrt{1-\gamma^2} - \varepsilon\frac{\gamma\pi}{\sqrt{1-\gamma^2}} + \mathcal{O}(\varepsilon^2)$, and $A_1 = \sqrt{1-\gamma^2} + \varepsilon\frac{\gamma\pi}{\sqrt{1-\gamma^2}} + \mathcal{O}(\varepsilon^2)$. Thus $A_0 \neq A_1$; hence we need two parameters c and $\hat{c}$ (modifying [2], where $\hat{c} = \pm 1$ following from the symmetry $A_0 = A_1$).

For small nonzero ε, if we can match exponentially decaying solutions (5.8) on both sides from either end of the lattice to a central site, then we obtain a candidate for an eigenfunction. With (5.6), the parameters ℓ and $\hat{c}$ will be determined up to order ε. For $n \neq 0, 1$, the relation (5.6) gives up to order ε

$$\Lambda = -\sqrt{1-\gamma^2} + \varepsilon\left(\ell - 2 + \tfrac{1}{\ell}\right). \tag{5.9}$$

At the central sites $n = 0, 1$ we get up to order ε

$$\Lambda = -\sqrt{1-\gamma^2} + \varepsilon\tfrac{\gamma\pi}{\sqrt{1-\gamma^2}} + \varepsilon(\ell - 2 + \hat{c}); \tag{5.10}$$

$$\Lambda = -\sqrt{1-\gamma^2} - \varepsilon\tfrac{\gamma\pi}{\sqrt{1-\gamma^2}} + \varepsilon\left(\ell - 2 + \tfrac{1}{\hat{c}}\right). \tag{5.11}$$

Combining (5.9), (5.10), and (5.11) shows that there are two possible values for $\hat{c}$, being $\hat{c}_\pm = -\frac{\pi\gamma}{\sqrt{1-\gamma^2}} \pm \frac{\sqrt{1+(\pi^2-1)\gamma^2}}{\sqrt{1-\gamma^2}} + \mathcal{O}(\varepsilon)$, and leads to the eigenvalue Λ and the decay exponent ℓ as a function of ε and γ; i.e.,

$$\ell_\pm = \pm\frac{\sqrt{1-\gamma^2}}{\sqrt{1+(\pi^2-1)\gamma^2}} + \mathcal{O}(\varepsilon), \tag{5.12}$$

$$\Lambda_\pm = -\sqrt{1-\gamma^2} + \frac{\varepsilon}{\ell_\pm}(\ell_\pm - 1)^2 + \mathcal{O}(\varepsilon^2). \tag{5.13}$$

General Sturm–Liouville theory states that a critical eigenfunction that corresponds to the largest eigenvalue of a continuous eigenvalue problem does not vanish, except probably at $x \to \pm\infty$. This theorem can also be extended to a discrete eigenvalue problem such that the most critical eigenvector does not have sign changes [1]. Thus, if we have a solution of the form (5.8) with $\ell > 0$, then it is the critical eigenvector.

From (5.12), we see that $\ell_+ > 0$; thus the largest eigenvalue Λ_+ from (5.13) is in the gap between zero and the interval associated with the continuous spectrum, i.e., $\Lambda_+ < 0$. ∎

Remark 5.3. From the details in the proof, note that weak coupling with strong bias current leads to one additional eigenvalue associated with ℓ_-, where $\ell_- < 0$ and $|\ell_-| < 1$. This

indicates that the eigenvector of the form (5.8) is localized but has out of phase configuration, i.e., has infinitely many sign changes. This is a typical characteristic of a "high-frequency" eigenvalue which is confirmed by the fact that Λ_- is indeed smaller than the phonon band. The presence of a high-frequency eigenvalue of a kink was previously reported by Braun, Kivshar, and Peyrard [6] in their study of the Frenkel–Kontorova model with the Peyrard–Remoissenet potential [26].

For γ close to 1, i.e., $\gamma = 1 - \varepsilon\widetilde{\gamma}$, with $\widetilde{\gamma} > \pi$, both the type 1 and the type 3 waves as given in (5.5) can be analyzed. In the following lemma we also show that both types have a high-frequency eigenvalue.

Lemma 5.4. *For $\gamma = 1-\varepsilon\widetilde{\gamma}$, with $\widetilde{\gamma} > \pi$, the largest eigenvalue of the operator $L^1(\varepsilon; 1-\varepsilon\widetilde{\gamma})$ is strictly negative and the largest eigenvalue of the operator $L^3(\varepsilon; 1-\varepsilon\widetilde{\gamma})$ is strictly positive.*

Proof. As before, we write for an eigenfunction

$$v_n = \begin{cases} c\ell^{-n}, & n \le 0, \\ \hat{c}\, c\ell^{n-1}, & n \ge 1, \end{cases}$$

for some c, $\hat{c}$, and $|\ell| < 1$, and we substitute this in the eigenvalue problem, which leads to the equations

$$\Lambda = -\sin\left(\sqrt{\varepsilon 2\widetilde{\gamma}} + \mathcal{O}(\varepsilon)\right) + \varepsilon(1/\ell - 2 + \ell); \tag{5.14}$$

$$\Lambda = \mp\sin\left(\sqrt{2\varepsilon(\widetilde{\gamma} - \pi)} + \mathcal{O}(\varepsilon)\right) + \varepsilon(\ell - 2 + \hat{c}); \tag{5.15}$$

$$\Lambda = -\sin\left(\sqrt{2\varepsilon(\widetilde{\gamma} + \pi)} + \mathcal{O}(\varepsilon)\right) + \varepsilon(\ell - 2 + 1/\hat{c}), \tag{5.16}$$

where the $\mp$-sign in the second equation is a minus sign for the eigenvalue problem associated with the type 1 wave and a plus sign in case of the type 3 wave. Again, by subtracting (5.16) from (5.15), we get a quadratic equation for $\hat{c}$, with two solutions, one of order $\frac{1}{\sqrt{\varepsilon}}$ and one of order $\sqrt{\varepsilon}$ (which is most easily found by writing the equation as a quadratic equation for $\frac{1}{\hat{c}}$):

$$\begin{aligned} \tfrac{1}{\hat{c}_1} &= \tfrac{1}{\sqrt{\varepsilon}}\left(\sqrt{2(\widetilde{\gamma} + \pi)} \mp \sqrt{2(\widetilde{\gamma} - \pi)} + \mathcal{O}(\sqrt{\varepsilon})\right); \\ \hat{c}_2 &= \tfrac{1}{\sqrt{\varepsilon}}\left(-\sqrt{2(\widetilde{\gamma} + \pi)} \pm \sqrt{2(\widetilde{\gamma} - \pi)} + \mathcal{O}(\sqrt{\varepsilon}\right). \end{aligned}$$

Combining (5.14) and (5.15), respectively, (5.16), and using the two expressions above give that in both cases ℓ is of order $\sqrt{\varepsilon}$ and is given by

$$\begin{aligned} \tfrac{1}{\ell_1} &= \tfrac{1}{\sqrt{\varepsilon}}\left(\sqrt{2\widetilde{\gamma}} \mp \sqrt{2(\widetilde{\gamma} - \pi)} + \mathcal{O}(\sqrt{\varepsilon})\right); \\ \tfrac{1}{\ell_2} &= \tfrac{1}{\sqrt{\varepsilon}}\left(\sqrt{2\widetilde{\gamma}} - \sqrt{2(\widetilde{\gamma} + \pi)} + \mathcal{O}(\sqrt{\varepsilon}\right). \end{aligned}$$

Finally, substitution into (5.14) shows that

$$\begin{aligned} \Lambda_1 &= \mp\sqrt{\varepsilon}\sqrt{2(\widetilde{\gamma} - \pi)} + \mathcal{O}(\varepsilon); \\ \Lambda_2 &= -\sqrt{\varepsilon}\sqrt{2(\widetilde{\gamma} + \pi)} + \mathcal{O}(\varepsilon). \end{aligned} \tag{5.17}$$

The eigenvalue that corresponds to $\ell > 0$ is Λ_1.

So clearly the largest eigenvalue Λ_1 is negative in case of the type 1 wave and is positive in case of the type 3 wave.

In addition, the operators $L^1(\varepsilon; 1 - \varepsilon\widetilde{\gamma})$ and $L^3(\varepsilon; 1 - \varepsilon\widetilde{\gamma})$ have the same high-frequency eigenvalue Λ_2 (up to order ε^2). ■

The proofs of Lemmas 5.2 and 5.4 show the presence of a high-frequency eigenvalue for a semikink in case the bias current is not small. In the following we will show that the eigenvalue appears when the bias current is larger than $\sqrt{\varepsilon} + \mathcal{O}(\varepsilon)$.

Because we do not have an analytic expression for the type 2 and type 3 semikinks in the small forcing limit, the analysis is done only for the type 1 semikink.

Lemma 5.5. *There is a critical value $\gamma_{\rm hf}$, $\gamma_{\rm hf} = \sqrt{\varepsilon} + \mathcal{O}(\varepsilon)$, such that for all $\gamma \in (\gamma_{\rm hf}, \gamma_{\rm cr})$ the operator $L^1(\varepsilon; \hat{\gamma})$ has a high-frequency eigenvalue that up to $\mathcal{O}(\varepsilon^3)$ is attached to the lowest boundary of the continuous spectrum. The corresponding eigenvector is localized and changes sign between any two adjacent sites.*

The appearance of this eigenvalue and the structure of its eigenvector is checked numerically in section 6.

Proof. Again, we write for an eigenvector

$$v_n = \begin{cases} c\ell^{-n}, & n \le 0, \\ \hat{c}\, c\ell^{n-1}, & n \ge 1, \end{cases}$$

for some c, $\hat{c}$, and $|\ell| < 1$, and we substitute this into the eigenvalue problem. We first consider $\gamma = \sqrt{\varepsilon}\hat{\gamma}$. This gives $A_n = 1 - \frac{\varepsilon\hat{\gamma}^2}{2} - \frac{\varepsilon^2\hat{\gamma}^4}{8} + \mathcal{O}(\varepsilon^{5/2})$, if $n \neq 0, 1$, $A_0 = 1 - \frac{\varepsilon\hat{\gamma}^2}{2} - \varepsilon^{3/2}\pi\hat{\gamma} - \varepsilon^2(\frac{\hat{\gamma}^4}{8} + \frac{\pi^2}{2}) + \mathcal{O}(\varepsilon^{5/2})$, and $A_1 = 1 - \frac{\varepsilon\hat{\gamma}^2}{2} + \varepsilon^{3/2}\pi\hat{\gamma} - \varepsilon^2(\frac{\hat{\gamma}^4}{8} + \frac{\pi^2}{2}) + \mathcal{O}(\varepsilon^{5/2})$.

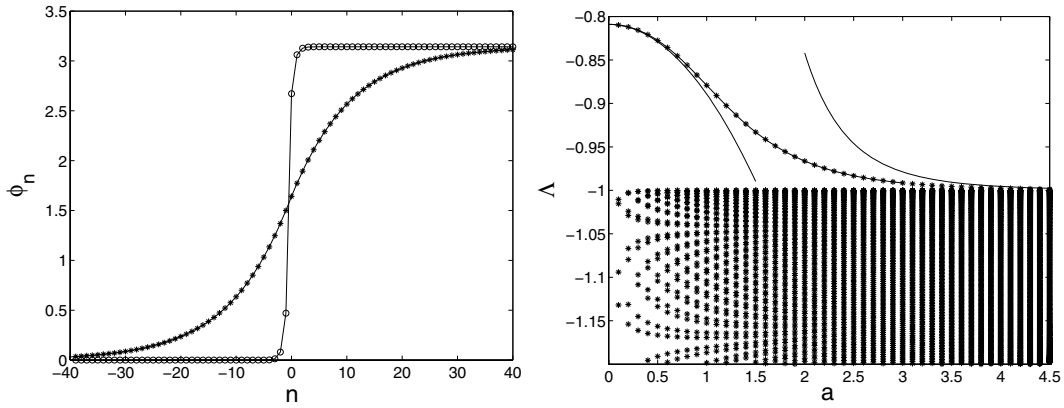
Using the same procedures, this implies $\hat{c}_\pm = -\sqrt{\varepsilon}\pi\hat{\gamma} - \varepsilon^{3/2}\hat{\gamma}^3\pi \pm \sqrt{\pi^2\hat{\gamma}^2\varepsilon + 2\pi^2\hat{\gamma}^4\varepsilon^2 + 1}$ and

$$\frac{1}{\ell_\pm} = \varepsilon\pi^2 \pm 2\sqrt{\pi^2\hat{\gamma}^2\varepsilon + 2\pi^2\hat{\gamma}^4\varepsilon^2 + 1} = \pm 1 + \frac{\varepsilon\pi^2(1 \pm \hat{\gamma}^2)}{2} + \mathcal{O}(\varepsilon^2).$$

For $\hat{\gamma} > 1$ there are two solutions $|\ell_\pm| < 1$; $\ell_+ > 0$ corresponds to the largest eigenvalue and also exists for $\hat{\gamma} \le 1$ (Lemma 5.2); $\ell_- < 0$, so that its associated eigenvector v_n indeed changes sign between any two adjacent sites.

The value $\gamma_{\rm hf} = \sqrt{\varepsilon} + \mathcal{O}(\varepsilon)$ indicates the appearance of this high-frequency eigenvalue from the continuous spectrum. It follows from a straightforward analysis that this eigenvalue exists, i.e., $\ell_- \in (-1, 0)$ exists, for all $\gamma \in (\gamma_{\rm hf}, \gamma_{\rm cr})$, and the corresponding eigenvalue is $\Lambda_- = -1 + (\frac{\hat{\gamma}^2}{2} - 4)\varepsilon + \hat{\gamma}^4/8\varepsilon^2 + \mathcal{O}(\varepsilon^3)$. Up to $\mathcal{O}(\varepsilon^3)$ this eigenvalue is nothing else but the lower boundary of the continuous spectrum. ■

6. Numerical computations of the discrete system. To accompany our analytical results, we have used numerical calculations. For that purpose, we have made a continuation program based on a Newton iteration technique to obtain the stationary kink equilibria of (2.3) and (2.4) and an eigenvalue problem solver in MATLAB. To start the iteration, one can either choose the continuum solutions discussed in section 3, i.e., the case where the lattice spacing parameter $a = 0$, or trace the equilibria from the uncoupled limit $\varepsilon = 0$ ($a \to \infty$) as discussed in the previous section. We use the number of computational sites $2N = 800$ for parameter values of $a = 0.05$ or larger ($\varepsilon = 20$ or lower).



(a) (b)

Figure 6. (a) *Two lattice semifluxons of type* 1 *with no bias current* ($\gamma = 0$) *are plotted as a function of the lattice index, namely, the kink for strong coupling with* $\varepsilon = 100$ *(equivalently, a very small lattice spacing* $a = 0.1$*)* $(-*-)$*, i.e., close to* (3.10)*, and the kink for weak coupling with* $\varepsilon = \frac{1}{4}$ *(equivalently, a large lattice spacing* $a = 2$*)* $(-\mathrm{o}-)$*.* (b) *Numerically computed spectrum of a lattice semifluxon against the lattice spacing parameter* a *with* $\gamma = 0$*. We used the number of sites* $2N = 300$*. We zoom in the plot of spectra around* -1 *for clarity. The bold solid line is the calculated approximate function for the point spectrum using perturbation theory for a small, respectively, for* ε *small.*

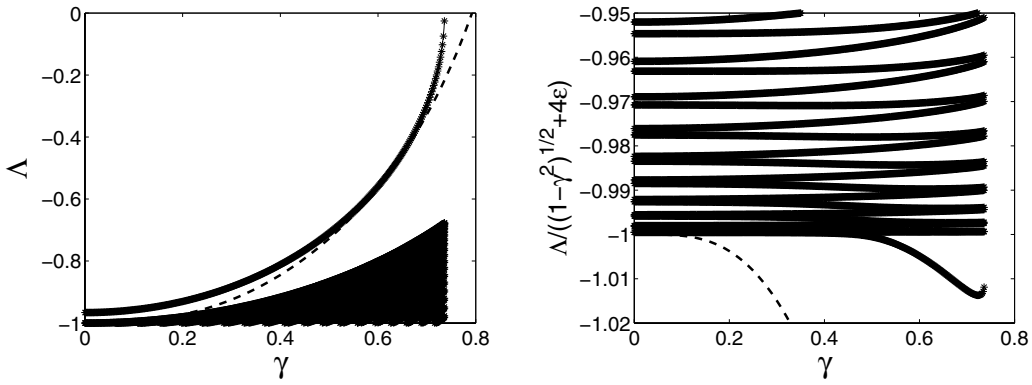
6.1. Stability of type 1 lattice semifluxon. The type 1 lattice semifluxon $\Phi_\pi^1(n;\varepsilon;0)$, $n \in \mathbb{Z}$, has been studied analytically both in the strong coupling limit ($a \ll 1$, or $\varepsilon \gg 1$) and the weak coupling limit ($a \gg 1$, $\varepsilon \ll 1$). In Figure 6(a), $\Phi_\pi^1(n;\varepsilon;0)$ is plotted for two different values of the coupling parameter ε. For a given value of ε, one can use as an initial guess in the numerical procedure a solution either from the continuous limit (3.10) or from the uncoupled limit that has been discussed in the preceding sections.

In Figure 6(b), we present the numerically calculated spectrum of the type 1 semifluxon with $\gamma = 0$ as a function of the lattice spacing parameter. The approximate largest eigenvalue (4.9), derived for a small, and the one derived in Lemma 5.2 for a large, are in good agreement with the numerically obtained largest eigenvalue. Any eigenvalue below $\Lambda = -1$ belongs to the continuous spectrum. For a close to zero we do not see dense spectra because of the number of sites we used. By increasing the site number we will obtain a more dense spectrum.

There is only one eigenvalue outside the phonon bands—the largest eigenvalue as studied in Lemma 5.2. This is in contrast to the case of an ordinary lattice 2π-kink [16, 19] where there is an internal mode bifurcating from the phonon band when the parameter a increases.

If Figure 6(b) shows the spectrum of the type 1 semifluxon as a function of the coupling parameter ε ($\varepsilon = 1/a^2$) for a fixed bias current γ, in Figure 7 we present the numerically calculated spectrum of the type 1 lattice semifluxon as a function of γ for a fixed ε, $\varepsilon = 0.25$.

Lemmas 5.2 and 5.5 established the existence of two eigenvalues (for ε small enough) for the stability problem associated to the type 1 semifluxon, the largest eigenvalue, and an additional eigenvalue which bifurcates from the lower edge of the phonon band for bias current $\gamma > \gamma_{\mathrm{hf}}$. It follows from the numerical simulations that these are indeed the only two eigenvalues (Figure 7). For $\varepsilon = 0.25$, this minimum bias current γ_{hf} is approximately 0.466.



(a) (b)

Figure 7. (a) *Spectrum of the type* 1 *semifluxon as a function of the applied bias current* γ *for a value of the coupling constant* $\varepsilon = 0.25$. *The dashed line is a theoretical prediction from* (5.13). *In* (b) *we zoom in on the spectrum around* -1 *for clarity. The spectrum is normalized to the lower edge of the phonon band, i.e.,* $\sqrt{1-\gamma^2}+4\varepsilon$, *such that the appearance of a high frequency eigenvalue can be seen clearly.*

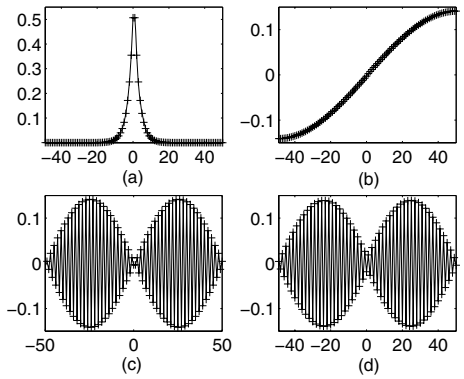


Figure 8. *The eigenvectors associated with the two largest eigenvalues and the two smallest eigenvalues of the truncated* $2N \times 2N$*-matrix associated with* $L^1(\varepsilon;\gamma)$ *(the type* 1 *discrete semikink) for* $\varepsilon = 0.25$ *and* $\gamma = 0$. *The results are shown with* $2N = 100$ *for clarity. Shown are the eigenvector of* (a) *the largest eigenvalue,* (b) *the second one, and* (c)–(d) *the last two eigenvalues. There is only one eigenvalue for* $L^1(\varepsilon;\gamma)$ *as there is only one localized eigenvector.*

Interestingly, according to Lemma 5.5 the bifurcation appears at $\gamma_{\text{hf}} = \sqrt{\varepsilon} = 0.5$ at leading order in ε. This is in remarkably good agreement with the numerical result, especially since the error is $\mathcal{O}(\varepsilon)$ and $\varepsilon = 0.25$.

To picture the appearance of the high-frequency eigenvalue, all eigenvalues for the truncated $2N \times 2N$-matrix associated with $L^1(\varepsilon;\gamma)$ are determined and the eigenvectors of the two largest eigenvalues and the two smallest eigenvalues are presented in Figures 8–10 for various values of γ and a fixed ε. It can be observed that there is always a localized eigenfunction associated with the largest eigenvalue. In Figure 8 ($\gamma = 0$), none of the other eigenvectors can be associated with localized eigenfunctions, and in Figures 9 and 10, the birth of the localized

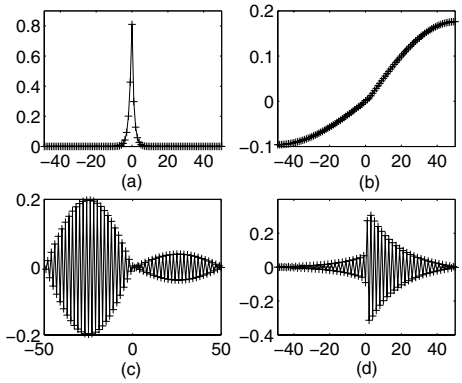


Figure 9. *The same as Figure 8 for $\gamma = 0.5$.*

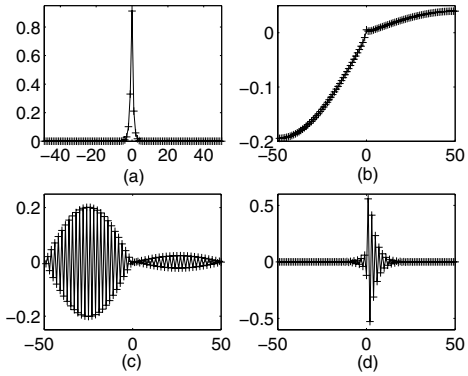


Figure 10. *The same as Figure 8 for $\gamma = 0.7$. Note that there are now two localized eigenvectors shown in* (a) *and* (d). *The smallest eigenvalue associated with* (d) *is* -1.7357, *while the lower edge of the phonon band is* -1.7141. *Note also that neighboring sites of the eigenvector in* (d) *move out of phase, indicating a high-frequency mode, contrary to the semikink's low-frequency mode in* (a).

eigenfunction associated with the smallest eigenvalue can be observed.

If we keep increasing γ further, then there is a critical applied bias current at which the largest eigenvalue becomes 0. Numerical computations show that this critical value is $\gamma_{\rm cr}(\varepsilon)$, above which static lattice semifluxons disappear.

The critical bias current for the existence of a static type 1 lattice semifluxon in the continuum limit and for a very weak coupling in the discrete system were discussed and analytical expressions were given in sections 3, 4, and 5, respectively. In Figure 11, the numerically calculated critical bias current $\gamma_{\rm cr}$ of the discrete system (2.3) as a function of the lattice spacing a is presented. The approximate functions, given in (4.2) for small a, and in (5.4) for large a (small ε), are presented as dashed lines.

6.2. Instability of type 2 lattice semifluxon. In the continuum models we have seen that for γ small, the instability of the type 2 semikink is mainly determined by the instability of the 3π-kink in the continuum models for $\gamma = 0$. So we start this section by looking at the

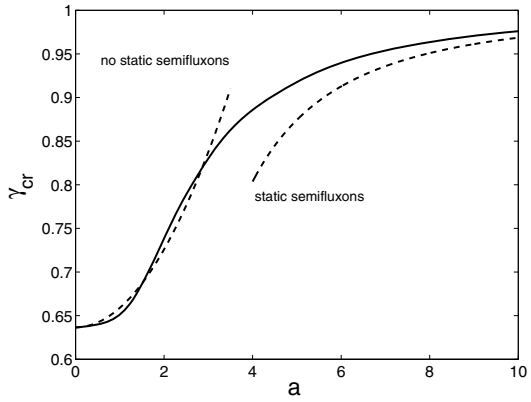


Figure 11. *The critical bias current of a static π-kink as a function of the lattice spacing parameter a. For γ above the critical current there is no static π-kink solution. The solid line is a numerically obtained curve. Dashed lines are the theoretical predictions* (4.2) *for $a \ll 1$, respectively,* (5.4) *for $a \gg 1$ ($\varepsilon \ll 1$).*

stability of the 3π-kink in the discrete model. We will denote the 3π-kink by $\Phi^2_{3\pi}(n;\varepsilon;0)$, where as before the coupling parameter ε and the lattice spacing a are related by $\varepsilon = \frac{1}{a^2}$.

Using our continuation program, we have followed a 3π-kink solution from the continuous limit $0 < a \ll 1$ up to the uncoupled situation $\varepsilon = 0$ (i.e., $a = \infty$). We obtain that $\Phi^2_{3\pi}(n;0;0)$ is given by

$$\Phi^2_{3\pi}(n;0;0) = \begin{cases} 0, & n = -1, -2, \ldots, \\ 2\pi, & n = 0, \\ \pi, & n = 1, \\ 3\pi, & n = 2, 3, \ldots. \end{cases} \tag{6.1}$$

Note that this discrete configuration is not monotonically increasing, as opposed to the continuum configuration, which is monotonic.

In Figure 12, we present the numerically obtained eigenvalues of a 3π-kink as a function of the lattice spacing a. For small a, the largest eigenvalue is indeed increasing as is predicted by the perturbation theory (4.10). As soon as the lattice spacing is of order one, the largest eigenvalue decreases and becomes zero at approximately $a = 1.7521$.

After establishing that increasing the lattice spacing can stabilize the 3π-kink at $\gamma = 0$, we continue by looking at the stability of the 2π-kink for $\gamma > 0$. Interestingly, increasing the lattice spacing does not stabilize a type 2 semikink for $\gamma > 0$. In Figure 13, we show a plot of the type 2 semikinks for two values of ε as well as a plot of the largest eigenvalue as a function of ε for two particular values of γ, namely, $\gamma = 0.01$ and $\gamma = 0.1$. We present the largest eigenvalue as a function of the coupling ε instead of the lattice spacing a as the eigenvalue changes most for small coupling (large lattice spacing). From Figure 13 it follows that the solutions are unstable even in the weak-coupling limit. This is interesting as in the limit for $\gamma \to 0$, the type 2 semikink can be seen as a concatenation of a 3π-kink and a -2π-kink. Both the 3π-kink and -2π-kink are stable for the coupling ε sufficiently small, while the type 2 semikink turns out to be unstable.

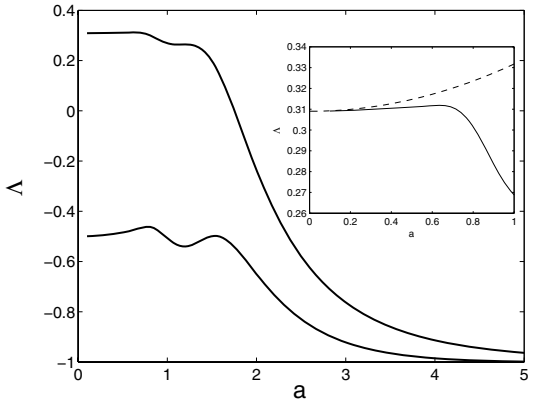
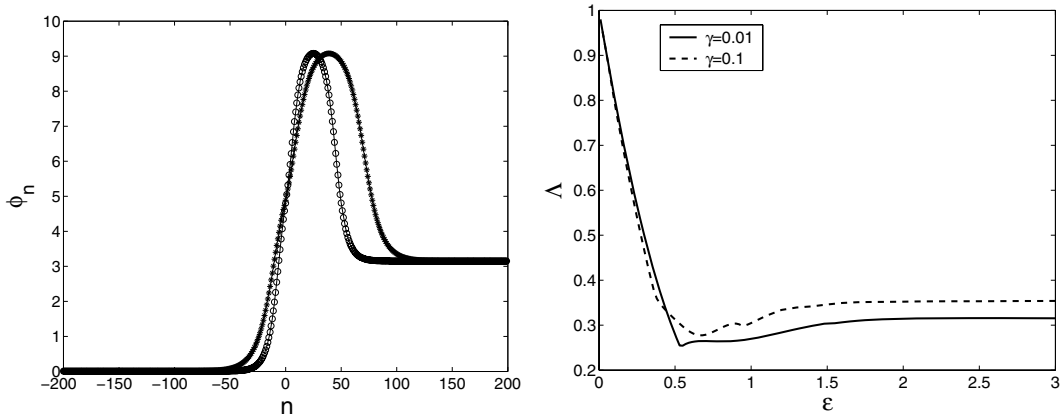


Figure 12. *Plot of the eigenvalues of a 3π-kink as a function of the lattice spacing parameter a. We zoom in on the region with $a \ll 1$, where it shows that turning the lattice spacing on destabilizes the kink. The dashed line depicts the analytically computed approximation (4.10) to the largest eigenvalue of the 3π-kink.*

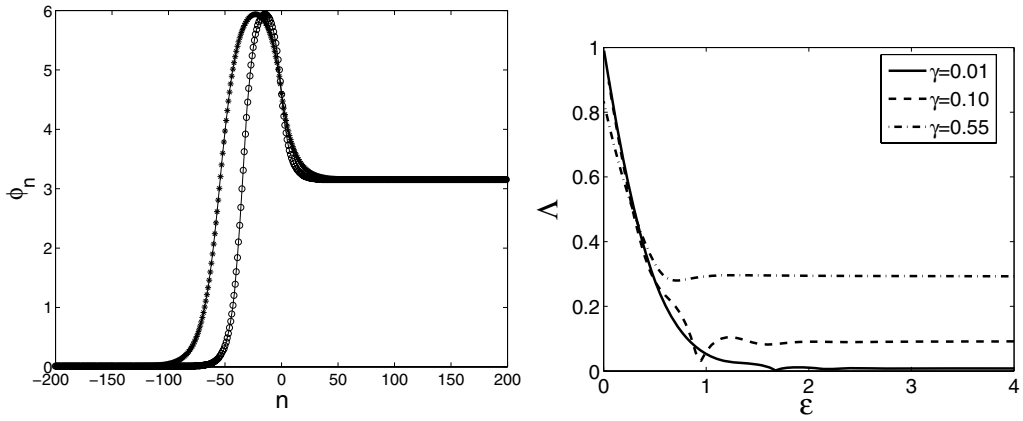


(a) (b)

Figure 13. (a) *Plot of a type 2 semikink with $\gamma = 0.01$ for $\varepsilon = 100$ $(-*-)$ and $\varepsilon = 40$ $(-\mathrm{o}-)$.* (b) *Plot of the largest eigenvalue of a type 2 semikink as a function of the coupling parameter ε. When $\varepsilon = 0$, the eigenvalue converges to $\Lambda = \sqrt{1-\gamma^2}$.*

This instability issue can be explained by looking at the expression of a type 2 semikink when it is uncoupled ($\varepsilon = 0$). For the two particular choices of γ above, we get from the simulations that the configurations of these semikinks are given by

$$\Phi^2_\pi(n;0;0.01) = \begin{cases} 0 + \arcsin(0.01), & n \leq -1, \\ \pi - \arcsin(0.01), & n = 0, \\ \pi + \arcsin(0.01), & n = 1, \\ 3\pi + \arcsin(0.01), & 2 \leq n \geq 8, \\ 2\pi - \arcsin(0.01), & n = 9, \\ \pi + \arcsin(0.01), & n \geq 10, \end{cases} \tag{6.2}$$



(a) (b)

Figure 14. (a) *Plot of a type* 3 *semikink with* $\gamma = 0.01$ *for* $\varepsilon = 100$ *(*$-*-$*) and* $\varepsilon = 40$ *(*$-$o$-$*).* (b) *Plot of the largest eigenvalue of a type* 3 *semikink as a function of the coupling parameter* ε*. When* $\varepsilon = 0$*, the eigenvalue converges to* $\Lambda = \sqrt{1-\gamma^2}$*.*

and

$$\Phi_\pi^2(n;0;0.1) = \begin{cases} 0 + \arcsin(0.1), & n \le -1, \\ \pi - \arcsin(0.1), & n = 0, \\ \pi + \arcsin(0.1), & n = 1, \\ 3\pi + \arcsin(0.1), & 2 \le n \ge 6, \\ 2\pi - \arcsin(0.1), & n = 7, \\ \pi + \arcsin(0.1), & n \ge 8. \end{cases} \tag{6.3}$$

We see that there are two sites, namely, $n = 0$ and $n = 9$ for $\gamma = 0.01$ and $n = 0$ and $n = 7$ for $\gamma = 0.1$, where Φ takes the value of an unstable fixed point of the discrete system (2.3). Looking only at sites numbered $n = 2$ to $n \to \infty$, $\Phi_\pi^2(n;0;\gamma)$ can be viewed as a -2π lattice kink sitting on a site which is known to be unstable. If we look only at sites numbered $n = 6$ to $n \to -\infty$, $\Phi_\pi^2(n;0;\gamma)$ can be seen as a deformed 3π lattice kink (at site $n = 0$, the phase Φ takes the value π instead of the value 2π as in the 3π-kink). Hence, it seems that coupling between the two kinks due to the presence of a nonzero γ is responsible for the instability.

It has been discussed in the previous sections that there is a critical bias current γ^* for the existence of a type 2 lattice semikink in the continuum models. However, we did not numerically calculate the critical bias current $\gamma^*(a)$ for discrete system (2.3).

6.3. Instability of type 3 lattice semifluxon. In this section, we will consider the type 3 semikinks, which will be denoted by $\Phi_\pi^3(n;\varepsilon;\gamma)$. In Lemmas 3.9 and 4.7 it has been shown that these kinks are unstable in the continuum models for small or zero lattice spacing.

The largest eigenvalue of a lattice type 3 semifluxon for three particular values of γ, i.e., $\gamma = 0.01, 0.1, 0.55$, is presented in Figure 14. Even though in the limit for $\gamma \to 0$, a semifluxon of this type is a concatenation of a 2π-kink and a $-\pi$-kink which can both be stable in the discrete system, the type 3 semikink is unstable for all parameter values from the zero lattice

spacing limit all the way to the zero coupling one. The explanation is similar to the one for a type 2 semikink discussed above.

Indeed, for the three particular choices of γ above, $\Phi^3_\pi(n;0;\gamma)$ is given by

$$\Phi^3_\pi(n;0;0.01) = \begin{cases} 0+\arcsin(0.01), & n=-1,-2,\dots, \\ \pi-\arcsin(0.01), & n=-6, \\ 2\pi+\arcsin(0.01), & n=-5,\dots,0, \\ \pi+\arcsin(0.01), & n=1,2,\dots, \end{cases} \tag{6.4}$$

$$\Phi^3_\pi(n;0;0.1) = \begin{cases} 0+\arcsin(0.1), & n=-1,-2,\dots, \\ \pi-\arcsin(0.1), & n=-2, \\ 2\pi+\arcsin(0.1), & n=-1,0, \\ \pi+\arcsin(0.1), & n=1,2,\dots, \end{cases} \tag{6.5}$$

and

$$\Phi^3_\pi(n;0;0.55) = \begin{cases} 0+\arcsin(0.55), & n=-1,-2,\dots, \\ \pi-\arcsin(0.55), & n=0, \\ \pi+\arcsin(0.55), & n=1,2,\dots. \end{cases} \tag{6.6}$$

One interesting point to note for the type 3 semikink is that the number of sites with value 2π is decreasing as γ increases. Starting from the continuum approximation of a type 3 semikink as the initial guess for the continuation program, the 2π-plateau disappears for $\gamma \geq \gamma^*(a)$ (see (4.1)). For $\gamma > \gamma^*$ the configuration at $\varepsilon = 0$ is similar to the stable type 1 π-kink (5.2), apart from the value of the phase at the site with $n = 0$ (where the phase takes the value of an unstable fixed point).

Because analytical calculation of the spectrum of the type 3 semikink has been obtained in the small coupling limit and bias current close to 1 (5.17), it is worth comparing the analytical predictions with numerical computations. The theoretical calculations show that for ε small and γ close to 1, the type 3 semikink has at least two eigenvalues, one of which corresponds to a high-frequency mode and the other to a positive eigenvalue.

Using the continuation of (6.6) for $\varepsilon = 0.25$, the spectrum of the type 3 lattice semikink is presented in Figure 15 as a function of the applied bias current γ. Our numerics show that when γ is very close to $\gamma_{\rm cr}$, the type 3 semikink has three eigenvalues, one of which corresponds to a high-frequency mode and is below the phonon band, while the other two are above the phonon band. The birth of this high-frequency mode is shown in Figure 15(c) and is qualitatively similar to the case of the type 1 lattice semikink. The two eigenvalues which exist for all values of γ can be observed in Figure 15(a) and (b).

7. Conclusions. We have performed an existence and stability analysis for three types of lattice π-kink solutions of the discrete 0-π sine-Gordon equation and its continuum limits. Analytical results have been established in the continuum limits and in the weak-coupling case. It has been shown that in the continuous 0-π sine-Gordon equation, π-kinks of type 1 are stable and the other types are unstable. The introduction of discreteness destabilizes the unstable π-kinks even more. An approximation to the largest eigenvalue of all types of π-kinks has been derived both in the continuum and the weak coupling limits.

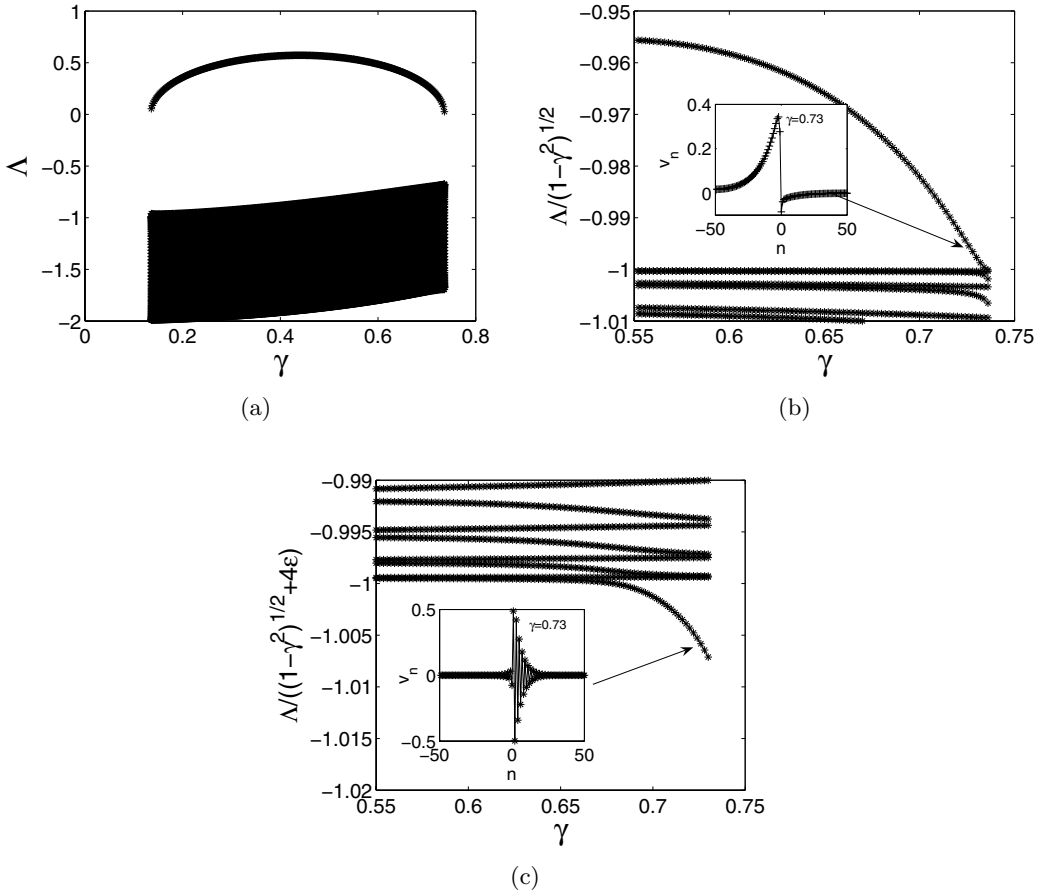


Figure 15. *Spectrum of the type* 3 *semifluxon as a function of the applied bias current* γ *for a value of the coupling constant* $\varepsilon = 0.25$. *In* (b) *and* (c) *we zoom in near the phonon band for clarity. In* (b), *the spectrum is normalized to the upper edge of the phonon band, i.e.,* $\sqrt{1-\gamma^2}$, *and in* (c) *it is normalized to the lower edge of the phonon band, i.e.,* $\sqrt{1-\gamma^2}+4\varepsilon$. *The disappearance of a high-frequency mode in the lower edge of the phonon band can be clearly observed in* (c). *The insets in* (b) *and* (c) *show the eigenfunctions of the two eigenvalues just above, respectively, just below, the phonon band for* $\gamma = 0.73$.

For future research, it is of interest to study the nucleation of kinks and antikinks when a constant force, or bias current, γ, that is above the critical value $\gamma_{\rm cr}$ is applied—see Figure 2(b). One question that can be addressed is the mechanism and the frequency of the nucleation as a function of the applied constant force, especially in the presence of a damping coefficient (which has not been considered in this paper). In work in progress, the stability of the type 3 semifluxons in the presence of defects is studied. These semifluxons are unstable, but the largest eigenvalue is close to zero. In fact, a type 3 semifluxon consists of a fluxon and a semifluxon with the opposite polarity. In experiments, the presence of a fluxon near a semifluxon can influence a junction measurement [12]. Because a fluxon can be pinned by a defect [17, 18], one can expect to have a stable type 3 semifluxon when there is a defect present in the system.

Acknowledgment. H. Susanto wishes to thank Panayotis Kevrekidis for numerous useful interactions and discussions.

REFERENCES

[1] F. V. Atkinson, *Discrete and Continuous Boundary Problems*, Math. Sci. Engrg. 8, Academic Press, New York, 1964.
[2] C. Baesens, S. Kim, and R. S. MacKay, *Localised modes on localised equilibria*, Phys. D, 113 (1998), pp. 242–247.
[3] N. J. Balmforth, R. V. Craster, and P. G. Kevrekidis, *Being stable and discrete*, Phys. D, 135 (2000), pp. 212–232.
[4] J. J. A. Baselmans, A. F. Morpurgo, B. J. van Wees, and T. M. Klapwijk, *Reversing the direction of the supercurrent in a controllable Josephson junction*, Nature, 397 (1999), pp. 43–45.
[5] O. M. Braun and Yu. S. Kivshar, *Nonlinear dynamics of the Frenkel-Kontorova model*, Phys. Rep., 306 (1998), pp. 1–108.
[6] O. M. Braun, Yu. S. Kivshar, and M. Peyrard, *Kink's internal modes in the Frenkel-Kontorova model*, Phys. Rev. E, 56 (1997), pp. 6050–6064.
[7] L. N. Bulaevskii, V. V. Kuzii, and A. A. Sobyanin, *Superconducting system with weak coupling to the current in the ground state*, JETP Lett., 25 (1977), pp. 290–294
[8] L. N. Bulaevskii, V. V. Kuzii, A. A. Sobyanin, and P. N. Lebedev, *On possibility of the spontaneous magnetic flux in a Josephson junction containing magnetic impurities*, Solid State Comm., 25 (1978), pp. 1053–1057.
[9] B. V. Chirikov, *A universal instability of many-dimensional oscillator systems*, Phys. Rep., 52 (1979), pp. 264–379.
[10] G. Derks, A. Doelman, S. A. van Gils, and T. P. P. Visser, *Travelling waves in a singularly perturbed sine-Gordon equation*, Phys. D, 180 (2003), pp. 40–70.
[11] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields*, 2nd ed., Springer-Verlag, New York, 1986.
[12] D. J. van Harlingen, *Phase-sensitive tests of the symmetry of the pairing state in the high-temperature superconductors—evidence for $d_{x^2-y^2}$ symmetry*, Rev. Modern Phys., 67 (1995), pp. 515–535.
[13] H. Hilgenkamp, Ariando, H. J. H. Smilde, D. H. A. Blank, G. Rijnders, H. Rogalla, J. R. Kirtley, and C. C. Tsuei, *Ordering and manipulation of the magnetic moments in large-scale superconducting π-loop arrays*, Nature, 422 (2003), pp. 50–53.
[14] T. Kato, *Perturbation Theory of Linear Operators*, Springer-Verlag, New York, 1976.
[15] T. Kato and M. Imada, *Vortices and quantum tunneling in current-biased 0-π-0 Josephson junctions of d-wave superconductors*, J. Phys. Soc. Japan, 66 (1997), pp. 1445–1449.
[16] P. G. Kevrekidis and C. K. R. T. Jones, *Bifurcation of internal solitary wave modes from the essential spectrum*, Phys. Rev. E, 61 (2000), pp. 3114–3121.
[17] Yu. S. Kivshar and B. A. Malomed, *Dynamics of solitons in nearly integrable systems*, Rev. Modern Phys., 61 (1989), pp. 763–915.
[18] Yu. S. Kivshar and B. A. Malomed, *Dynamics of solitons in nearly integrable systems: Addendum*, Rev. Modern Phys., 63 (1991), p. 211.
[19] Yu. S. Kivshar, D. E. Pelinovsky, T. Cretegny, and M. Peyrard, *Internal modes of solitary waves*, Phys. Rev. Lett., 80 (1998), pp. 5032–5035.
[20] A. B. Kuklov, V. S. Boyko, and J. Malinsky, *Instability in the current-biased 0-π Josephson junction*, Phys. Rev. B, 51 (1995), pp. 11965–11968.
[21] A. B. Kuklov, V. S. Boyko, and J. Malinsky, *Instability in the current-biased 0-π Josephson junction: Erratum*, Phys. Rev. B, 55 (1997), p. 11878.
[22] R. S. MacKay and J. A. Sepulchre, *Multistability in networks of weakly coupled bistable units*, Phys. D, 82 (1995), pp. 243–254.
[23] E. Mann, *Systematic perturbation theory for sine-Gordon solitons without use of inverse scattering methods*, J. Phys. A, 30 (1997), pp. 1227–1241.

[24] Y. Nomura, Y. H. Ichikawa, and A. T. Filippov, *Stochasticity in the Josephson map*, J. Plasma Phys., 56 (1996), pp. 493–506.
[25] D. E. Pelinovsky, P. G. Kevrekidis, and D. J. Frantzeskakis, *Stability of discrete solitons in nonlinear Schrödinger lattices*, Phys. D, 212 (2005), pp. 1–19.
[26] M. Peyrard and M. Remoissenet, *Solitonlike excitations in a one-dimensional atomic chain with a nonlinear deformable substrate potential*, Phys. Rev. B, 26 (1982), pp. 2886–2899.
[27] N. R. Quintero, A. Sánchez, and F. G. Mertens, *Anomalous resonance phenomena of solitary waves with internal modes*, Phys. Rev. Lett., 84 (2000), pp. 871–874.
[28] P. Rosenau, *Hamiltonian dynamics of dense chains and lattices: Or how to correct the continuum*, Phys. Lett. A, 311 (2003), pp. 39–52
[29] V. V. Ryazanov, V. A. Oboznov, A. Yu. Rusanov, A. V. Veretennikov, A. A. Golubov, and J. Aarts, *Coupling of two superconductors through a ferromagnet: Evidence for a π junction*, Phys. Rev. Lett., 86 (2001), pp. 2427–2430.
[30] A. Scott, *Nonlinear Science: Emergence and Dynamics of Coherent Structures*, Oxford University Press, Oxford, UK, 1999.
[31] H. Susanto and S. A. van Gils, *Instability of a lattice semifluxon in a current-biased 0-π array of Josephson junctions*, Phys. Rev. B, 69 (2004), pp. 092507–092510.
[32] H. Susanto, S. A. van Gils, T. P. P. Visser, Ariando, H. J. H. Smilde, and H. Hilgenkamp, *Static semifluxons in a long Josephson junction with π-discontinuity points*, Phys. Rev. B, 68 (2003), pp. 104501–104508.
[33] E. C. Titchmarsh, *Eigenfunction Expansions Associated with Second-Order Differential Equations*, 2nd ed., Oxford University Press, Oxford, UK, 1962.
[34] C. C. Tsuei and J. R. Kirtley, *Pairing symmetry in cuprate superconductors*, Rev. Modern Phys., 72 (2000), pp. 969–1016.
[35] O. Vávra, S. Gaži, D. S. Golubović, I. Vávra, J. Dérer, J. Verbeeck, G. Van Tendeloo, and V. V. Moshchalkov, *The 0 and the π phase Josephson coupling through an insulating barrier with magnetic impurities*, Phys. Rev. B, 74 (2006), pp. 020502–020505.
[36] K. Yosida, *Functional Analysis*, Springer-Verlag, New York, 1995.

SIAM J. APPLIED DYNAMICAL SYSTEMS
Vol. 6, No. 1, pp. 142–207

A Parameterization Method for the Computation of Invariant Tori and Their Whiskers in Quasi-Periodic Maps: Explorations and Mechanisms for the Breakdown of Hyperbolicity*

A. Haro† and R. de la Llave‡

Abstract. In two previous papers [*J. Differential Equations*, 228 (2006), pp. 530–579; *Discrete Contin. Dyn. Syst. Ser. B*, 6 (2006), pp. 1261–1300] we have developed fast algorithms for the computations of invariant tori in quasi-periodic systems and developed theorems that assess their accuracy. In this paper, we study the results of implementing these algorithms and study their performance in actual implementations. More importantly, we note that, due to the speed of the algorithms and the theoretical developments about their reliability, we can compute with confidence invariant objects close to the breakdown of their hyperbolicity properties. This allows us to identify a mechanism of loss of hyperbolicity and measure some of its quantitative regularities. We find that some systems lose hyperbolicity because the stable and unstable bundles approach each other but the Lyapunov multipliers remain away from 1. We find empirically that, close to the breakdown, the distances between the invariant bundles and the Lyapunov multipliers—which are natural measures of hyperbolicity—depend on the parameters, with power laws with universal exponents. We also observe that, even if the rigorous justifications in [*J. Differential Equations*, 228 (2006), pp. 530–579] are developed only for hyperbolic tori, the algorithms work also for elliptic tori in Hamiltonian systems. We can continue these tori and also compute some bifurcations at resonance which may lead to the existence of hyperbolic tori with nonorientable bundles. We compute manifolds tangent to nonorientable bundles.

Key words. invariant tori, invariant manifolds, quasi-periodic systems

AMS subject classifications. 37M99, 37D10, 70K40, 70K43

DOI. 10.1137/050637327

1. Introduction.

In this paper we implement the numerical algorithms for the computation of invariant tori and their whiskers in quasi-periodically forced systems presented in [HdlL05a]. The whiskers include in particular the stable and unstable manifolds, but also the nonresonant and slow invariant manifolds attached to the whiskers.

The algorithms are based on the parameterization method introduced in [HdlL06b] for quasi-periodic systems (see also [CFdlL03a, CFdlL03b, CFdlL05] for the parameterization

*Received by the editors August 1, 2005; accepted for publication (in revised form) by T. Kaper September 11, 2006; published electronically March 15, 2007.
http://www.siam.org/journals/siads/6-1/63732.html

†Departament de Matemàtica Aplicada i Anàlisi, Facultat de Matemàtiques, Universitat de Barcelona, Gran Via de les Corts Catalanes 585, 08007 Barcelona, Spain (alex@maia.ub.es). This work was initiated when this author was enjoying a Fulbright Scholarship at The University of Texas at Austin in the year 2000. He has been supported by the MCyt/FEDER grant BFM2000–805 and the Catalan grant 2000SGR–27 and the INTAS project 00-221. His work is also supported by the MCyT/FEDER grant BFM2003-07521-C02-01.

‡Dept. of Mathematics, The University of Texas at Austin, 1 University Station C1200, Austin, TX 78712-0257 (llave@math.utexas.edu). The work of this author has been partially supported by N. S. F. grants. He enjoyed a Dean's Fellowship at U.T. Austin in spring of 2003. Visits of this author to Barcelona were supported by FBBV and ICREA.

method in other contexts).

Using the numerical procedures developed, we analyze several examples that have been considered in the literature in regimes where the results of [HdlL06b] apply.

Since the results in [HdlL06b] give "a posteriori estimates," we know that, in some regimes, the numerical computations correspond to true solutions of the problem close by. We will also use the term "validating" as a synonym for "a posteriori."

It is important to realize that once we have validating results, to guarantee the validity of the calculations, we do not need to study the algorithms. We just need to check—after the calculation is done—that the calculation leads to objects that solve the desired equation with a small error and that they satisfy the nondegeneracy assumptions required for the theorem. Of course, an analysis of the algorithms may be useful for obtaining an idea of how many steps or what number of modes will be needed.

Once the algorithms have been benchmarked in regions where they are backed up by a rigorous analysis, we explore situations which are not yet covered by a rigorous analysis. This allows us to formulate conjectures, which we hope will stimulate new analytical results. The main empirical findings of this exploration are as follows.

• We have identified an apparently robust scenario for the breakdown of exponential dichotomies and, in particular for the breakdown of hyperbolicity.

In this "bundle merging" scenario, the invariant bundles approach each other and their distance tends to zero as the parameters approach a critical value. Nevertheless, before the critical value, the spectrum remains uniformly separated.

We have identified several scaling properties in this scenario and we have found that, in the examples considered, the exponents of these scaling laws are universal.

In sections 3.2 and 6 we report in detail these phenomena. See also [HdlL06a].

• Even if the mathematical justifications in [HdlL06b] are stated only for normally hyperbolic tori, we note that some of the algorithms in [HdlL05a] work for tori whose linearizations are elliptic in the normal directions.

Hence, we can continue these tori with respect to parameters of the map. This has already been done in [CJ00], using the Newton method. In this paper we also use algorithms based on reducibility. The refined numerical study here allows us to explore bifurcations that happen when there are resonances between the normal modes and the internal modes. This leads to a bifurcation in which the torus becomes hyperbolic. We note that, in this case, the stable and unstable bundles of the tori may be nonorientable. The appearance of nonorientable bundles in applications is the main reason why the algorithms developed in the first part took special pains to include nonorientable bundles.

There are justifications of persistence of elliptic tori in the literature [Eli88, JS96, JV97, Pös89], in rather general contexts, but they are not designed to validate numerics. An a posteriori result that can be found in a slightly more restrictive context is in [JdlLZ99]. We hope that an a posteriori result will be available in the near future. Some results on the bifurcations of elliptic tori at resonance are available in [BHJ$^+$03], but they do not include the study of the hyperbolic manifolds generated. We hope that an a posteriori result tailored toward our numerics will be available soon.

1.1. Quasi-periodic maps. The examples we study in this paper are quasi-periodic perturbations of two-dimensional (2D) maps. We will take the external perturbation to be quasi-

periodic with one frequency. That is, we will consider maps of the form

$$\begin{pmatrix} x \\ y \\ \theta \end{pmatrix} \longrightarrow \begin{pmatrix} f_\varepsilon(x,y,\theta) \\ g_\varepsilon(x,y,\theta) \\ \theta+\omega \end{pmatrix}, \tag{1}$$

where $x, y \in \mathbb{R}$ or $\mathbb{T}$ are the state variables, $\theta \in \mathbb{T}$ is the external angle, and ε is a parameter (for $\varepsilon = 0$, the functions f_0, g_0 do not depend on θ). The frequency $\omega \in \mathbb{R}$ is supposed to be irrational (in fact, we assume that ω is Diophantine, although for some of the algorithms and results this extra assumption is not necessary). By introducing the 2D variable $z = (x, y)$, we write (1) in a more compact way as

$$\begin{pmatrix} z \\ \theta \end{pmatrix} \longrightarrow \begin{pmatrix} F_\varepsilon(z,\theta) \\ \theta+\omega \end{pmatrix}. \tag{2}$$

This is the simplest nontrivial case, and it is a natural first step to test the methods. It seems that adding more variables to the phase space will not be too complicated, but we thought it would be better to study in detail some new phenomena that appear already. Of course, we plan to come back to these issues.

Maps of the form (2) appear also as the surface of section maps of flows which are quasi-periodically perturbed. In this case, ω is the ratio of the two external frequencies. We will not present any specific examples of flows (see [CJ00]). The papers [HdlL06b, HdlL05a] present some algorithms that are specifically designed for flows.

The examples we consider are quasi-periodic perturbations of two very well-known maps: the *Hénon map* and the *standard map*. We have included examples in which the perturbations are analytic and another one in which the perturbation is just Lipschitz. As we will see in the examples, the algorithms yield high precision results (in some cases, the errors are a few times the order of magnitude of the machine epsilon). We will see that the algorithms can handle 10^5 Fourier coefficients in hours of work of a by now outdated desktop computer. (High order expansions are required in the study of the Lipschitz example here, but a large number of Fourier coefficients are needed in the computation of higher dimensional invariant tori.)

1.2. Some similar calculations in the literature. Let us end this introductory section recalling some of the areas in which the computation of invariant manifolds has proved fruitful. Since invariant manifolds organize the global behavior of a dynamical system, their importance is both theoretical and practical. We will focus on the practice in this paper. For instance, codimension 1 invariant manifolds are barriers in the phase space and are useful in describing transport and formation of resonances. For 2D systems, stable and unstable manifolds of hyperbolic periodic orbits have been used to study the geometrical structure of the attractors in the Hénon map [Sim79], the mechanisms of destruction of invariant tori for the standard map [OS87], transport phenomena in area preserving maps [Mei92], etc. In four dimensions, the stable and unstable manifolds of a center manifold of a hyperbolic-elliptic fixed point of the Froeschlé map have been used to bound a stability region around the totally elliptic fixed point [Har99].

In celestial mechanics, stable and unstable manifolds of quasi-periodic solutions in the Solar System can be used also to perform transfer orbits in space missions [Sim99, GJMS91b, GJMS91a, ESA, NAS]. The list of applications is far from complete.

We mention some papers dealing with a quasi-periodically forced system. In [BLW91], the whiskers of invariant tori are the tool used to explain transport in the homoclinic and heteroclinic tangle regions in the quasi-periodic 2D case. In [OF00] these manifolds explain the structure of the attractors in a quasi-periodically forced Hénon map (what we call the rotating Hénon map), and the computational method is based on a general algorithm [BOV97]. In these papers, the manifolds are approximated using simplicial complexes. The papers [LM00, LM03] compute invariant manifolds for volume preserving maps. In [Har05], these manifolds are related to the breakdown of KAM tori in a quasi-periodically forced standard map (what we call the rotating standard map).

Since the objects we compute are invariant tori (and their whiskers), we will use coordinates that are more adapted to them. It is natural to parameterize invariant tori using angular coordinates, so the functions are periodic in the angular coordinates and we use truncated Fourier series in their approximations. It is also natural to parameterize invariant manifolds attached to invariant tori using angular Cartesian coordinates, and then use truncated Fourier–Taylor series in their approximations.

The use of truncated Fourier series to compute invariant tori using Newton's method for the invariance equation appears, for instance, in [WR87, Har02] (KAM tori), [dlLT94, CJ00] (invariant tori in quasi-periodic systems), [GJ04] (lower dimensional tori in autonomous conservative systems), and [MKM97] (PDE approach in continuous systems). Other different approaches to compute invariant tori can be found, for instance, in [DLR91, DL95, Sim98, ERS00, Tru00].

We also note that the use of high order expansions (Taylor, Fourier, Fourier–Taylor, Lindstedt, etc.) to compute invariant manifolds has been standard in the work of C. Simó since the 1970's and it has been used in many of his papers, often without mention.

2. Some implementation details.

In this section, we provide some details about the implementation we have carried out.

At the moment, we have a rather complete toolkit to manipulate periodic functions of one angle variable. This allows us to consider, rather comfortably, one-dimensional (1D) invariant tori and their whiskers. The numerical results reported in this paper are only for the case that the ambient space is three-dimensional (3D) (2D maps are subject to a quasi-periodic perturbation of one frequency).

The restriction to two dimensions allows some mathematical simplifications, and allows more complete visualization. It is a natural first step. Increasing the dimension of the maps seems rather straightforward but increasing the dimension of the tori requires more effort. Both of these improvements are actively pursued.

2.1. The parameterization equations for tori and their whiskers.

In this section, we recall concisely the main algorithms developed in [HdlL05a].

2.1.1. Equations for invariant tori and algorithms to solve them.

Following the parameterization method, to find a 1D torus invariant for (1) we seek a mapping $K : \mathbb{T} \to \mathbb{R}^2$ satisfying

$$F(K(\theta), \theta) = K(\theta + \omega). \tag{3}$$

In [HdlL05a] we consider several algorithms to solve (3). For the purposes of this paper, the only one we will consider is the following:

- *Newton method.* Given a K that approximately solves (3), we obtain $\hat{K}$, an improved approximate solution of (3), by solving the linearization of (3).
  More explicitly, if K satisfies

$$r(\theta) \equiv F(K(\theta-\omega), \theta-\omega) - K(\theta)$$

  (we should think of r as the error), an improved solution is $\tilde{K} = K + h$, where h is obtained by solving

$$\mathrm{D}F(K(\theta-\omega), \theta-\omega)h(\theta-\omega) - h(\theta) = -r(\theta). \tag{4}$$

The equations (4), which define the Newton method, can be treated in different ways.

- *Large matrix method.* The most straightforward method is simply to discretize (4) in a basis of discretization and then apply a linear equation solver.
  This method is relatively easy to implement and is quite robust. Even if the justifications of this method in [HdlL05a] are for normally hyperbolic tori, we have found empirically that the Newton method also works well for normally elliptic tori in Hamiltonian systems. See section 4.3 and [CJ00].
  Even if this procedure is adequate for many problems, it has the disadvantage that it scales badly with the number of discretization parameters. If we discretize the equation using N coefficients, the storage required for a full matrix is N^2 and the number of operations in the inversion of a matrix to accomplish one step of Newton method $O(N^3)$. We refer to this shortcoming as the *large matrix problem.*
  In slightly out of date desktop computers, the method works very well for N being a few hundreds, even about one thousand, but it quickly becomes impractical.
  The following algorithms overcome this large matrix problem, are faster, and use less storage. In section 5 we present an example of a problem which is very impractical for the large matrix method.
- *Projection method.* This method works for normally hyperbolic tori. The first step is to seek parameterizations of the stable and unstable bundles. In the case that they are 1D, this can be accomplished just by the linearized equation.
  Notice that to store the $\mathrm{D}F \circ K(\theta)$ by storing the matrix elements, we need only $O(N)$ storage. One step of application of $\mathrm{D}F$ requires just the multiplication of several functions. As we will see later, much of the work reported here is done with algorithms of multiplications which require $O(N^2)$ operations (after considering several issues we decided not to use multiplication algorithms in [Knu97] which are theoretically $O(N^{1+\delta})$ or even $O(N \log N)$). In any case, irrespective of which multiplication algorithm is used, the application of $\mathrm{D}F$ is significantly faster than the inversion of a matrix.
  Once we have the stable and unstable bundles, we can take the projection of the error into them and solve the resulting equations by iterations. The rate of convergence of the method depends on the size of the gap between the stable and unstable modes.
- *Reducibility method.* This method tries not only to solve (3) but also to find a constant

matrix Λ and a periodic matrix $P(\theta)$ such that

$$P(\theta+\omega)^{-1}M(\theta)P(\theta)-\Lambda=0, \tag{5}$$

where $M(\theta)=\mathrm{D}F(K(\theta),\theta)$. In this case, we say that the torus is reducible.
Note that having a solution of (5) makes it very easy to solve (4). If we substitute (5) into (4), it becomes

$$\Lambda P^{-1}(\theta)h(\theta)-P^{-1}(\theta+\omega)h(\theta+\omega)=-P^{-1}(\theta+\omega)r(\theta). \tag{6}$$

Equation (6) is an equation for $H(\theta)=P^{-1}(\theta)h(\theta)$, which, as we will see, is block-diagonal in a Fourier series; hence it can be solved very quickly if we discretize in Fourier series. The desired function h can be obtained by multiplying $h(\theta)=P(\theta)H(\theta)$. The storage of P, P^{-1}, $\mathrm{D}F$, etc. is only $O(N)$ and operations required are only linear in N and multiplications of Fourier series so that they are much faster than inverting full matrices.
For the sake of simplicity, assume that $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2)$ is diagonal. Then, denoting $R(\theta)=-P^{-1}(\theta+\omega)r(\theta)$ and denoting the Fourier coefficients by subindices separated by ; from the indices denoting the vector components, (6) becomes in Fourier terms

$$(\lambda_j-\exp(2\pi\mathbf{i}\,k\omega))H_{j;k}=R_{j;k}. \tag{7}$$

In summary, if we know P solving (5), we can reduce the solution of (4) to several multiplication of Fourier series and to a solution of (7), which is linear in the number of terms used in the discretization.
Equation (5) can be treated also via a Newton method, which as we will see can be given a fast implementation in terms of Fourier discretizations. We give some details. Given a P,Λ for which

$$P(\theta+\omega)^{-1}M(\theta)P(\theta)-\Lambda=S(\theta),$$

with S small, we seek improved solutions of the form $\hat{P}=P(\mathrm{Id}+Q)$, $\hat{\Lambda}=\Lambda+\Delta$, with $\Delta=\mathrm{diag}(\delta_1,\delta_2)$.
Expanding (5) up to first order in the corrections we obtain

$$\Lambda Q(\theta)-Q(\theta+\omega)\Lambda-\Delta=-S(\theta). \tag{8}$$

Equation (8) is diagonal in Fourier coefficients so that it can be solved in a time proportional to the number of coefficients used in the discretization. Equation (8) in Fourier coefficients amounts to

$$\begin{aligned}&(\lambda_i-\exp(2\pi\mathbf{i}\,k\omega)\lambda_j)Q_{i,j;k}=-S_{i,j;k}\quad\text{when } k\neq 0 \text{ or } i\neq j;\\ &Q_{i,i;0}=0,\ \delta_i=S_{i,i;0}.\end{aligned} \tag{9}$$

In practice, since inverting a Fourier matrix from scratch is time consuming, we keep an extra matrix $\bar{P}(\theta)$ which is supposed to be the inverse of the matrix $P(\theta)$, that is,

$$\bar{P}(\theta)P(\theta)-\mathrm{Id}=0. \tag{10}$$

When P changes, it is very easy to obtain the new $\bar{P}$ as a perturbation of the other one. Hence, we write a Newton method for $K, P, \bar{P}, \Lambda$ [HdlL05a].

In [HdlL06b] it is shown that if one uses a Newton step for both (3) and (5), one obtains a method which solves both equations with quadratic convergence. This requires $O(N)$ storage and the cost of one step of the method is comparable to the cost of one multiplication.

Remark 2.1. The theory has to consider several normalizations that take care of the fact that the solution of (5) is not unique. For example, multiplying P by a constant—or by any matrix that commutes with Λ—is also a solution.

Nevertheless, the algorithm described above converges to a solution of (5), which is all that is needed.

We note that the reducibility method also gives us very detailed information on the invariant subspaces.

We also note that we have found empirically that this method works well when λ_1, λ_2 are complex imaginary numbers of modulus 1 for which the small divisors in (9) and in (7) do not vanish except for k moderately large. This situation happens for elliptic tori in Hamiltonian systems. The empirical results are presented in section 4.3.

On the other hand, we note that it can happen that (5) may fail to have solutions even if the system is hyperbolic (and, therefore, the projection method works), so that the reducibility method, even if more effective, has a more reduced range of applicability. The phenomena that happen when (5) fails are quite interesting. We will study these failures in sections 3.2 and 6.

Our most efficient method to continue tori along a parameter is to use the reducibility method whenever possible but keep an eye on when (5) runs into trouble. Near these places, the method switches to the projection method or the large matrix method. Since the later methods require more effort, they may have to be run with fewer Fourier coefficients.

Remark 2.2. The condition that the coefficients of the unknowns that appear in (7) do not vanish is usually called the first Melnikov condition. The condition that the coefficients of the unknowns that appear in (9) do not vanish is usually called the second Melnikov condition.

Remark 2.3. We will see that, in this 2D context, reducibility is closely related to spectral properties of the transfer operator. In particular, it is closely related to the fact that the spectrum of the transfer operator consists of two circles. See section 2.2.5.

We will see that, in the dissipative context, there are quasi-periodic 2D mappings for which reducibility fails because of topological reasons. Similar examples happen in four-dimensional (4D) symplectic mappings. There are also intervals in which reducibility seems to be very hard to compute numerically. In these intervals, there seem to be hierarchies of subintervals in which the reducibility is harder and harder to compute. Moreover several topological and geometric properties seem to change in each level of the hierarchy. See section 6.3.4.

On the other hand, one can expect that reducibility is quite abundant among elliptic situations. In [dlLGJV05] it is shown that Lagrangian KAM tori are reducible, so that the algorithms derived here work also in this case with slight modifications.

2.1.2. Equations for invariant whiskers. In this paper we will be interested only in the study of whiskers which have one dimension more than the tori. We will also restrict ourselves

to the nonresonant case discussed in [HdlL06b]. This includes as particular cases the stable and unstable manifolds in a saddle-type torus and the fast stable manifolds. We also include the slow stable manifolds in case there is no resonance.

We will look for a parameterization for the torus given by an angle variable that measures the position with respect to the torus and another variable that moves in a direction roughly along the bundle. We will require that, in the chosen parameterization, we have that the angle variable moves by a rigid rotation and the normal variable contracts linearly. The fact that these normalizations can be done without loss of generality is discussed in [HdlL06b].

Hence, we will be looking for functions $W : \mathbb{T}^1 \times \mathbb{R}^1 \to \mathbb{R}^2$ in such a way that

$$F(W(s,\theta),\theta) = W(\lambda s, \theta + \omega), \tag{11}$$

where the unknowns are both W and λ, with $|\lambda| \neq 1$.

This equation can be adapted also to deal with nonorientable whiskers. As we will see, this generality is natural since in section 4.3 we will see that nonorientable bundles appear naturally in the bifurcations of elliptic tori. To deal with nonorientable bundles we just consider

$$F\left(W\left(s, \frac{1}{2}\theta\right), \theta\right) = W\left(\lambda s, \frac{1}{2}(\theta + \omega)\right).$$

Equation (11) will be discretized in Fourier–Taylor series. We will write

$$W(s,\theta) = \sum_{i=0} W_i(\theta) s^i$$

and match similar coefficients in s^i on both sides of (11).

We obtain

$$\begin{aligned} &F(W_0(\theta),\theta) = W_0(\theta+\omega),\\ &\mathrm{D}F(W_0(\theta),\theta)W_1(\theta) = \lambda W_1(\theta+\omega)\\ &\dots\\ &\mathrm{D}F(W_0(\theta),\theta)W_k(\theta) + S_k[W_0, W_1, \dots, W_{k-1}](\theta) = \lambda^k W_1(\theta+\omega), \end{aligned} \tag{12}$$

where S_k is an explicit polynomial in $W_0, W_1, \dots, W_{k-1}$ whose coefficients are derivatives of F evaluated at W_0. The polynomials S_k can be readily computed using the methods of *automatic differentiation* or *series manipulation*.

The first equation admits the solution $W_0 = K$, where K is the parameterization of an invariant circle. (That is, we are choosing an invariant circle to study.)

The second equation tells us that λ is an eigenvalue of the transfer operator defined by

$$\mathcal{M}_\omega V(\theta) = M(\theta - \omega) V(\theta - \omega) \tag{13}$$

and that W_1 is an eigenvector. This equation determines the space that we are choosing. This determines λ and determines W_1 up to a multiple. Even if all the choices of W_1 are mathematically equivalent (see [HdlL06b, HdlL05a]) and correspond to the nonuniqueness of

the solutions of (11), the choice affects the numerical properties of the algorithm. This will be discussed in more detail in section 2.1.5.

The other equations are quite straightforward. Note that, proceeding by induction, we can assume that $W_0, \ldots, W_{k-1}$ are known. The resulting equations for W_k can be solved provided that λ^k is not an eigenvalue of the transfer operator, which is the main assumption in the theorems about existence of nonresonant whiskers in [HdlL06b]. See also [CFdlL03a, dlL97].

We note that the nonresonance assumption above is automatically satisfied for the strong stable manifolds (resp., the strong unstable manifolds) of an attractive (resp., repelling) torus or by the stable and unstable manifolds of a saddle-type torus. In [HdlL06b] it is shown that all the classical manifolds just mentioned can be obtained through a parameterization as above. The only thing to check is that given a classical manifold, there are some coordinates in which the motion can be made to be a multiplication by a number. In [HdlL06b] we also consider cases in which reduction to a multiplication is not possible, but reduction to a polynomial dynamics is possible.

2.1.3. Fourier–Taylor series. Since the components of the parameterization of a 1D torus are periodic functions of one angle variable θ, it is natural to consider Fourier series [dlLT94, CJ00]. We store the periodic functions of one variable as cosine and sine Fourier series. If $f(\theta)$ is a 1-periodic function, we will write

$$f(\theta) = a_0 + \sum_{k>0} \left(a_k \cos(2\pi k\theta) + b_k \sin(2\pi k\theta)\right).$$

The finite computer memory forces us to cut off these expansions. The objects that we compute are smooth, or even analytic; hence, the coefficients in these expansions decrease eventually faster than powers or even faster than exponentially. Of course, when we move parameters so that the torus is close to breakdown, this eventually fast convergence will only be apparent for the coefficients of very high order. In this paper we also present an example of a Lipschitz map whose invariant torus is only Lipschitz, so the Fourier coefficients of its parameterization decrease very slowly.

The norm we have used in order to estimate the errors in the computations is

$$\|f\| = |a_0| + \sum_{k>0} \sqrt{a_k^2 + b_k^2} \geq \|f\|_\infty.$$

The numbers $c_k = \sqrt{a_k^2 + b_k^2}$ are called the amplitudes of the Fourier modes. For a periodic matrix $M(\theta) = (M_{i,j}(\theta))$, we denote $\|M\| = \max_{i,j} \|M_{i,j}\|$.

For the computation of whiskers of invariant tori, we use Fourier–Taylor expansions of the form

$$f(s, \theta) = \sum_{m=0}^{\infty} f_m(\theta) s^m,$$

with

$$f_m(\theta) = a_{m,0} + \sum_{k>0} \left(a_{m,k} \cos(2\pi k\theta) + b_{m,k} \sin(2\pi k\theta)\right).$$

We have used a package of routines that perform operations on series as above. Certainly, it is elementary to implement the elementary algebraic functions (sum, product, and product by scalar), since there are ready formulas for the coefficients of the result in terms of the coefficients of each of the terms. It is not too hard to implement the division of series and elementary transcendental functions (exp, sin, cos, log, power, etc.). Some hints on how to implement the elementary transcendental functions are given in [Knu97]. These techniques are known under the name of automatic differentiation algorithms. Similar packages have been implemented many times in the literature [GC91, RJB83, MS91, Koc99, Har02].

Remark 2.4. The most straightforward algorithm for multiplication of Fourier series is the use of the Cauchy formula. If we implement it in the most straightforward way, evaluating the product of two series of N coefficients requires $O(N^2)$ operations to complete.

In [Knu97], one can find a discussion of algorithms that can be asymptotically faster. The asymptotically fastest is to compute the FFT, multiply, and then take the FFT back. This would give an asymptotic cost of $O(N \log(N))$. Preliminary testing suggests that these different methods have different numerical properties.

In this package we have used always the straightforward Cauchy formula.

It would be quite natural to implement the Fourier manipulation routines using the BLAS library. We have not done so, but rather have used explicit loops. This has the advantage that the precision of the programs can be changed from `float` to `double` to `long double` very easily.

Using these routines, one can implement the operators involved in (3), (11), and their derivatives provided that the function F can be written using algebraic operations and the elementary transcendental functions.

If the map is not given explicitly (for instance, if it is given by the flow of a vector field) or its computation is hard, we could follow the following strategy [CJ00]: take a mesh of points on the torus, apply the map to the mesh of points, and compute the torus corresponding to the image of the new mesh by applying FFT methods. To compute first derivatives (the first order terms in the Taylor expansions), one has to integrate the first order variational equations of the flow [CJ00]. For higher order derivatives, one has to integrate higher order variational equations.

For flows, one can also use functional equations to determine the invariant manifolds. The invariance equation of a torus is a first order linear PDE [DLR91, DL95, MKM97, ERS00, Tru00]. See also [HdlL06b] for the invariance equations for invariant tori and their whiskers in quasi-periodic differential equations.

As the examples that we have studied here are maps that involve only a few operations in this paper we have used only the Fourier–Taylor methods. We think that a detailed comparison between the FFT method and the Fourier–Taylor method would be quite interesting, and we plan to come back to this issue. Preliminary studies indicate that the numerical properties of both methods are quite different, especially if one looks at properties which involve derivatives.

We have not undertaken a comparison of the Fourier–Taylor method with other methods of discretization (e.g., splines). On the other hand, we note the following.

- As will be seen in the section on numerical results, we will routinely obtain residuals which are of the order 10^{-15}–10^{-20} using only 100 terms or so (and the error is possibly an overestimate). Achieving a similar accuracy would seem to require a method of

interpolation of rather high degree whose stability properties would be problematic.

- Many of the equations we have to solve in a Newton step are diagonal—or quasi-diagonal—when discretized in Fourier series. In particular, the reducibility method, when discretized in Fourier series, makes the computational effort of a Newton step essentially linear in the number of terms.
- Fourier methods have the disadvantage that they are not adaptive and it is hard to increase the accuracy in the places where the manifold becomes more oscillatory. In the present examples, the fact that the dynamics on the torus is a rigid rotation makes the torus very homogeneous, so that there is little need for adaptivity. Our implementation can increase or decrease the number of Fourier terms used dynamically.

The above reasons, especially the last two, are very closely tied to the models that we are considering in this paper. It is quite possible that, when applying the parameterization method to other problems, other methods could prove to be more convenient.

2.1.4. Assessment of the error of the calculation of invariant tori. We have introduced several measurements that give an idea of what accuracy is achieved by our calculations of invariant tori and their invariant bundles.

Given $K, P, \bar{P}, \Lambda$ approximate solutions of the equations of the invariance equation (3), the reducibility equation (5), and the inverse equation (10), we give names to the remainders as follows:

$$\begin{aligned} R(\theta) &= \bar{P}(\theta+\omega)(F(K(\theta),\theta) - K(\theta+\omega)), \\ S(\theta) &= \bar{P}(\theta+\omega)\mathrm{D}F(K(\theta),\theta)P(\theta) - \Lambda(\theta), \\ T(\theta) &= \bar{P}(\theta)P(\theta) - \mathrm{Id}. \end{aligned}$$

The quality of the approximate solutions is given by the norms of these Fourier matrices, as defined in section 2.1.3.

Remark 2.5. If we apply the Newton method, then we take $P = \bar{P} = \mathrm{Id}$ and $\Lambda(\theta) = \mathrm{D}F(K(\theta),\theta)$, so $S = 0$ and $T = 0$.

2.1.5. Domains of the parameterization of the whiskers. We have explained in [HdlL06b, HdlL05a] how to compute the Fourier–Taylor expansions of the parameterizations of the whiskers. Given such an expansion, we have to obtain a domain in which such approximations are good enough to allow reliable evaluations.

In principle, the parameterization gives a complete description of the manifold and the parameterization can follow the twists and turns of the manifold.

In practice, however, the numerical evaluation could become numerically unstable at larger values of s. Hence, it is sometimes more advantageous to evaluate W only for values of s smaller than a certain s_0 and then use the functional equation (11).

We refer to this last step as *globalizing the manifold*. This has been always an ingredient of algorithms to compute invariant manifolds since all the algorithms give only some local piece. See, for example, [KO98]. We note that, in our case, since the parameterization covers, in principle, the whole manifold, the globalization is to a large extent a matter of numerical convenience and that even without performing any extension, the parameterization can follow a large number of turns in the manifold.

Remark 2.6. We note that if $W(s,\theta)$ is a solution of (11), so is $\tilde{W}(s,\theta) = W(r \cdot s, \theta)$ for any $r \in \mathbb{R}$. The choice of r is closely related to the choice of the first coefficient in the expansion. Hence, the choice of the multiple in the first step is related to the domain. From the mathematical point of view, it is equivalent to choosing a larger W_1 and then evaluating in a proportionally small domain. Nevertheless, from the numerical point of view, both methods are inequivalent. If we have coefficients of different sizes, they are much more prone to round-off error. Hence, a rule of thumb is to choose W_1 so that the coefficients computed have comparable sizes. In practice, one can do a preliminary run that gives some idea of how the coefficients grow. Since choosing ρW_1 instead of W_1 makes W_k become $\rho^k W_k$, it is easy to choose ρ so that the coefficients W_k have a comparable size. Similar adjustments are common also in Lindstedt series [FdlL92].

Remark 2.7. The computation of slow manifolds [CFdlL03a, CFdlL03b, CFdlL05] deserves a few words. Notice that the dynamics around an attracting invariant circle is dominated by the normal directions associated to the "eigenvalue" with largest modulus. The situation is analogous to the situation in linear maps. The components along the most contractive eigenvalues disappear faster than those along the slow eigenvalue and the component along the slow eigenvalue is the dominant one for long term behavior. Hence, from the point of view of asymptotic behavior, the manifold associated to the spectral values closest to the unit circle is the most observable. In [HdlL06b] it is shown that under some appropriate nonresonance conditions, the nonresonant manifold exists and is unique among the tangent manifolds of a degree of regularity that depends only on the spectrum. Then, this unique manifold is as smooth as the map. Hence, as soon as we know that there is one invariant manifold with a moderate regularity, we can bootstrap the regularity to the regularity of the map. These results are sharp because examples in [HdlL06b] show that there are examples with resonances with no smooth invariant manifolds. Moreover, there may be many invariant manifolds with low regularity.

In the 2D set-up of this paper, assume the transfer operator associated to the invariant circle has a spectrum that is two circles of radii $\rho_- < \rho_+ < 1$ (see section 2.2). The slow manifold is tangent to the spectral subbundle associated to the circle of radius ρ_+. The nonresonance conditions in [HdlL06b] in our simpler case just amount to $\rho_+^i \neq \rho_-$ for all $i = 2, \ldots, L$, with L such that $\rho_+^{L+1}\rho_- < 1$.

If we use only a first order approximation of the whisker, the domain where it is a good approximation can be very small and the globalization is harder. If we are computing the fast manifold, the iteration backward aligns the point with the fast manifold and reduces the error, so that it is not so crucial. On the other hand, iterating backward makes the slow manifold unstable. Hence, to obtain global slow manifolds, it is important to obtain high accuracy which can be best obtained by computing the expansion to a rather high order.

2.2. Analysis of the dynamics of linearization. In the analysis of an invariant torus, it is quite natural to study the dynamics of the linearization. As was discussed in great generality in [HdlL06b], the study of the linearization is very closely related to the properties of the linearization of the invariance equations. It is important to realize that the invariance equations (3) are functional equations and that therefore one should think of them as equations in a Banach space. One of the features of normal hyperbolicity theory is that there is a close

connection between the geometric properties of the variational equations and the functional analysis properties of (3). This connection was realized in [Mat68, Sac78]. A treatment tailored for the examples appearing in this paper is in [HdlL03b].

At the geometric level, we observe that the linearized dynamics is given by

$$\begin{aligned} \bar{v} &= M(\theta) v, \\ \bar{\theta} &= \theta + \omega, \end{aligned} \tag{14}$$

where $v = (v_x, v_y) \in \mathbb{R}^2$, and $M(\theta)$ a 2×2 periodic matrix. In our applications, $M(\theta) = \mathrm{D}F(K(\theta), \theta)$.

We use the notation

$$\begin{aligned} M(\theta, m) &= M(\theta + (m-1)\omega) \dots M(\theta), \\ M(\theta, -m) &= M(\theta - m\omega)^{-1} \dots M(\theta - \omega)^{-1} \end{aligned}$$

for $m > 0$ and $M(\theta, 0) = \mathrm{Id}$.

On the functional analysis side, the cocycle (14) induces a transfer operator $\mathcal{M}_\omega$ defined in (13). We see that inverting $\mathcal{M}_\omega - \mathrm{Id}$ is precisely solving the Newton step in (4). Hence, it is quite important for the algorithms to decide whether 1 is in the spectrum or not.

In the subsequent sections, we will develop certain numerical observables that are relevant for the analysis of the invariant torus and which are readily computable and which have an influence on the behavior of the linearized dynamics and on the solvability of the functional equations for the parameterizations we are interested in.

Other issues, such as reducibility of the cocycle, have been discussed above (see also [Jor01, HdlL05a]).

In section 2.2.5, we will discuss the relation between the geometric properties and the spectrum of the linearization of the functional equations we are interested in.

2.2.1. Projectivization of the cocycle. To describe the 1D linear subspaces we are considering, it is natural to give a set of directions at every point. That is, to parameterize an invariant bundle, we give a mapping from the torus to the 1D projective space $\mathbb{P}^1$ which we represent as an angle in $[0, \pi)$ with $0, \pi$ identified.

Corresponding to the linear transformation (14), we can consider its action on directions given by the *"projective cocycle"*

$$\begin{aligned} \bar{\alpha} &= m(\alpha, \theta), \\ \bar{\theta} &= \theta + \omega, \end{aligned} \tag{15}$$

where $\bar{\alpha} = m(\alpha, \theta)$ is computed as follows:

- Given $\alpha \in [0, \pi[$, we consider

$$v = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}.$$

- We apply the cocycle, to obtain

$$\bar{v} = M(\theta) v = |M(\theta) v| \begin{pmatrix} \cos \tilde{\alpha} \\ \sin \tilde{\alpha} \end{pmatrix},$$

where $\tilde{\alpha} \in [0, 2\pi[$ is an angle in the unit circle ($0 = 2\pi$).

- Finally, if $\tilde{\alpha} \in [0, \pi[$, we define $\bar{\alpha} = \tilde{\alpha}$, and if $\tilde{\alpha} \in [\pi, 2\pi[$, we define $\bar{\alpha} = \tilde{\alpha} - \pi$.

Notice that a subbundle in $\mathbb{R}^2 \times \mathbb{T}$ is projectivized into a graph of a function $\theta \to \alpha(\theta)$ in $\mathbb{P}^1 \times \mathbb{T}$. It is also very easy from this graph to know if the subbundle is orientable or not, and what is its index. The index of a subbundle E is a semi-integer number $\frac{k}{2}$ that gives the number of turns that gives the fiber E_θ when θ goes from 0 to 1. If k is even, the subbundle is orientable, and if k is odd, it is nonorientable. Doubling the period of θ is equivalent to taking a double cover so that we can orientate any 1D bundle.

In the case that the linear cocycle (14) has an exponential dichotomy, the projective cocycle (15) will have an attractor and a repellor. The attractor will be a representation of the unstable bundle (in the case of a hyperbolic torus) or the slow bundle (in the case of an attractive torus). We denote this bundle E^+. The repellor (i.e., the attractor under inverse iteration) will correspond to the stable bundle (in the case of a hyperbolic torus) or the fast bundle (in the case of an attractive torus). We denote this bundle E^-.

Remark 2.8. Note that the only bundles we are considering are continuous bundles. According to Oseledec's theorem [Ose68, Pes77], it could be natural to consider also measurable subbundles, especially in situations where there is bundle collapse (see section 6).

Nevertheless, we note that the motion on the circle admits only one invariant measure (the usual Lebesgue measure). Therefore, Oseledec's theorem produces only one set of Lyapunov multipliers and the Oseledec bundles can only be discontinuous in a set of measure zero (if this set is not empty, it is dense). It could happen that this is what is observed in section 6.

The theory of reduction in the measurable category for 2D cocycles is studied in [Thi97].

2.2.2. Lyapunov multipliers. The maximal and minimal Lyapunov multipliers λ_+ and λ_- are computed by taking a random vector $v \in \mathbb{R}^2$ and a random angle $\theta \in \mathbb{T}$ and applying the formula

$$\lambda_+ = \lim_{m\to\infty} |M(\theta, m)v|^{\frac{1}{m}}, \ \lambda_- = \lim_{m\to\infty} |M(\theta, -m)v|^{\frac{1}{-m}}.$$

To avoid the growth of the vectors, we do, as usual, scalings at each step. To compute the accuracy, we just compute the Lyapunov multipliers at the convergents of the continued fraction of ω and check that there is no appreciable change. An estimate of the error is the difference of two consecutive estimates.

We note that when the bundles are 1D, we have that $\gamma(\theta, m) = \prod_{i=0}^{m-1} \gamma(\theta + i\omega)$, where $\gamma(\theta)$ is the—scalar—block of the matrix M in the invariant bundle. We have that the corresponding Lyapunov multiplier is given by

$$\begin{aligned} \lambda &= \lim_{m\to\infty} \exp\left[\frac{1}{m}\sum_{i=0}^{m-1} \log|\gamma(\theta + i\omega)|\right] \\ &= \exp\left[\int_0^1 \log(|\gamma(\theta)|)\, d\theta\right]. \end{aligned} \tag{16}$$

Furthermore, since in many cases the system is analytic, then the 1D bundles are also analytic, and the exponents for finite m converge to their limit (16) faster than any power in

m. (Of course, as the analyticity properties become weaker, the exponential convergence may take longer to manifest itself.)

The exponential convergence of the Lyapunov multipliers to their limiting value in the uniformly hyperbolic case can be used to speed the calculation or increase the accuracy.

One could consider also the use of quadrature algorithms to evaluate (16) (this is useful to make checks).

2.2.3. Separation of the bundles. An observable that will play an important role in our considerations (see, in particular, section 6) is the distance between two bundles in a exponential dichotomy. We will consider

$$\Delta = \min_{\theta} \operatorname{dist}(E^+_\theta, E^-_\theta); \tag{17}$$

this is the minimum angle between the bundles $E^{\pm}$.

As is well known, the distance between the stable and unstable bundles plays a very important role in hyperbolicity theory. See [Fen72, HPS77]. The fact that to measure hyperbolicity one cannot just use the Lyapunov multipliers but has to use also the distance of the splittings is emphasized in [CL00].

2.2.4. Rotation numbers. When considering normally elliptic tori in Hamiltonian systems, it will be useful to develop fast diagnostics that detect empirically the lack of reducibility. Besides the Lyapunov multipliers and the distance between invariant bundles, the observables that we will use are roughly patterned on the rotation number which measures the averaged amount of rotation per turn.

We note that, when the cocycle is reducible, the cocycle is a rotation modulo a change of coordinates. Therefore, there are many ways to compute the rotation number. There are several definitions of rotations numbers in the mathematical literature. See [Rue85, NNO98, FJN03a, FJN03b] for a review.

Unfortunately, the average amount of rotation requires us to consider the angles not modulo 2π but rather on a lift of the circle. When iterating maps, the information needed to lift the angles seems to involve global considerations (e.g., choosing an origin for the angles and propagating this choice along the circle, e.g., by interpolating the map by a globally defined family). We have not succeeded in implementing any of these rigorous definitions in a numerically efficient manner because the global considerations required by the definitions are hard to implement efficiently.

Since our purpose is just to get some rough understanding of reducibility, we have used several diagnostics that work well in the reducible case. When they start to show changes or erratic behavior, this is a sign that reducibility has broken down. The thresholds computed in this way can subsequently be refined.

The first method is to compute the rotation number γ by the formula

$$\gamma_a = \lim_{m\to\infty} \frac{1}{m} \sum_{i=0}^{m-1} \operatorname{ang}(M(\theta, i)v, M(\theta, i+1)v), \tag{18}$$

where $\operatorname{ang}(v, w)$ denotes the oriented angle from v to w, from $-\pi$ to π.

An easy way to compute the *Sturmian* rotation number inspired by [DS83b] (see also [FJN03a]) is defined as follows. If $v^i = (v_x^i, v_y^i)$ denotes $M(\theta, i)v$, for $m > 0$ we define $N(m)$ as the number of times in which v_y^i changes sign for $i = 0 \div m$. Then we define the rotation number as

$$\gamma_s = \pm\pi \lim_{m\to\infty} \frac{N(m)}{m}. \tag{19}$$

In order to specify a sign for the rotation number, we can see if, "on average," the vector spins counterclockwise (sign $+1$) or clockwise (sign -1).

We note that both definitions, even if they make sense for differential equations or small steps, may run into trouble in some cases. For example, if we were taking long steps in a differential equation, the angle could rotate by more than a turn. Similarly, the Sturmian values could miss several changes of sign. In the reducible case, these mistakes happen in a very systematic way. On the other hand, in the nonreducible case, it is not possible to say that they will occur with a fixed frequency. For us, the above observables are just quick diagnostics for resonances and lack of reducibility.

2.2.5. Spectral properties of the transfer operator. The transfer operators that appear in our case are very special for several reasons: The motion in the base is a rotation and the nontrivial bundles are 1D; hence the restriction of the derivative is commutative.

One of the consequences of the theory developed in [HdlL03b] is that the spectrum is largely independent of the spaces considered so that we will not specify the space we are considering the transfer operator acting on.

Putting together results from [HdlL03b], we obtain that there are only a few possibilities for the spectrum of the transfer operator $\mathcal{M}_\omega$ introduced in (13):

(a) *Two circles of radii* $\rho_- < \rho_+$. In this case, we have a decomposition into continuous invariant bundles $\mathbb{R}^2 = E_\theta^+ \oplus E_\theta^-$ for all $\theta \in \mathbb{T}$, characterized by the uniform rates of growth

$$\begin{aligned} v \in E_\theta^- \quad &\Leftrightarrow \quad |M(\theta, m)v| \le C(\rho_- + \delta)^{+m}|v|, \ m \ge 0, \\ v \in E_\theta^+ \quad &\Leftrightarrow \quad |M(\theta, -m)v| \le C(\rho_+ - \delta)^{-m}|v|, \ m \ge 0, \end{aligned} \tag{20}$$

where δ is small enough, and the constant C is uniform (it does not depend on θ or v). The vectors in $E_\theta^\pm$ have a Lyapunov multiplier equal to $\lambda_\pm = \rho_\pm$. We say that the cocycle M has an exponential dichotomy. In particular, if $\rho_- < 1 < \rho_+$, the cocycle is uniformly hyperbolic and of saddle type, and E^- and E^+ are the stable and the unstable subbundles; if $\rho_- < \rho_+ < 1$, the cocycle is uniformly hyperbolic and of attracting node type, and E^+ and E^- are the fast stable and the slow stable subbundles.

The subbundles $E_\theta^\pm$ are as smooth as M. Moreover, if the external rotation is Diophantine, the dynamics on each bundle can be reduced to a multiplication by a constant [JS81].

(b) *One circle of radius* ρ. All the vectors have Lyapunov multipliers equal to ρ. This case appears naturally in the elliptic case in Hamiltonian systems, with $\rho = 1$.

(c) *One annulus of radii* $\rho_- < \rho_+$. In this case, there is no continuous invariant splitting but, according to [Ose68, Pes77], there is a measurable splitting characterized by the fact that the corresponding Lyapunov multipliers are $\lambda_\pm = \rho_\pm$ (see [CL99] and the references therein). If $\rho_- < 1 < \rho_+$, the cocycle is nonuniformly hyperbolic.
The forward and backward Lyapunov multipliers of vectors lie between ρ_- and ρ_+ [Sac78, HdlL03a], but we emphasize that the maximal and minimal Lyapunov multipliers $\lambda_\pm = \rho_\pm$ are attained in a set of full Lebesgue measure.
Moreover, for all $\rho \in [\rho_1, \rho_2]$ there is a vector $v \in \mathbb{R}^2$ and an angle $\theta \in \mathbb{T}$ whose scaled orbit $\{\frac{1}{\rho^m} M(\theta, m)v\}_{m\in\mathbb{Z}}$ is bounded [Mn78, CL99, HdlL03a].

Remark 2.9. If a cocycle is reducible (or almost reducible), then the spectrum of the transfer operator is one of the alternatives (a) and (b) above [HdlL03b]; that is, the spectrum is either two circles or one circle.

Almost conversely, in our case, when we are in case (a), the cocycle will be reducible.

2.2.6. Rotating transformations. Assume the cocycle (14) is reducible to a matrix $\Lambda = \rho R_\gamma$, with R_γ being a rotation matrix of angle γ. This is that there exists a periodic matrix $P(\theta)$, of period 1 (or 2), such that

$$M(\theta)P(\theta) = \rho P(\theta + \omega)R_\gamma. \tag{21}$$

For $k \in \mathbb{Z}$, consider the rotating matrix $R_k(\theta) = R_{\pi k\theta}$. For k even, this is of period 1, but for k odd, this is of period 2. Then, it is easy to see that the matrix $P_k(\theta) = P(\theta)R_k(\theta)$ satisfies

$$M(\theta)P_k(\theta) = \rho P_k(\theta + \omega)R_{\gamma-\pi k\omega}. \tag{22}$$

So, the cocycle is also reducible to $\Lambda_k = \rho R_{\gamma-\pi k\omega}$.

This freedom in choosing the reduced matrix is very useful in numerical computations. For instance, assume that the torus has a cocycle that is reduced to a matrix ρR_γ. When perturbing the system, the rotation number γ crosses resonances. Assume we are near a resonance $k_1 : k_2$, that is,

$$\left| \frac{\gamma}{2\pi} - \frac{k_1 + k_2\omega}{2} \right| \simeq 0. \tag{23}$$

Then, the rotating transformation R_{k_2} lets us reduce the system to the matrix

$$\Lambda_{k_2} = \rho R_{\gamma-\pi k_2\omega} \simeq \rho R_{k_1\pi}.$$

So, the matrix Λ_{k_2} is close to a diagonal matrix with the same entries in the diagonal.

It is obvious that k_1 gives the sign of the elements of the diagonal. The number k_2 is related to the topology (the index) of the dynamics around the torus. The transformation R_{k_2} straightens the dynamics of the cocycle but keeps the topological information. The index of R_{k_2} is $\frac{k_2}{2}$. We will clarify these ideas with some examples later.

2.3. Computer programs. The algorithms have been implemented in the C language and have been run under the Linux environment. We have applied them to compute invariant manifolds for 2D quasi-periodic maps, with one external frequency.

The algorithms have been applied to quasi-periodic perturbations of the Hénon map and the standard map (although with slight modifications they can be applied to other examples). In the examples the system is coupled with a quasi-periodic external perturbation via a parameter ε.

With this set of programs we can do the following:

1. *Compute invariant and periodic circles*, with continuation with respect to the parameter ε. The input is the set of parameters of the autonomous system (a and b in the case of the Hénon map, and K in the case of the standard map), the period with some specification of the initial periodic point (in the case of the standard map, for instance, that is its rotation number and some symmetry), and a list of ε's; the output is a set of files containing the Fourier expansions of the circles for these ε's. Other programs let us obtain meshes of the circles to make plots. In the implementation of the projection method and the reducibility method, the programs include in the output files the reducing transformations and the reduced matrices.
2. *Compute whiskers of invariant and periodic circles*: The input is one of the outputs of the previous programs, that is, a file containing the parameters and the Fourier expansions giving the circle; the output is, in the case that the circle is hyperbolic, the files containing the Fourier–Taylor expansions of the whiskers. If in the input file there are the reducing frames (because we have used the projection method or the reducibility method), these programs use them to do the computations. If not, the programs implement the power method to compute the bundles, and we refine the computation using reducibility [HdlL05a] (we have also done tests and computations by discretizing the transfer operator and computing the eigenvalues of the discretization [Jor01]). These bundles are the frames where the whiskers are constructed.
3. *Globalize the whiskers*: A first program computes the intersection of a whisker with a given plane $\theta = \theta_0$. A second program computes a 2D grid of points of the surface, by doing several sections of the object with equidistant time planes. In both cases we have to specify the total arc-length of the sections, the maximum distance between consecutive points in the curves, and the local error to compute the fundamental domain. This is very standard (see, for instance, [KO98]).
4. Once we have computed the meshes and grids of the objects, we can plot them using standard packages. To produce the figures of this paper we have used `xmgr`, `gnuplot`, and MATLAB.

We have also written many other test programs—for example, computation of Lyapunov multipliers and rotation numbers of a cocycle, computation of attractor and repellor of a projective cocycle, and power method.

We have also written the versions of the programs to work in the double cover. This is just substituting the system (2) by

$$\begin{pmatrix} z \\ \theta \end{pmatrix} \longrightarrow \begin{pmatrix} F_\varepsilon(z, 2\theta) \\ \theta + \frac{\omega}{2} \end{pmatrix}. \tag{24}$$

This trick lets us work with 1-periodic functions and avoid 2-periodic functions.

The timings reported correspond to work in different machines that we will specify. For instance, some results have been produced with a desktop computer with a 700MHz processor. From this, the reader can guess timings in other machines.

In the following we will discuss some efficiency properties of the implementation in some representative examples and, much more importantly, we will present some conjectures that we have obtained by observing phenomena that are not yet justified by a full mathematical theory. We hope that reporting these conjectures will stimulate mathematical work.

3. Example 1: The rotating Hénon map. The first example we consider is a quasi-periodic perturbation of the Hénon map. The *rotating Hénon map* is the quasi-periodic dissipative map on the plane given by

$$\begin{aligned} \bar{x} &= 1 + y - a\, x^2 + \varepsilon \cos(2\pi\theta), \\ \bar{y} &= bx, \\ \bar{\theta} &= \theta + \omega \qquad (\mathrm{mod}\ 1), \end{aligned} \tag{25}$$

where a, b are the parameters of the Hénon map and ε leads to a quasi-periodic forcing. The motion of the angle variable θ is a rotation by the irrational number ω. We have taken the golden mean $\omega = \frac{1}{2}(\sqrt{5} - 1)$.

For $\varepsilon = 0$, and we suppress the angle variable, we have the classical Hénon map

$$\begin{aligned} \bar{x} &= 1 + y - a\, x^2, \\ \bar{y} &= bx. \end{aligned} \tag{26}$$

This map has two fixed points

$$p_\pm = \left(x_\pm = \frac{b - 1 \pm \sqrt{(b-1)^2 + 4a}}{2a},\ y_\pm = bx_\pm \right),$$

provided that $(b-1)^2 + 4a > 0$ (for instance, if $a > 0$). There exists also a 2-periodic orbit, provided that $4a - 3(1-b)^2 > 0$ (at $a = \frac{3}{4}(1-b)^2$ there is a period doubling bifurcation of p_+). Both points of the 2-periodic orbit correspond to the solutions of the quadratic equation

$$a^2x^2 - a(1-b)x - a + (1-b)^2 = 0.$$

These fixed points and periodic orbits turn into invariant curves and periodic curves under the quasi-periodic perturbation. (See Figure 1.)

The map (25) was studied in [SFKA96] as an example of system with a strange non-chaotic attractor (SNA). In [OF00] the structure of this attractor was studied by means of computation of stable and unstable manifolds of invariant curves of saddle type, based on the algorithms appearing in [KO98]. The computation of slow manifolds considered here does not seem to have been considered in the literature before.

In the following, we fix the parameters of the Hénon map to be $a = 0.68$ and $b = 0.1$. We have used several computer programs to continue invariant and periodic tori with respect to ε and to compute their whiskers.

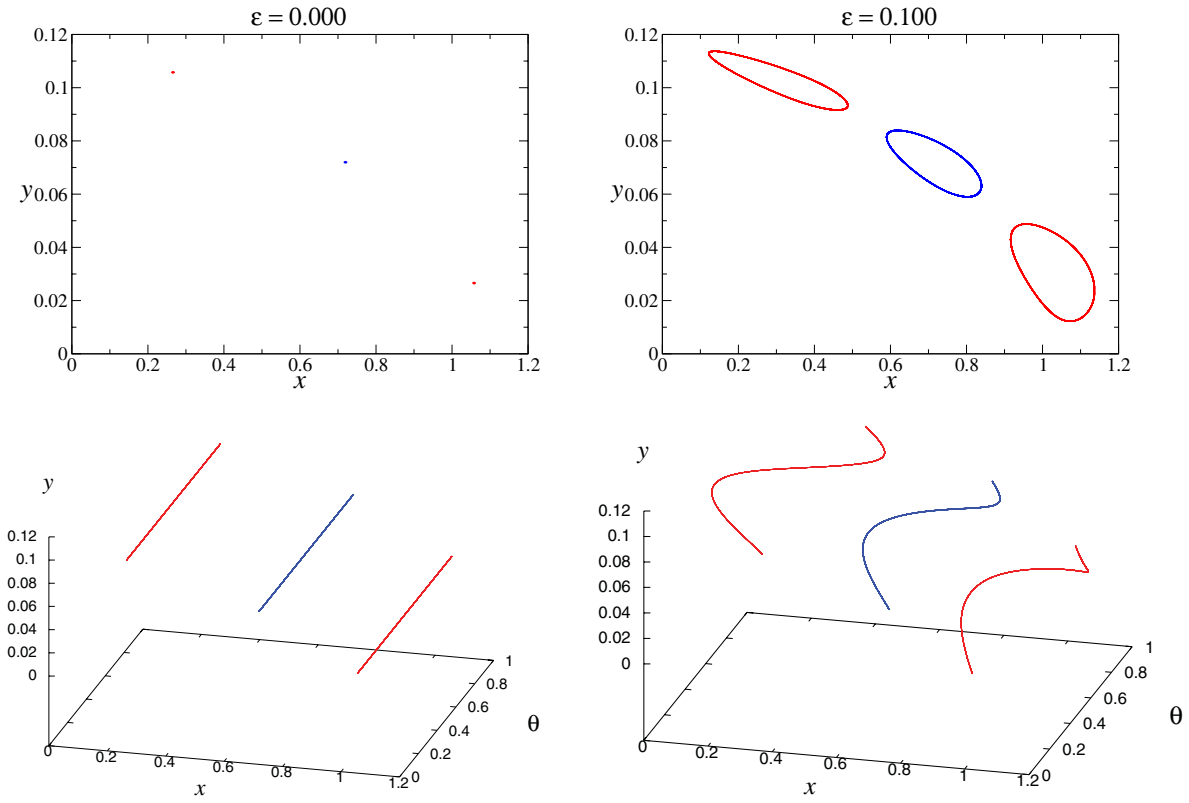


Figure 1. *A saddle-type fixed point (blue) and an attracting 2 periodic point (red) of the Hénon map with $a = 0.68$ and $b = 0.1$ become closed curves when we add the external perturbation. The lower pictures are 3D views in the extended phase space.*

Starting from the fixed point p_+, the continuation of the invariant curve for $\varepsilon = 0., 0.1, 0.2, 0.3, 0.4, 0.5$ is done in less than one second using 50 harmonics in the Fourier expansions and running on a Pentium III at 700 MHz. The errors in the functional equations are of the order 10^{-19}. (See Figure 1 for some examples.)

For $\varepsilon = 0.1$, the invariant curve is of saddle type. We can also compute the whiskers of this invariant curve up to order 30 (the degree of the Fourier–Taylor polynomials), in less than 8 seconds. If the linear approximations of the manifolds are not provided as an input, the program uses the power method to find them. We use these local expansions to globalize the manifolds. For instance, to compute points of the curve of intersection with the plane $\theta = 0$ of one of the branches of the unstable manifold (length = 10, distance between points = 0.01, local error = 10^{-10}) takes about 4 seconds. With this technology we can also produce 3D grids (taking essentially different sections).

Figure 2 shows the section with the plane $\{\theta = 0\}$ of some invariant curves and their whiskers for $\varepsilon = 0.1$.

3.1. Continuation of an invariant torus: A brief description. In the following sections we will explain the bifurcations and transitions we have encountered in continuing the invariant

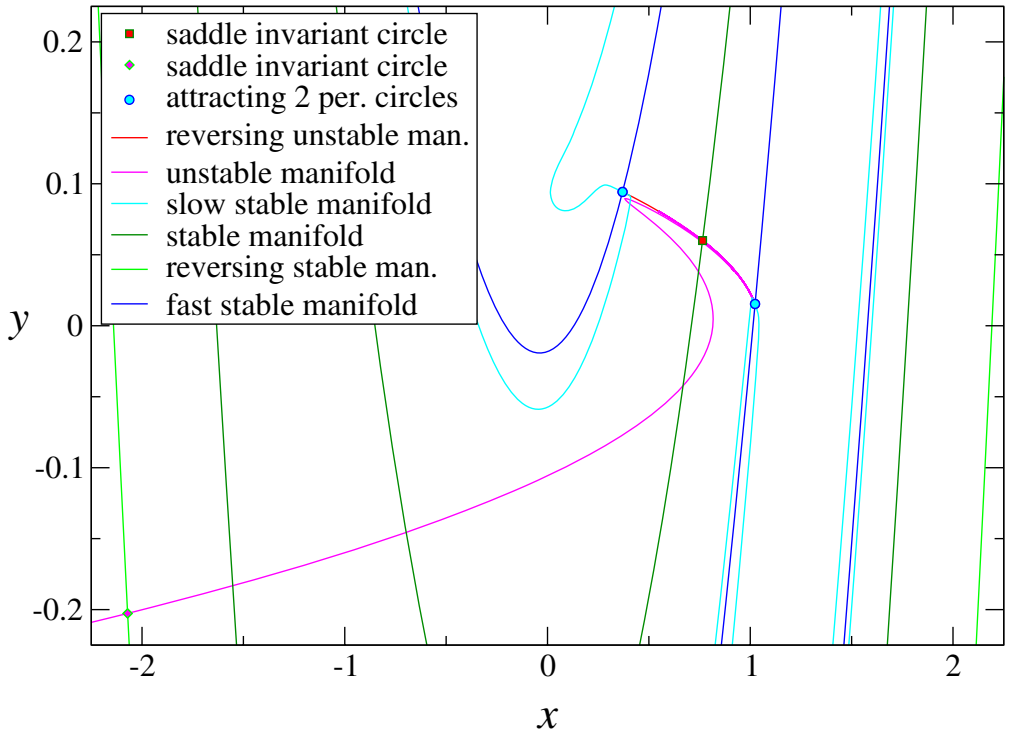


Figure 2. *Sections with the plane $\theta = 0$ of invariant 2D manifolds associated to invariant curves. The parameters of the rotating Henón map are $a = 0.68$, $b = 0.1$, and $\varepsilon = 0.1$.*

circle corresponding to the fixed point p_+ of the Hénon map. In particular, in section 6, we will report on a mechanism of loss of exponential dichotomy and breakdown.

In the following exploration, which is representative of many others, we have fixed $a = 0.68$ and $b = 0.1$ and have increased ε from 0. Figure 3 shows the continuation of the invariant torus, whose analysis is detailed below.

To measure the hyperbolicity around the torus, we have used two observables: the maximal Lyapunov multiplier $\Lambda = \lambda_+$ (the minimal Lyapunov multiplier is $\lambda_- = |b|/\Lambda$) and the distance Δ between the attractor and repellor of the projective cocycle. The results are displayed in Figure 4, where we observe several transitions, labeled with the letters `a,b,c,d,e`.

The fixed point p_+ of the Hénon map is of saddle type. When we take the Cartesian product with the external rotation, the fixed point becomes a "straight" invariant circle (see the first picture of Figure 3, where we also draw the 2-periodic attracting circle).

A bird's eye analysis of the bifurcations detected easily in Figure 4 is as follows.

`(a)` At $\varepsilon_{\mathtt{a}} \simeq 0.2549$, the spectrum of the linearization crosses 1 and -1 is an eigenvalue. The torus goes from being unstable to being stable, and there is a period doubling bifurcation [BHTB90]. The leading spectral bundle goes from being unstable to being slowly stable.

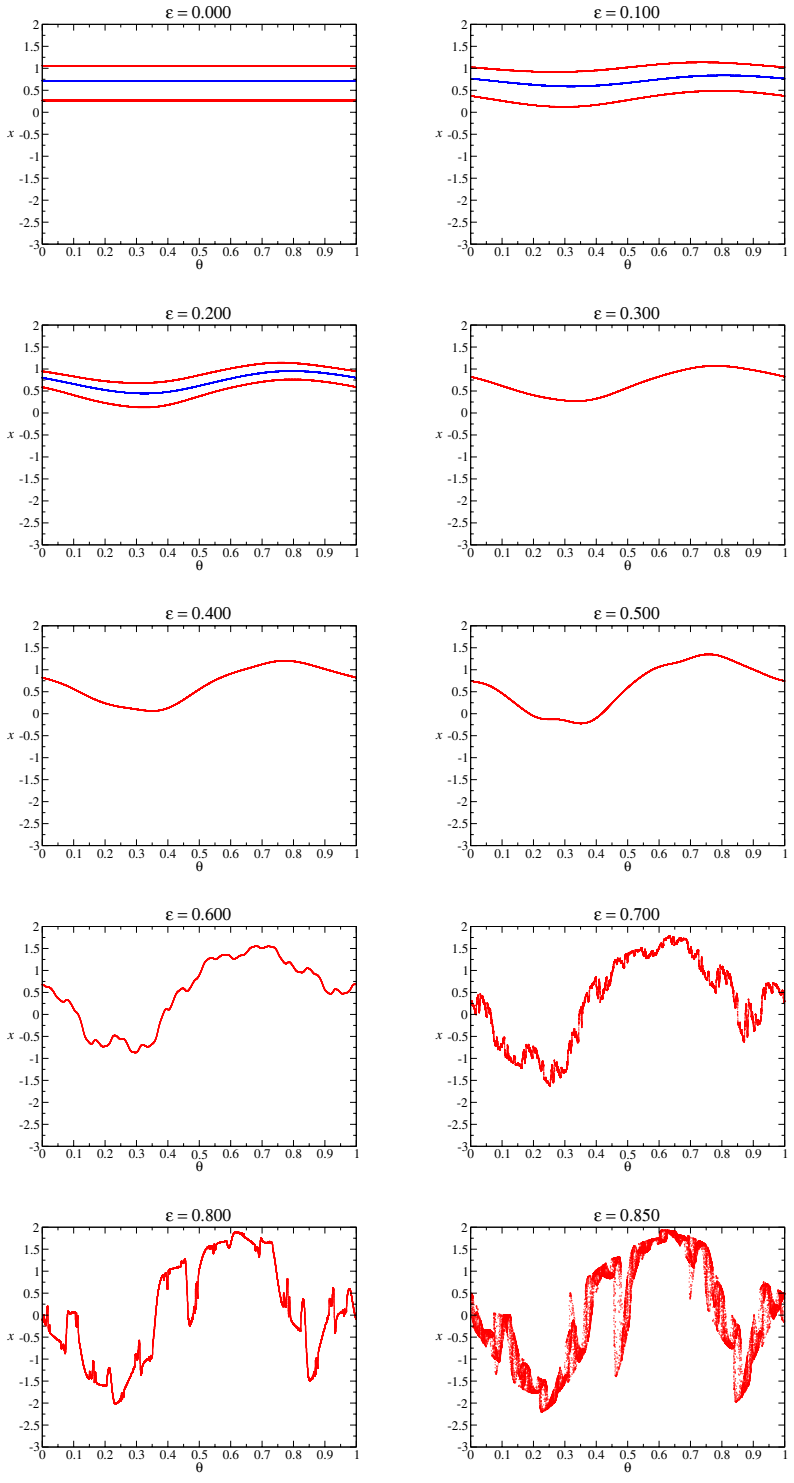


Figure 3. *θx projection of an attractor (red).*

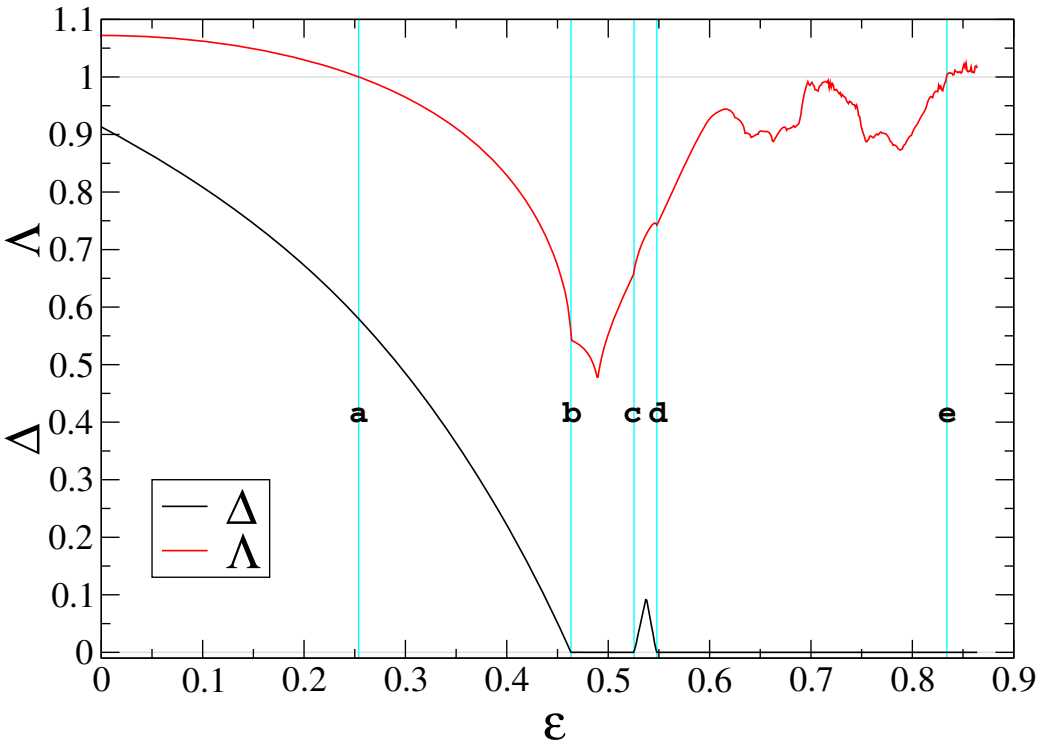


Figure 4. *Distance (Δ) between the invariant bundles and Lyapunov multiplier (Λ) of an invariant torus for the rotating Hénon map, continued with respect to ε.*

(b) This is the first instance of a global bifurcation, which we call the bundle merging scenario. This bifurcation will be described in more detail in section 6. In particular, we will identify quantitative regularities.

The spectrum remains two separate circles for $\varepsilon < \varepsilon_b \simeq 0.4633$. Nevertheless, the separation between the bundles becomes smaller and seems to tend to zero as ε approaches ε_b from below.

For ε slightly bigger than ε_b, it seems that there are no separate continuous invariant fast and slow bundles.

(c, d) The transition c, which happens at $\varepsilon_c \simeq 0.5256$, seems to be a reverse of the bundle merging scenario found in b. The bifurcation in d, which happens at $\varepsilon_d \simeq 0.5475$, is a direct bundle merging bifurcation.

(e) At $\varepsilon_e \simeq 0.8337$, the spectrum of the linearization touches 1, and the torus seems to disappear completely. In contrast with the quasi-periodic saddle-node bifurcation, there does not seem to be any companion unstable circle nearby. For ε slightly smaller than ε_e it seems that there are no continuous invariant bundles.

This transition seems hard to understand. We have not found any quantitative reg-

ularity. We just make here the simple remark that this transition seems visually similar to the *fractalization mechanism* that has been described several times in the literature [Kan84, NK96, SFKA96] (see also the review [PNR01]). In [SFKA96] there is a study of this transition in the rotating Hénon map.

Moreover, it seems that the torus is continuous (even analytic [HdlL06b]) until the outer radius of the spectrum (the maximal Lyapunov multiplier) touches 1 [HdlL06a, HdlL05c]. Notice that a normally hyperbolic torus persists under a sufficiently small perturbation, so the torus could survive until the maximal Lyapunov multiplier crosses 1 and the torus loses its normal hyperbolicity.

We would also like to refer to [HS05], which shows that some of the conclusions before may have been based on numerics which are not accurate.

Remark 3.1. We emphasize that the above list of the bifurcations is not exhaustive. As we will see in section 6, if we look very carefully in the interval $[\varepsilon_{\mathtt{b}}, \varepsilon_{\mathtt{c}}]$, there are other bundle merging bifurcations. Indeed, it seems that bundle merging bifurcations generate other "daughter" bundle merging bifurcations close by. See section 6.3.4.

The beginning of Figure 4 is computed using the reducibility method. This method can continue over the period doubling bifurcation. Nevertheless, it gets stuck on the bifurcations b, which correspond to lack of reducibility. To get an idea of the accuracy we display some typical results in Table 1.

We note that in this case, the fact that the reducibility fails does not imply that the torus cannot be continued. Since the torus is attractive, either the full matrix Newton method or even direct iteration can be used to compute the torus accurately.

Note that in the intervals $\varepsilon_{\mathtt{b}} \leq \varepsilon \leq \varepsilon_{\mathtt{c}}$ and $\varepsilon_{\mathtt{d}} \leq \varepsilon \leq \varepsilon_{\mathtt{e}}$ it seems that there are no continuous invariant subbundles and that the spectrum is not reducible (it is a full annulus). Indeed, it seems empirically that this lack of reducibility is abundant; that is, nonreducibility holds in a set of parameters of positive measure. Moreover, if we change slightly the other two parameters (a and b), the same phenomena is observed.

There are mathematical results showing that reducibility is dense for cocycles taking values in $SU(2)$ or in other compact groups [Kri99c, Kri99b, Kri99a, Kri01]. Some recent results for noncompact groups are in [AK03].

These mathematical results suggest that the regions where reducibility fails have a complicated structure. If we perform a numerical experiment with a prescribed accuracy, we obtain a partition of parameter space in intervals where the finite precision criterion for reducibility applies or not. In the cases where the criterion applies, the validating results in [HdlL06b] show that reducibility is indeed true. In the cases where the finite precision criterion fails, however, it means only that we need more delicate exploration. A higher precision computation may reveal other intervals where the criterion for reducibility holds. In the intervals not covered by this refined criterion, we can, of course, try more refined observations, and so on. Some of these explorations motivated by the mathematical insights are reported in section 6.3.4.

3.2. Breakdown of exponential dichotomies. To study the mechanism of loss of reducibility, we study the projective cocycle associated to the torus (see (15) in section 2.2) around the critical value. We will see that the breakdown of reducibility is associated to the

Table 1

Continuation with respect to ε of the invariant curve from the fixed point p_+ of the Hénon map, with $a = 0.68$, $b = 0.10$, using the reducibility method. The second column displays the two eigenvalues of the reduced cocycle. R, S, and T are estimates of the errors in the computation of the torus, the reducing change of variables, and its inverse, respectively. NF is the number of Fourier modes in the representation of the objects. tail *is an estimate of the maximum of the last* 10 *harmonics of the Fourier series. As a numerical test, notice that the product of both eigenvalues should be $-b = -0.10$ (test!). There is a "period halving" bifurcation in $\varepsilon_{\mathtt{a}} \simeq 0.2539339$.*

Rotating Hénon map: $a = 0.68$, $b = 0.1$

ε	Eigenvalues	R	S	T	NF	tail
0.000	-1.0721039594050942856 1e+00, 9.3274536599500630873 4e-02	0.0e+00	0.0e+00	0.0e+00	100	0.0e+00
0.020	-1.0717116868010741828 6e+00, 9.3308677353783033275 2e-02	1.7e-21	1.7e-19	1.7e-21	100	1.6e-144
0.040	-1.0705314982592398839 9e+00, 9.3411543857053337958 4e-02	1.1e-25	1.7e-19	4.4e-21	100	4.7e-117
0.060	-1.0685531696126862502 1e+00, 9.3584486803072760599 8e-02	5.8e-20	2.3e-19	1.2e-19	100	6.5e-101
0.080	-1.0657592783127116403 3e+00, 9.3829188295559262173 e-02	5.9e-20	2.4e-19	8.7e-21	100	2.1e-89
0.100	-1.0621245972127069526 8e+00, 9.4150912484680412544 2e-02	6.2e-20	1.4e-19	6.3e-20	100	2.3e-80
0.120	-1.0576151754647731314 7e+00, 9.4552349776992008309 1e-02	1.5e-20	1.8e-19	7.2e-20	100	6.7e-73
0.140	-1.0521870241789555498 6e+00, 9.5040138019219713758 5e-02	3.8e-21	2.4e-19	7.4e-20	100	1.7e-66
0.160	-1.0457842818201174949 9e+00, 9.5622014729420778467 8e-02	1.1e-21	4.2e-19	1.9e-20	100	7.1e-61
0.180	-1.0383366736601483286 2e+00, 9.6307876372601665698 4e-02	1.5e-20	8.4e-19	3.7e-20	100	8.2e-56
0.200	-1.0297559932648378499 0e+00, 9.7110384065792468023 2e-02	1.1e-19	6.9e-19	2.2e-19	100	3.4e-51
0.220	-1.0199312131703441745 6e+00, 9.8045827707499005569 0e-02	7.8e-20	4.6e-19	1.5e-19	100	6.2e-47
0.240	-1.0087216679974989665 7e+00, 9.9135374179597722743 8e-02	2.2e-20	4.2e-19	1.7e-19	100	5.4e-43
0.260	-9.9594753702695340482 2e-01, 1.0040689522513844808 7e-01	9.2e-20	2.2e-19	2.3e-19	100	1.3e-39
0.280	-9.8137655667180805909 5e-01, 1.0189768577632938586 4e-01	5.4e-21	4.7e-19	3.9e-20	100	4.0e-37
0.300	-9.6470538425355144475 0e-01, 1.0365859010663220290 2e-01	2.4e-21	2.2e-19	1.1e-19	100	6.7e-35
0.320	-9.4553280956143860764 9e-01, 1.0576047598642554470 4e-01	8.5e-20	7.1e-19	2.7e-19	100	9.5e-33
0.340	-9.2331857969474279146 4e-01, 1.0830497966699736652 5e-01	1.5e-19	4.3e-19	2.5e-19	100	8.0e-31
0.360	-8.9731239199462547594 3e-01, 1.1144390837811918187 2e-01	3.1e-20	7.3e-19	3.3e-19	100	3.2e-29
0.380	-8.6641457246764635687 7e-01, 1.1541818798729196630 2e-01	9.6e-20	7.1e-19	4.5e-19	100	1.1e-27
0.400	-8.2886932914411732078 7e-01, 1.2064627859165577493 6e-01	1.1e-19	1.6e-18	4.1e-19	100	3.7e-26

breakdown of the existence of exponential dichotomies. That is, the fast and slow bundles merge. See the discussion in section 2.2.5.

By direct iteration of (15), the orbits will converge to an attractor. Similarly, the iteration of the projectivization of the inverse cocycle will converge to a repellor. If there are slow and fast invariant bundles, the attractor corresponds to the slow bundle and the repellor to the fast bundle. The results of this experiment are depicted in Figure 5. Compare the results with those in Figure 4.

For $\varepsilon = 0.400$ the attractor and repellor of the projective cocycle are smooth curves. These correspond to the invariant subbundles of the cocycle corresponding to the least and most attractive eigenvalues, respectively.

We note that, as ε approaches a critical value $\varepsilon_{\mathtt{b}}$, the attractor and the repellor in the projective space approach each other. We note that the theory in [JS81, HdlL03b] ensures that, provided that the distance remains positive, the invariant bundles will be analytic.

In fact, the distance between both curves for $\varepsilon = 0.463$ is of the order of 10^{-3}. For $\varepsilon = 0.463254$ the distance is of the order of 10^{-6}. One can presume that the distance converges to zero as ε converges to the critical value $\varepsilon_{\mathtt{b}}$. We have detected that the critical parameter is

$$\varepsilon_{\mathtt{b}} \simeq 0.4632544711720$$

and that the corresponding maximal Lyapunov multiplier (the spectral radius of the transfer operator) is $\Lambda_{\mathtt{b}} \simeq 0.542306556$. Figure 6 depicts more accurately the attractors and repellors for the projective cocycle at the presumed critical value.

In section 6 we will present some quantitative regularities of the phenomenon, which we have found empirically, as well as other phenomena.

3.3. Topological evidence of bundle collapse (I). An indirect proof that there is a bundle collapse is the following.

Notice that when the cocycle has two continuous invariant bundles, it is reducible to a constant diagonal matrix, since the external rotation is Diophantine [JS81].

In the continuation of the attracting torus, we observe that before the transition $\mathtt{b}$, the leading eigenvalue is negative and the other is positive. In particular, we have the following.

- For $\varepsilon = 0.463$, the torus is attracting and the cocycle is reducible to a constant diagonal matrix

$$\operatorname{diag}(-0.603430499529439903989, 0.165719167456701022933).$$

After the transition $\mathtt{c}$ reducibility is recovered, but the leading eigenvalue is positive. In particular, we have the following.

- For $\varepsilon = 0.530$, the torus is attracting and the cocycle is reducible to a constant diagonal matrix

$$\operatorname{diag}(0.694546750046480781363, -0.143978789035162504966).$$

We emphasize that the calculations mentioned above are very reliable since reducibility methods apply and the results can be validated using the methods in [HdlL06b]. Hence, we

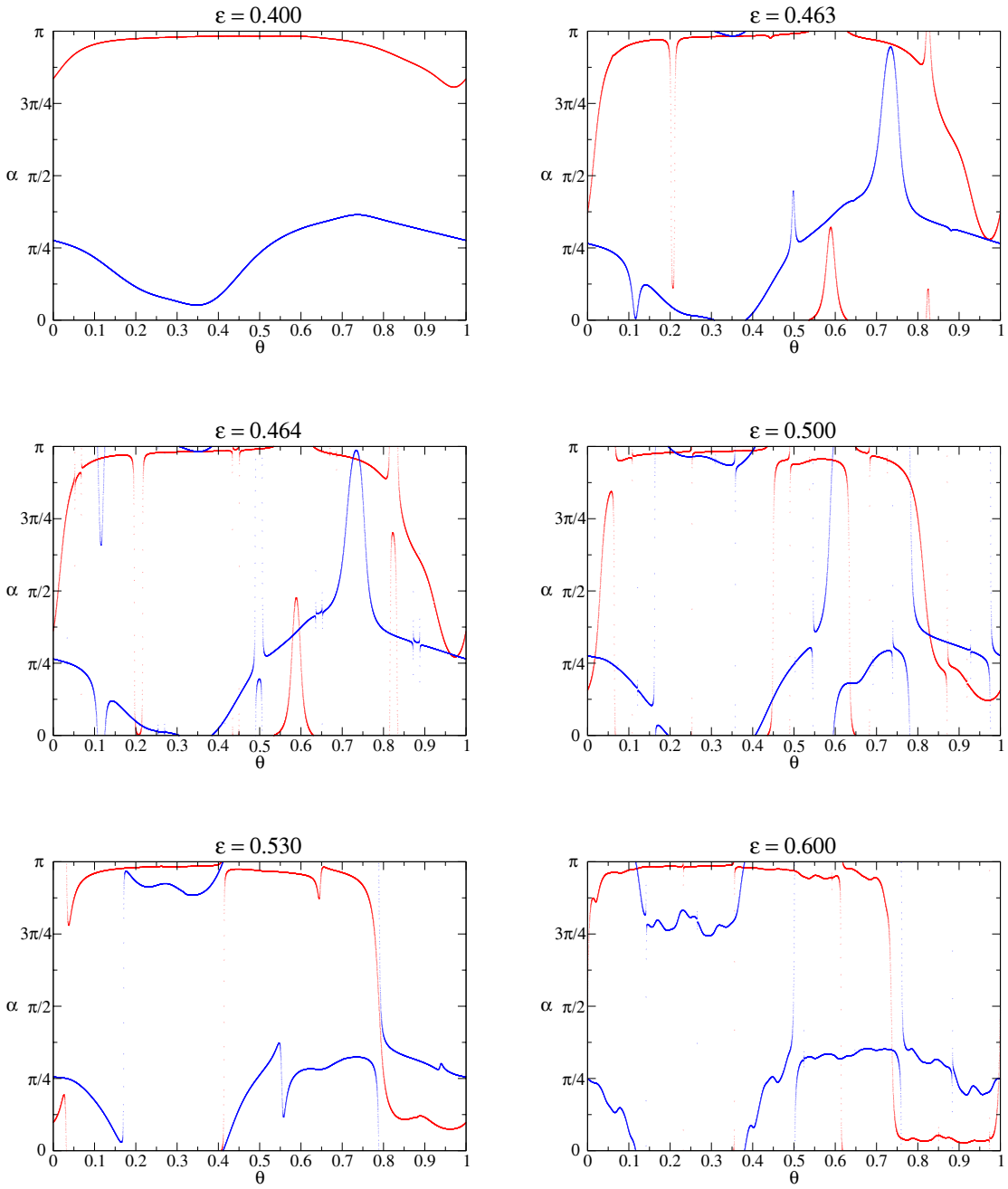


Figure 5. *Attractor (red) and repellor (blue) of the projective cocycle of the attracting invariant curve for the Hénon map.*

are quite sure that from b to c in Figure 4 there is a change of topology/dynamics of the invariant bundles.

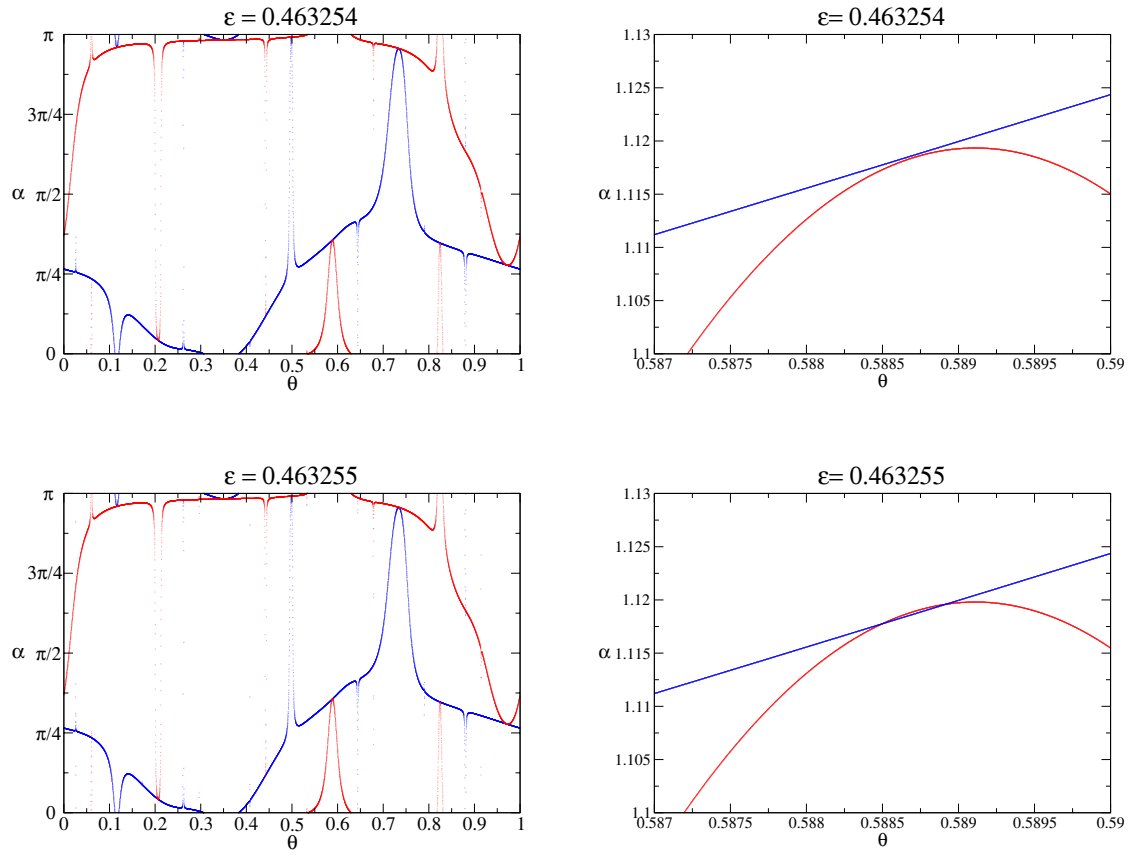


Figure 6. *Slow (red) and fast (blue) directions before and after the transition of the lack of reducibility, and the magnification of a very small region. The distance between the attractor and the repellor for* $\varepsilon = 0.463254$ *is* $\Delta \simeq 1.86 \cdot 10^{-6}$.

We also observe that the maximal and minimal Lyapunov multipliers are different during the continuation. If they were equal, they would be equal to $\sqrt{|b|} \simeq 0.3162$, which is not the case (see Figure 1).

As a result, the torus cannot remain reducible during the whole continuation, but the Lyapunov exponents cannot become close. The only possibility is that at some parameter value the torus becomes nonreducible (hence there are no 1D invariant bundles) even if the Lyapunov multipliers remain different.

Remark 3.2. Note that the argument above can only hope to show that there is a point in the interval where reducibility fails but nevertheless the spectrum is not trivial. Topological methods cannot hope to show that the transition is sharp—indeed, we do not expect it is based on our numerics—and much less the quantitative regularities discussed in section 6.2.

We think that the argument above can be made rigorous. The only ingredients that we need are that the index changes at the end of the interval and that the Lyapunov exponents do not degenerate.

Note that, given the validation results in [HdlL06b], to validate the results on the index, all that we need to do is show that the invariance equations for the torus and for the bundles are satisfied with enough accuracy that we can assess the sign of the averaged multipliers. The calculations we have performed can routinely be transformed into interval calculations [KSW96] and we certainly expect that the precision will be enough.

Since the Jacobian of the transformation is constant, the fact that the Lyapunov exponents do not coincide can be established if we prove lower bounds for one of the Lyapunov exponents. These lower bounds can be achieved by applying the method of [Her83] to the finite approximation computed.

We certainly hope that the present detailed numerics inspire computer assisted proofs and, of course, more rigorous studies.

Remark 3.3. The arguments above to show that there is a value of the parameter without reducibility can be modified slightly to show that they do not rely on the fact that the rotation is Diophantine.

Just notice that a normalization of an orientable bundle in $\mathbb{S}^1 \times \mathbb{T}$, instead of $\mathbb{P}^1 \times \mathbb{T}$, produces two curves ($\mathbb{S}^1$ is a double covering of $\mathbb{P}^1$). (If the bundle were nonorientable, we would also double the base torus $\mathbb{T}$.)

In the representation of the bundles in $\mathbb{S}^1 \times \mathbb{T}$, for $\varepsilon = 0.463$ the slow bundle corresponds to a 2-periodic curve, and the fast bundle corresponds to two fixed curves. Moreover, for $\varepsilon = 0.530$, the slow bundle corresponds to two fixed curves, and the fast bundle corresponds to a 2-periodic curve. These two cases are homotopically nonequivalent. Hence, it is impossible that the torus remains reducible throughout the continuation. We also note that this breakdown in reducibility has to happen while the Lyapunov multipliers remain different.

Hence, the mechanism of bundle collapse is independent of the Diophantine properties of the rotation. The above version of the argument applies to Liouville numbers.

We have observed the same mechanism with the external frequency $\omega = e/4$ (with a relative distance less than 10% of the golden mean). This number has very different Diophantine properties from those of the golden mean; in particular, it does not have a periodic continuous fraction expansion. (The continued fraction expansion of $e/4$ is known explicitly and has unbounded entries.)

4. Example 2: The rotating standard map.

Our second example is the *rotating standard map*. This is a quasi-periodic perturbation of the standard map [Chi79, Gre79], that is, a symplectic (area preserving) map of the annulus $\mathbb{T} \times \mathbb{R}$.

We will study the continuation of hyperbolic invariant tori, whose theory has been developed. Moreover, we will also study continuation of elliptic tori and the bifurcation of elliptic tori. We think that it would be quite desirable to develop rigorous a posteriori estimates for these cases. We will also describe another version of the bundle merging bifurcation, which has some quantitative differences from the one studied in section 3.2.

4.1. Preliminaries and notation.

The rotating standard map is given by

$$\begin{aligned} \bar{x} &= x + y - s(x)(K + \varepsilon c(\theta)) \qquad (\mathrm{mod}\ 1), \\ \bar{y} &= \qquad y - s(x)(K + \varepsilon c(\theta)), \\ \bar{\theta} &= \theta + \omega \qquad\qquad\qquad\quad (\mathrm{mod}\ 1), \end{aligned} \tag{27}$$

where

$$s(x) = \frac{1}{2\pi}\sin(2\pi x) \tag{28}$$

is a 1-periodic odd function,

$$c(\theta) = \cos(2\pi\theta) \tag{29}$$

is a 1-periodic even function, and K and ε are positive parameters. We will refer to y, x as the action-angle variables and ω as the external frequency.

So, the extended phase space is $\mathbb{T} \times \mathbb{R} \times \mathbb{T}$. We have taken $\omega = \tau - 1$, where τ is the largest root of the polynomial $p(t) = t^3 - t^2 - t - 1$. That is,

$$\omega = \sqrt[3]{\frac{19}{27} + \sqrt{\frac{11}{27}}} + \sqrt[3]{\frac{19}{27} - \sqrt{\frac{11}{27}}} - \frac{2}{3}.$$

This choice of an irrational number is motivated by this number's usefulness in the study of 2D KAM tori. See [Har98]. For our purposes, it is useful to check the properties when the frequency is different from the golden mean. Note that the number above does not have a periodic continued fraction expansion.

The rotating standard map is reversible (as is the standard map), since it is the composition $F = I_1 \circ I_0$ of two involutions I_0, given by

$$\begin{aligned} \bar{x} &= -x && \pmod 1, \\ \bar{y} &= y - s(x)(K + \varepsilon c(\theta)), \\ \bar{\theta} &= -\theta && \pmod 1, \end{aligned}$$

and I_1, given by

$$\begin{aligned} \bar{x} &= -x + y && \pmod 1, \\ \bar{y} &= y, \\ \bar{\theta} &= -\theta + \omega && \pmod 1. \end{aligned}$$

As far as we know, the rotating standard map was introduced in [ACS91, ACS92], where some properties of the map and existence of KAM tori were investigated. This example was also studied in [Tom96], which used a higher dimensional extension of Greene's criterion [Gre79] to study the breakdown of a 2D invariant KAM torus. The thesis [Har98] contains a study of the breakdown of the 2D KAM tori using variational principles.

The standard map, corresponding to $\varepsilon = 0$, is an area preserving map that has periodic orbits of all rotation numbers. We say that the *rotation number* of an n-periodic orbit through (x_0, y_0) is p/n if $x_n = x_0 + p$, $y_n = y_0$, where $(x_n, y_n) = F^n(x_0, y_0)$ and F is the lift of the standard map to $\mathbb{R} \times \mathbb{R}$ (that is, we think of x as a real variable instead of as an angle).

Analogously, we can define the rotation number of a n-periodic curve $K_0(\theta) = (x_0(\theta), y_0(\theta))$ of the rotating standard map as the rational number p/n such that

$$x_n(\theta) = x_0(\theta + n\omega) + p, \;\; y_n(\theta) = y_0(\theta + n\omega),$$

where $K_n(\theta) = F(\theta, n)(K_0(\theta)) = (x_n(\theta), y_n(\theta))$ is the n-iterate of the torus K_0.

Notice that the tori $\{x = 0,\ y = 0\}$ and $\{x = \frac{1}{2},\ y = 0\}$ are invariant under the rotating standard map. Their rotation numbers are both $0/1$. If $\varepsilon = 0$, the torus $\{x = 0,\ y = 0\}$ corresponds to an elliptic fixed point if $0 < K < 4$ and to a hyperbolic fixed point (with reflection) if $K > 4$. It is parabolic if $K = 4$. The torus $\{x = \frac{1}{2},\ y = 0\}$ for $\varepsilon = 0$ corresponds to a hyperbolic fixed point if $K > 0$.

Notice also that the torus $\{x = 0,\ y = \frac{1}{2}\}$ is a 2-periodic curve under the rotating standard map, whose rotation number is $1/2$. If $\varepsilon = 0$, it corresponds to an elliptic 2-periodic orbit if $K < \sqrt{2}$ and to a hyperbolic 2-periodic orbit if $K > \sqrt{2}$.

In our numerical experiments that are described below, we have fixed $K = 0.2$ and changed the forcing constant ε.

4.2. A 2-periodic normally hyperbolic torus. The simplest nontrivial periodic torus for the rotating standard map with $K > 0$ is generated from $\varepsilon = 0$ by the hyperbolic periodic orbit of rotation number $1/2$. We also compute this 2-periodic saddle curve using the reducibility method, and the results are displayed in Table 2 and Figure 7.

Once we have computed a 2-periodic saddle curve, we compute its stable and unstable invariant manifolds. We have expanded them to order 30 and then we have globalized them. Our programs produce 2D grids of those objects, which we display in Figure 8, for $K = 0.2$, $\varepsilon = 0.5$. We have also made some slices of those objects, producing the familiar homoclinic web. This is displayed in Figure 9. Keep in mind that these are slices of 2D objects in a 3D space. Notice that the objects we produce are symmetric. This is due to the reversibility of the rotating standard map.

4.3. Continuation of reducibility. In this section we will analyze the applicability of the reducibility method to normally elliptic tori. Later, we will discuss some bifurcations that can arise when the reducibility fails.

As a first example, we will consider the torus $\{x = 0,\ y = 0\}$ of the rotating standard map, which is invariant for all the values K, ε. The normal behavior is given by the cocycle

$$M(\theta) = \begin{pmatrix} 1 - (K + \varepsilon\cos(2\pi\theta)) & 1 \\ -(K + \varepsilon\cos(2\pi\theta)) & 1 \end{pmatrix}, \tag{30}$$

with rotation ω. So then, the reducibility method provides approximations of the reduced matrices and the corresponding Floquet transformations.

This example has the advantage that the torus is fixed for all the values of the parameters, and so then the only thing we have to do is study the cocycle (30). We will refer to it as the *standard cocycle*. This example serves as a test and gives some insight about the general case, and the effects that happen at breakdown of reducibility are easier to study. A continuation of elliptic tori that are not fixed by the symmetry will be undertaken later in this section.

Remark 4.1. The standard cocycle (30) is equivalent to the *almost Mathieu cocycle*

$$A(\theta) = \begin{pmatrix} a - b\cos(2\pi\theta) & -1 \\ 1 & 0 \end{pmatrix}, \tag{31}$$

Table 2

Continuation of a saddle 2-periodic torus for the rotating standard map, using the reducibility method, for $K = 0.2$. The second column displays the two eigenvalues of the reduced cocycle. R, S, and T are estimates of the errors in the computation of the torus, the reducing change of variables, and its inverse, respectively. NF is the number of Fourier modes in the representation of the objects. tail *is an estimate of the maximum of the last* 10 *harmonics of the Fourier series. The whole computation took around* 2*m* 30*s on a Pentium III at* 900 *MHz.*

Rotating standard map: $K = 0.200$

ε	Eigenvalues	R	S	T	NF	tail
0.000	8.1893486357896750 2551e-01, 1.2210983369662988 0250e+00	2.7e-20	0.0e+00	0.0e+00	100	0.0e+00
0.100	8.0683605625638765 6202e-01, 1.2394091615586286 9636e+00	5.5e-20	8.8e-19	1.1e-19	100	7.0e-97
0.200	7.7566250567187736 4351e-01, 1.2892204956249650 5027e+00	1.8e-21	2.3e-19	4.7e-21	100	8.2e-77
0.300	7.3451571885142741 5026e-01, 1.3614412521541596 1203e+00	5.6e-20	3.8e-19	2.3e-19	100	5.7e-61
0.400	6.8981024373323136 2798e-01, 1.4496740358450338 8026e+00	9.3e-22	3.7e-19	1.0e-20	100	2.4e-55
0.500	6.4493953891992135 7654e-01, 1.5505329409245051 2950e+00	5.9e-20	9.6e-19	1.3e-20	100	1.6e-49
0.600	6.0156108516048039 4491e-01, 1.6623415720669942 7532e+00	3.2e-20	8.0e-19	1.3e-19	100	3.8e-45
0.700	5.6045523415134464 3524e-01, 1.7842638253066278 3850e+00	5.8e-20	2.4e-19	1.3e-19	100	1.1e-41
0.800	5.2195628041378457 2255e-01, 1.9158692739691585 8879e+00	3.2e-20	4.5e-19	1.2e-19	100	8.2e-39
0.900	4.8616262844891683 0503e-01, 2.0569248672825007 9975e+00	1.0e-19	1.4e-18	2.5e-19	100	2.1e-36
1.000	4.5304348481304667 2187e-01, 2.2072936341037127 8277e+00	5.5e-21	8.9e-19	1.2e-19	100	2.6e-34
1.100	4.2249636498496400 8530e-01, 2.3668842690175297 7652e+00	4.7e-20	6.4e-19	4.1e-19	100	1.7e-32
1.200	3.9437999153548720 5742e-01, 2.5356255932421402 5869e+00	3.9e-20	9.3e-19	7.5e-20	100	7.0e-31
1.300	3.6853404324704392 0327e-01, 2.7134535284428467 3521e+00	4.5e-20	1.1e-18	7.2e-20	100	1.9e-29
1.400	3.4479138015823431 6120e-01, 2.9003045248436091 9260e+00	1.4e-19	1.4e-18	4.3e-19	100	3.6e-28
1.500	3.2298569086591774 7972e-01, 3.0961123922209102 4392e+00	4.4e-20	1.9e-18	1.7e-19	100	5.0e-27
1.600	3.0295622021268269 9930e-01, 3.3008069591638536 5358e+00	7.6e-21	1.2e-18	5.9e-20	100	5.5e-26
1.700	2.8455057726625387 5027e-01, 3.5143137280101187 4672e+00	8.2e-20	1.1e-18	2.4e-19	100	4.8e-25
1.800	2.6762626218050565 9319e-01, 3.7365540730286433 5647e+00	5.4e-20	3.2e-18	1.9e-19	100	3.5e-24
1.900	2.5205133693229238 3442e-01, 3.9674457282034821 8576e+00	1.3e-19	1.7e-18	5.2e-19	100	2.1e-23
2.000	2.3770453008499079 9479e-01, 4.2069034176271353 0468e+00	7.2e-20	2.2e-18	1.8e-19	100	1.1e-22

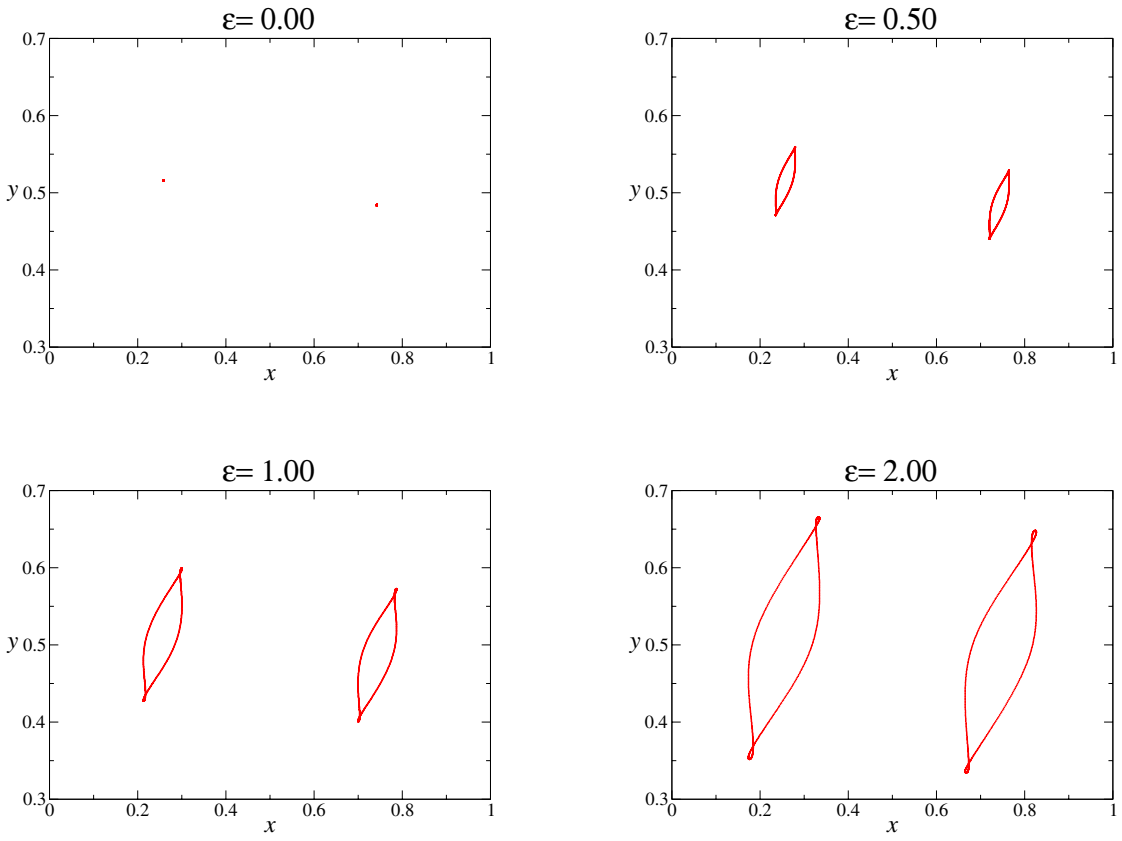


Figure 7. *xy projection of a 2-periodic saddle curve for the rotating standard map ($K = 0.20$).*

with rotation ω, where $a = 2 - K$, $b = \varepsilon$. In fact, $M(\theta)P = PA(\theta)$, where P is the constant antisymplectic matrix

$$\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}.$$

The almost Mathieu cocycle has called the attention of many authors, among them [CEY90, BS82, Las94, AS82, AS83, DS83a, Eli01, Jit99, AK03, Pui04].

We have analyzed numerically the standard cocycle for $K = 0.2$, moving the parameter ε. We have applied the reducibility method to reduce the linear cocycle. The problem of reducibility has been considered in [Eli88, Pös89, JS92, Pui05]. There is a Cantor family of parameters for which the cocycle is reducible and elliptic. There is also an open set of parameters for which the cocycle is normally hyperbolic (and reducible). The results are displayed in Table 3.

4.3.1. Detection of resonances in elliptic tori. We have observed that the program implementing the reducibility method cannot continue beyond the critical value

$$\varepsilon_{crit} \simeq 0.0963488851723619376131.$$

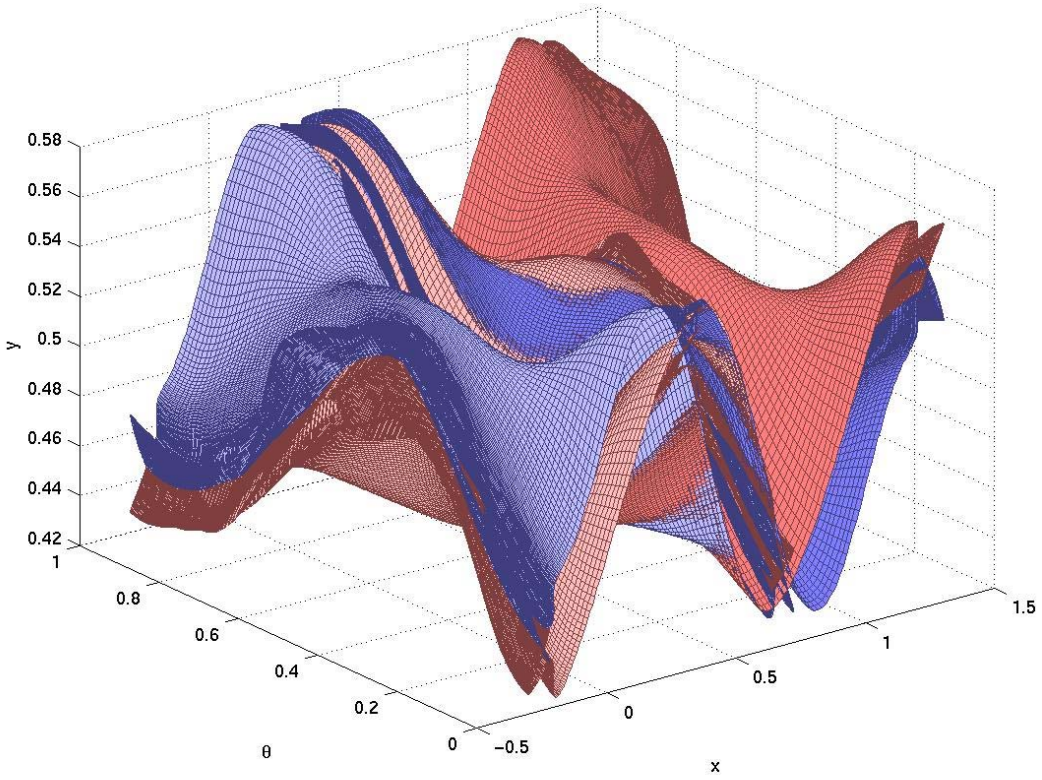


Figure 8. *3D view of the stable (blue) and unstable (red) manifolds of a hyperbolic 2-periodic circle for the rotating standard map. The parameters are $K = 0.2$ and $\varepsilon = 0.5$, and the frequency of the forcing is $\omega = \tau - 1$ (see text for details).*

The reason is that for such a value the rotation number γ of the cocycle (the argument of the eigenvalue) and the external frequency ω, which are

$$\gamma \simeq -0.504895542331356775420,\ \omega \simeq 0.83928675521416126683,$$

are very close to $-1:1$ resonance:

$$\left| \frac{\gamma}{2\pi} - \frac{k_1 + k_2\omega}{2} \right| \simeq 1.085821\ 10^{-9} \tag{32}$$

for $k_1 = -1$, $k_2 = 1$. It is not hard to show that, when there is a resonance, if there is a nonlinear term, this provides an obstruction to reducibility. So, it is quite possible that there is a dense set of parameter values for which reducibility—in the mathematical sense—fails. On the other hand, we note that, when the order of the resonance is high, the interval shows only on high order perturbation theory. Hence, it seems that the measure of the gaps except those of resonances whose numerator k_2 is small will be visible.

4.3.2. Analysis of the cocycle. We have computed the dominant Lyapunov multiplier and the rotation number of the standard cocycle [BS98]. To do so, we have considered orbits

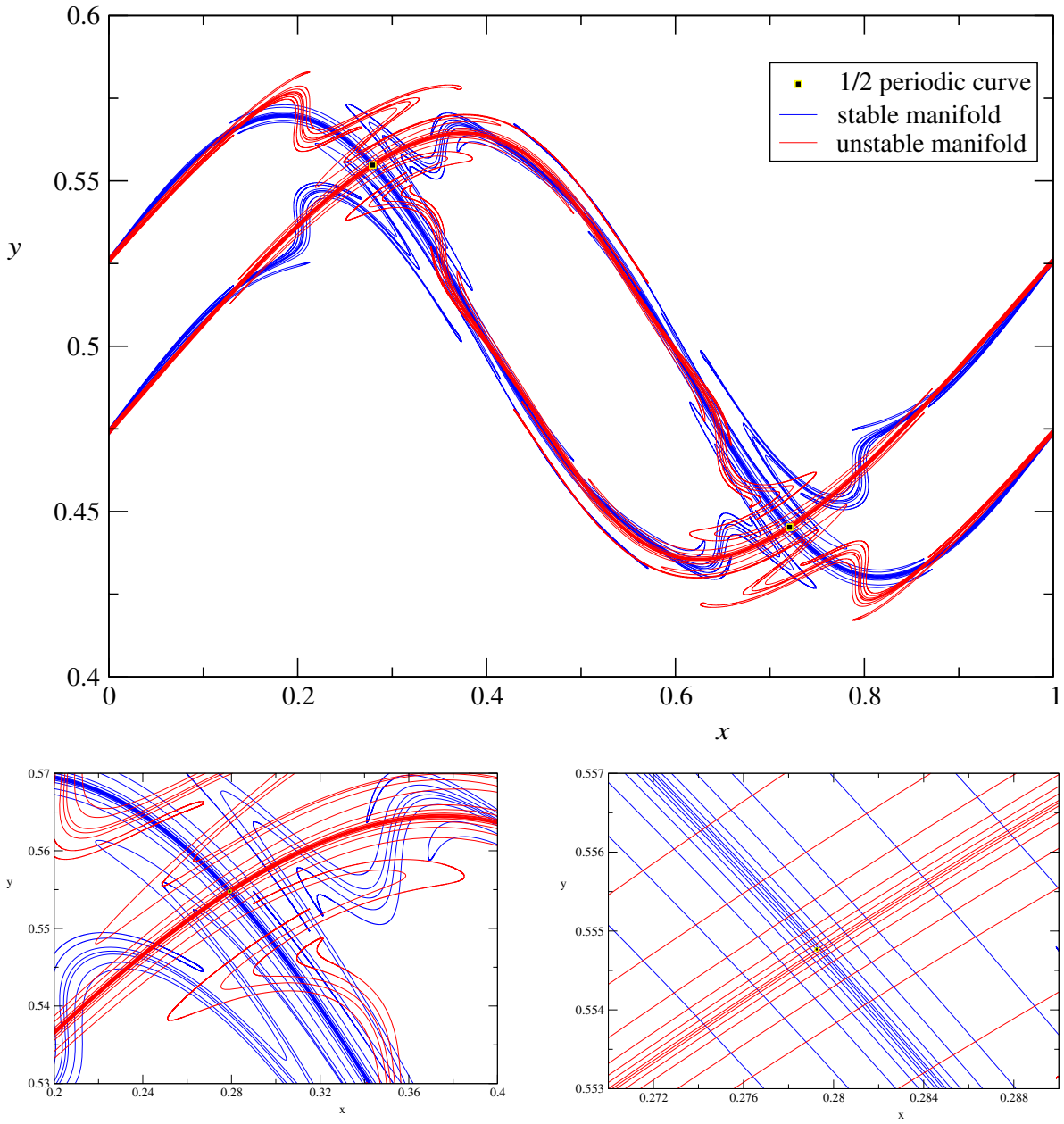


Figure 9. *Heteroclinic web associated with the hyperbolic 2-periodic circle. The section of the stable and unstable manifolds is with $\theta = 0$. Both of the lower figures are magnifications of the first one.*

of length 10^8 of the cocycle (applied to a random vector). The results are displayed in the two last columns of Table 3 (and agree with the computed eigenvalues in about 7 digits) and in Figure 10. We note that just after the critical value ε_{crit}, the maximal Lyapunov multiplier is appreciably bigger than 1, and the rotation number is phase locked in the angle γ, so the torus bifurcates into a normally hyperbolic torus.

Table 3

Continuation of reducibility of the torus $\{x = 0,\ y = 0\}$ *of the rotating standard map for* $K = 0.2$. *This torus is normally elliptic for* $\varepsilon = 0$. *The last two columns display the Lyapunov multiplier and the rotation number, computed directly from the cocycle (compare the results). The reducibility method stops for* $\varepsilon = \varepsilon_{crit} \simeq 0.0963488851723619376131$, *where there is a strong resonance between the internal and external frequencies. After the critical value, the maximal Lyapunov multiplier is appreciably bigger than* 1, *and the rotation number is phase locked, so the torus is normally hyperbolic.*

Rotating standard map: $K = 0.200$. Torus $x = 0$, $y = 0$.

ε	Eigenvalues	S	T	NF	Multiplier	Rot. number
0.000	1.0000000000000000000 exp(±0.4510268117962624452 70**i**)	0.0e+00	0.0e+00	100	0.99999997	-0.45102678
0.005	0.9999999999999999989 exp(±0.4511002127121101543 78**i**)	1.7e-20	6.4e-20	100	1.0000000	-0.45110021
0.010	0.9999999999999999995 exp(±0.4513209918744906085 08**i**)	1.6e-19	1.7e-19	100	1.0000000	-0.45132095
0.015	0.9999999999999999995 exp(±0.4516909025130547745 03**i**)	1.8e-19	3.4e-20	100	1.0000000	-0.45169083
0.020	0.9999999999999999995 exp(±0.4522129489701396602 27**i**)	1.3e-19	1.7e-19	100	1.0000000	-0.45221294
0.025	0.9999999999999999989 exp(±0.4528951875953176495 8**i**)	1.5e-19	1.8e-19	100	1.0000000	-0.45289149
0.030	0.9999999999999999995 exp(±0.4537325863245850024 24**i**)	3.5e-19	2.2e-19	100	0.99999997	-0.45373253
0.035	1.0000000000000000011 exp(±0.4547440116927199499 31**i**)	2.0e-19	9.0e-20	100	1.0000000	-0.45474394
0.040	0.9999999999999999989 exp(±0.4559359715436447212 53**i**)	2.1e-19	1.6e-19	100	1.0000000	-0.45593589
0.045	0.9999999999999999995 exp(±0.4573215837425207971 94**i**)	3.5e-19	3.9e-20	100	1.0000000	-0.45732150
0.050	1.0000000000000000000 exp(±0.4589178271574768586 06**i**)	2.2e-19	2.1e-19	100	0.99999997	-0.45891779
0.055	1.0000000000000000000 exp(±0.4607469330491622167 58**i**)	3.9e-19	1.6e-19	100	1.0000001	-0.46074693
0.060	0.9999999999999999984 exp(±0.4628385686037936012 51**i**)	6.0e-19	3.8e-20	100	1.0000001	-0.46283853
0.065	0.9999999999999999984 exp(±0.4652334328941057376 63**i**)	3.6e-19	1.6e-19	100	1.0000001	-0.46523339
0.070	0.9999999999999999995 exp(±0.4679895639245924774 67**i**)	5.0e-19	2.6e-19	100	1.0000001	-0.46798955
0.075	1.0000000000000000011 exp(±0.4711943694898515523 62**i**)	1.2e-18	2.9e-19	100	1.0000001	-0.47119433
0.080	0.9999999999999999995 exp(±0.4749904173933330268 23**i**)	6.8e-19	2.6-19	100	1.0000001	-0.47499034
0.085	0.9999999999999999984 exp(±0.4796412611450507015 20**i**)	6.1e-19	5.3e-19	100	1.0000001	-0.47964124
0.090	0.9999999999999999995 exp(±0.4857579944362734340 93**i**)	8.4e-19	5.7e-19	100	1.0000001	-0.48575788
0.095	1.0000000000000000011 exp(±0.4959617641014786752 83**i**)	2.9e-18	1.5e-18	100	1.0000001	-0.49596162
0.096	0.9999999999999999989 exp(±0.5003407231996228728 35**i**)	4.3e-18	2.5e-18	100	1.0000002	-0.50034075
0.09625	1.0000000000000000000 exp(±0.5024691304144695623 39**i**)	6.7e-18	6.4e-18	100	1.0000003	-0.50246901
ε_{crit}	1.0000000999835361949 exp(±0.5048955423313567754 20**i**)	0.6e-11	3.1e-11	450	1.0000010	-0.50489537
0.097	-	-	-	-	1.0062575	-0.50489537
0.098	-	-	-	-	1.0100077	-0.50489537
0.099	-	-	-	-	1.0127294	-0.50489537
0.100	-	-	-	-	1.0149916	-0.50489537
0.150	-	-	-	-	1.0654805	-0.50489538
0.200	-	-	-	-	1.1004674	-0.50489539
0.250	-	-	-	-	1.1325480	-0.50489539
0.300	-	-	-	-	1.1631207	-0.50489539
0.500	-	-	-	-	1.2746950	-0.50489539
1.000	-	-	-	-	1.4820201	-0.50489563
1.500	-	-	-	-	1.5881951	-0.50489566
2.000	-	-	-	-	1.5650877	-0.50489569

Remark 4.2. From [Her83], the maximal Lyapunov multiplier of the almost Mathieu cocycle (31) (and so of the standard cocycle (30)) is bounded from below by $\frac{\varepsilon}{2}$. In fact, if the almost Mathieu cocycle is not uniformly hyperbolic, then the maximal Lyapunov multiplier is $l(\varepsilon) = \max(1, \frac{\varepsilon}{2})$ [BJ02]. See Figure 10, where we also display the lower bound $l(\varepsilon)$.

We distinguish several regions in Figure 10. Notice that after the critical value, the dominant Lyapunov multiplier is bigger than 1, so the torus is not going to be elliptic anymore. In such a case, the spectrum of the transfer operator associated to the cocycle is going to be either two circles or a fat annulus (notice that the radii are inverses of each other).

When the rotation number is phase locked then the spectrum is two circles and the cocycle has stable and unstable invariant subbundles, so the torus is normally hyperbolic. (And the type of resonance has to do with the index of the stable and unstable subbundles.)

In the regions where the rotation number is not phase locked (and the Lyapunov multiplier is bigger than 1), the torus is nonuniformly hyperbolic. Notice that a nonuniformly hyperbolic torus is destroyed by generic perturbations [Mn78], but in the present example the torus survives because the perturbation vanishes on the torus.

We can confirm our expectations using the trick of projectivization described in section 2.2.1. The examples we display in Figure 11 correspond to the regions detected in Figure 10.

The following is an analysis of both Figure 10 and Figure 11.

- From $\varepsilon = 0$ to $\varepsilon \simeq 0.096$ the torus is "almost" elliptic (there are small gaps in which the torus loses its ellipticity). See Figure 11 for $\varepsilon = 0.050$ and $\varepsilon = 0.095$, where the dynamics of the projective cocycle is "almost" ergodic and a strong resonance is born.
- From $\varepsilon \simeq 0.096$ to $\varepsilon \simeq 2.345$ the torus is normally hyperbolic. This happens for a whole open interval that in this case is a big gap corresponding to a $-1:1$ resonance. Moreover, the indices of the stable and unstable subbundles are $\frac{1}{2} = \frac{k_2}{2}$, where $k_2 = 1$ in (32), and as a result those subbundles are nonorientable. See Figure 11 for $\varepsilon = 0.100$, $\varepsilon = 0.500$, and $\varepsilon = 1.000$.
- From $\varepsilon \simeq 2.345$ to $\varepsilon \simeq 2.697$ there are two different behaviors that coexist: the torus is normally hyperbolic in the small bumps in Figure 10, which are open intervals (moreover, the unstable and stable bundles probably have a large index), and the torus is nonuniformly hyperbolic in the complementary intervals. See Figure 11 for $\varepsilon = 0.250$, where the attractor and repellor of the projective cocycle collapse.
- From $\varepsilon \simeq 2.697$ to $\varepsilon \simeq 4.332$ the torus is again normally hyperbolic for a large open interval. The resonance is $-2:2$, and the invariant subbundles have index $1 = \frac{2}{2}$, so they are orientable. See Figure 11 for $\varepsilon = 4.000$.
- From $\varepsilon \simeq 4.332$ to $\varepsilon = 5.000$, where we stop our analysis, we are again in a normally/nonuniformly hyperbolic region. See Figure 11 for $\varepsilon = 5.000$.

To finish the analysis of the standard cocycle we come back to our starting point in this section: the detection of a $-1:1$ resonance between the internal and normal frequencies that prevent the continuation of the reducibility (notice that we enter in a region that, in a sense, is "more reducible" than in the first elliptic region). We emphasize that we have detected this phenomenon numerically because the resonance is very strong, producing a big gap. But it occurs also in very small gaps. Numerical and rigorous results on these kinds of gaps and the corresponding bifurcations appear in [BHJ$^+$03] for Hamiltonian systems.

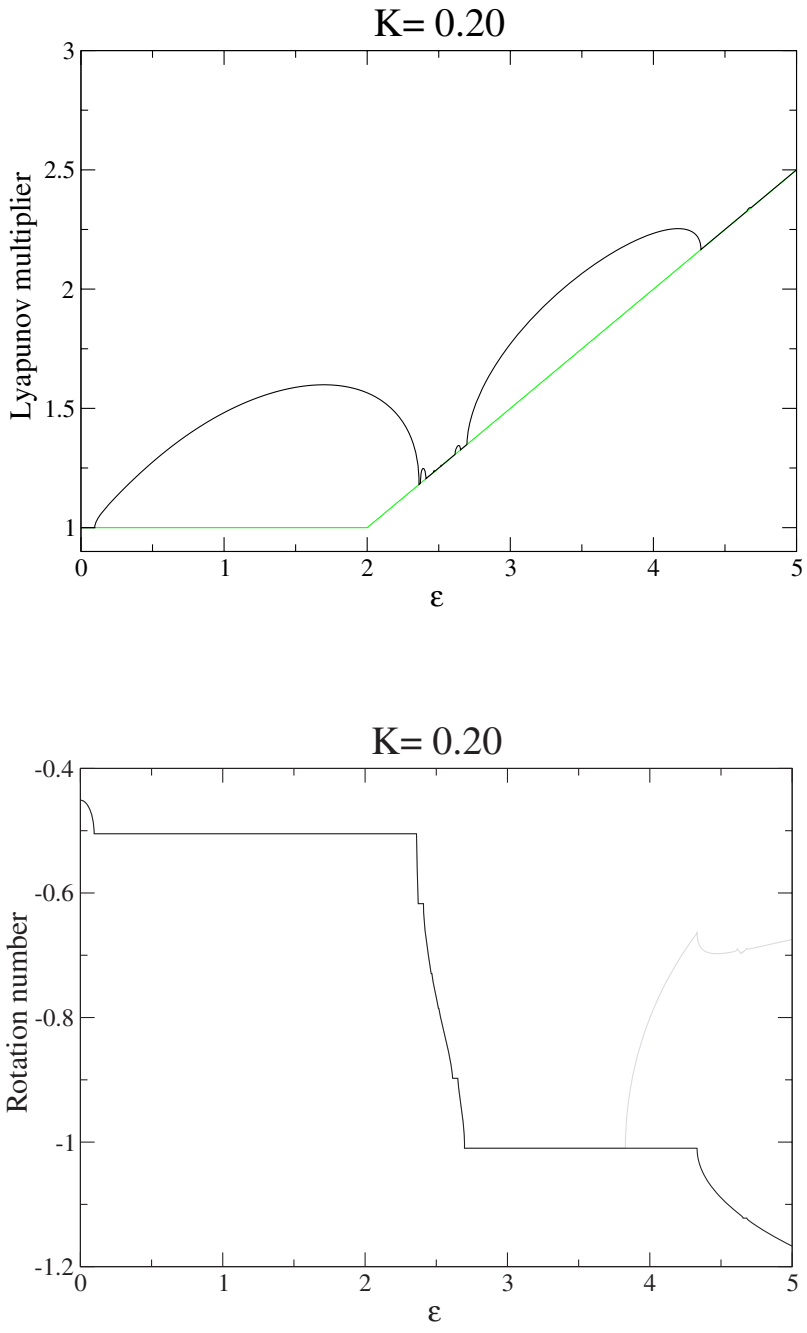


Figure 10. *Lyapunov multiplier and rotation number as functions of ε of the standard cocycle (30) with $\omega = \tau - 1$, $K = 0.2$. The green function is a lower bound of the Lyapunov multiplier: $l(\varepsilon) = \max(1, \frac{\varepsilon}{2})$. The rotation number has been computed with the Sturmian definition (19). We also display the average rotation number (grey color) that produces incorrect estimates beyond $\varepsilon \simeq 3.829 \simeq 4 - K$.*

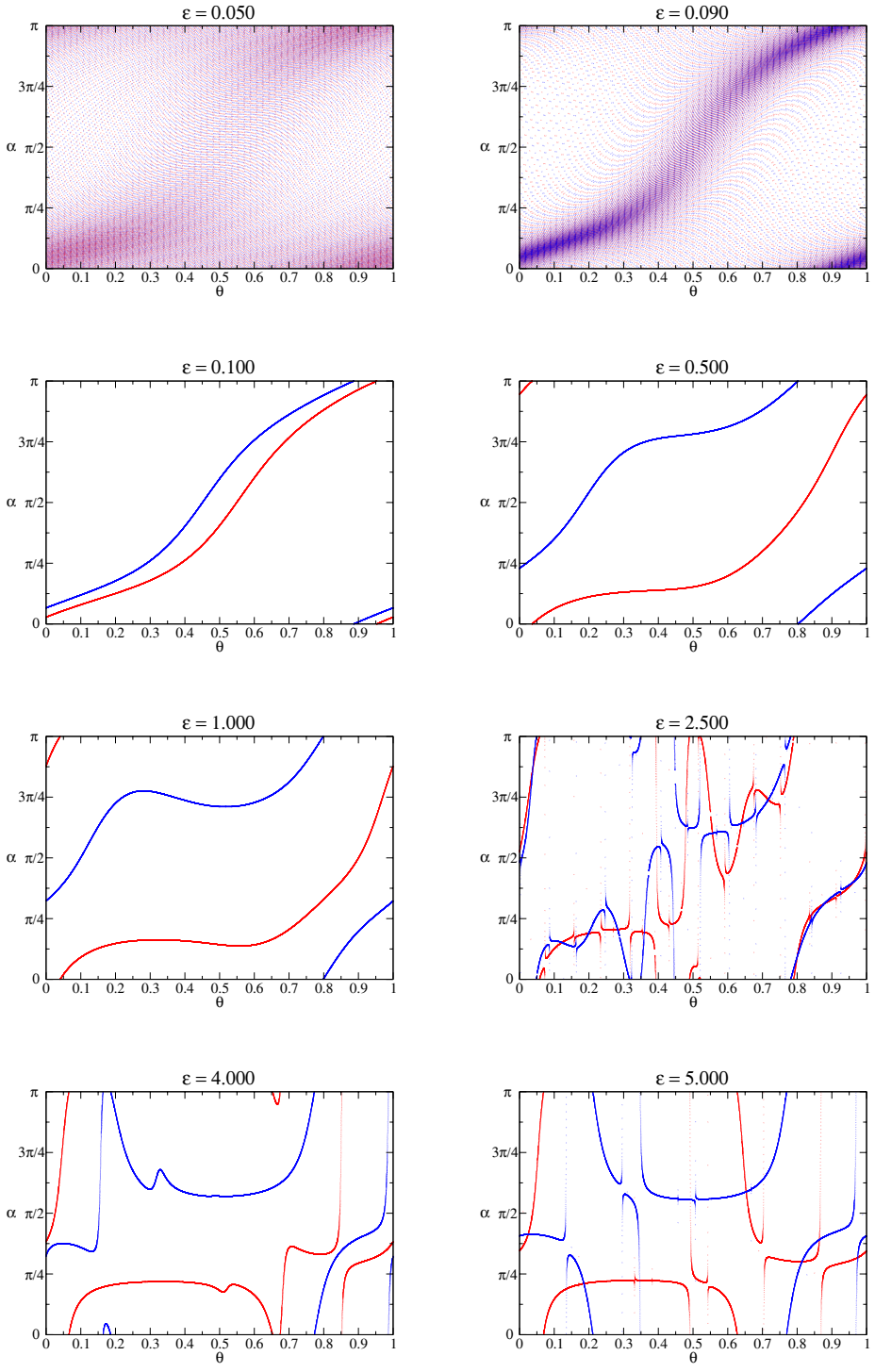


Figure 11. *Attractor (red) and repellor (blue) of the projective standard cocycle with* $K = 0.2$. *This is associated to the torus* $\{x = 0, y = 0\}$ *that is invariant for the rotating standard map.*

Table 4

Continuation of reducibility and Floquet transformation of the torus $\{x = 0, y = 0\}$ of the rotating standard map for $K = 0.2$. We improve reducibility using rotating transformations, so we can cross the resonance $-1 : 1$ that appears in $\varepsilon = \varepsilon_{crit} \simeq 0.096348885172361937 6131$.

Rotating standard map: $K = 0.200$. Torus $\{x = 0, y = 0\}$.

ε	Eigenvalues	S	T	NF
0.000	1.0000000000000000000 exp(±0.45102681179626244527**i**)	0.0e+00	0.0e+00	100
0.010	0.99999999999999999995 exp(±0.45132099187449060824**i**)	7.7e-20	6.9e-20	100
0.020	0.99999999999999999984 exp(±0.45221294897013966028**i**)	1.4e-19	1.0e-20	100
0.030	1.0000000000000000011 exp(±0.45373258632458500232**i**)	1.6e-19	9.3e-20	100
0.040	1.0000000000000000000 exp(±0.45593597154364472120**i**)	2.1e-19	1.5e-19	100
0.050	1.0000000000000000000 exp(±3.09561493159350098947**i**)	5.1e-10	3.2e-19	100
0.060	1.0000000000000000000 exp(±3.09953567303981773143**i**)	3.4e-19	2.3e-19	100
0.070	0.99999999999999999995 exp(±3.10468666836061660871**i**)	5.2e-19	2.5e-19	100
0.080	1.0000000000000000000 exp(±3.11168752182935715836**i**)	2.6e-19	3.0e-19	100
0.090	1.0000000000000000000 exp(±3.12245509887229756346**i**)	3.8e-19	5.1e-19	100
0.100	-0.985229910849008448355, -1.0149915151665093364 5	3.0e-19	2.7e-19	100
0.200	-0.908704733983905314765, -1.1004674704574738989 5	3.3e-19	3.6e-19	100
0.300	-0.859755912104609600457, -1.1631208182704841881 1	4.6e-19	3.2e-19	100
0.400	-0.819034502317305008415, -1.2209497855959510212 8	5.7e-19	2.2e-19	100
0.500	-0.784499412393390937147, -1.2746982141760295865 4	3.7e-19	2.5e-19	100
0.600	-0.755089133897168369623, -1.3243469613167347702 8	5.4e-19	4.7e-19	100
0.700	-0.729998333578531503093, -1.3698661407867753073 7	4.6e-19	3.4e-19	100
0.800	-0.708575796553274791698, -1.4112816227484765148 4	7.5e-19	3.1e-19	100
0.900	-0.690298554554738629245, -1.4486485498220798666 4	4.8e-19	3.7e-19	100
1.000	-0.674754590255442611413, -1.4820202995898536448 2	8.0e-19	2.9e-19	100
1.500	-0.629645514281034828278, -1.5881952262327431226 7	1.1e-16	7.8e-17	100
2.000	-0.638941822211735963705, -1.5650877204100365257 3	1.8e-09	1.1e-09	100
2.300	-0.721367120200611090905, -1.3862566950957593873 4	5.4e-08	2.4e-08	300

4.3.3. Bifurcations at resonance in elliptic tori. The reducibility method stops at the resonance because the Floquet transformation is not well adapted to the topology of the stable and unstable subbundles produced after the critical value ε_{crit}. Nevertheless, we will see that it is possible to continue the torus and its reducibility if we take into account that invariant bundles can appear, and therefore the torus can become normally hyperbolic.

We have to take into account the somewhat unexpected phenomenon that the invariant bundles generated by the resonance have an index which is related to the resonance. In particular, they may be nonorientable.

The key idea is to use the rotating transformations of section 2.2.6. To apply such a transformation it is not necessary to wait until the resonance is more evident (say, to values very close to ε_{crit}). In fact, our experiments suggest that it is better numerically to do this transformation before the transition happens. In Table 4 we display the results when doing the transformation for $\varepsilon = 0.050$, quite far from $\varepsilon_{crit} \simeq 0.0963489$.

We see that the method continues the reducibility of the normally hyperbolic torus, whose cocycle has a spectrum composed by two circles, until there is a new transition near $\varepsilon \simeq 2.345$. Again, reducibility is lost and the method stops. Notice, however, that in this case, the torus is nonuniformly normally hyperbolic. The spectrum grows and it is a full annulus enclosed by two circles, and the Oseledec bundles are not continuous.

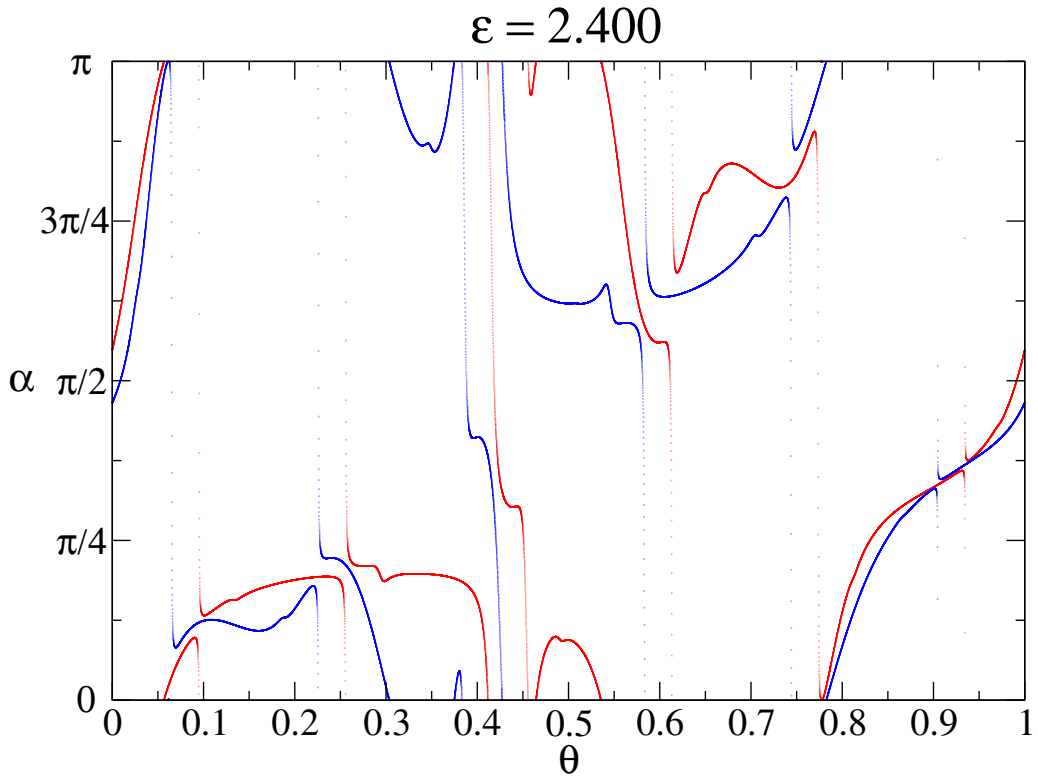


Figure 12. *Attractor (red) and repellor (blue) of the projective cocycle of the invariant torus* $\{x = 0, y = 0\}$ *of the rotating standard map with* $K = 0.2$ *and* $\varepsilon = 2.400$.

Notice, moreover, that reducibility holds in the small bumps between 2.345 and 2.697, and we can do a direct analysis by using projectivization. We can compute the invariant bundles using the power method [HdlL05a] or by discretizing the transfer operator [Jor99].

4.3.4. Topological evidence of bundle collapse (II). The invariance of the index of the bundles under homotopies has some consequences. By observing at some ends, we may conclude that something has happened in the middle.

For example, we observe that the torus is normally hyperbolic for $\varepsilon = 2.400$, with resonance $4 : -5$, so its invariant manifolds have index $\frac{-5}{2}$ and are nonorientable (see Figure 12). Since for $\varepsilon = 2.340$ the torus is normally hyperbolic and the invariant manifolds have index $\frac{1}{2}$, the torus cannot be normally hyperbolic for all the values between $\varepsilon = 2.340$ and $\varepsilon = 2.400$. Since the Lyapunov multiplier is bounded away from 1 (see Remark 4.2), the only possibility is that the stable and unstable bundles merge at some values of the parameter ε between 2.340 and 2.400.

Notice also that the torus is again normally hyperbolic for $\varepsilon = 0.4$, and the index of the invariant bundle in this case is 1, so that there are further transitions.

4.4. Nonorientable whiskers generated at resonances. We note that the rotating transformations introduced in section 2.2.6 lead to the existence of nontrivial invariant bundles in the linearization of the invariant circle $\{x = 0, y = 0\}$ in the standard map. In particular, we have found cases in which those bundles are nonorientable.

As remarked before, the methods of this paper can compute invariant manifolds modeled in these invariant bundles for the linearization.

Figure 13 shows the stable and unstable manifolds of the invariant torus $\{x = 0, y = 0\}$ for the rotating standard map, with $K = 0.2$, $\varepsilon = 0.5$.

4.5. Continuation of an invariant torus until breakdown. In this section we consider the continuation of the elliptic 3-periodic orbit of the standard map with $K = 0.2$.

Using the reducibility method, we detect a strong resonance for the value

$$\varepsilon_{crit} \simeq 0.06078370936559205612,$$

for which the rotation number is

$$\gamma \simeq -0.112219351279490344844.$$

In fact,

$$\left| \frac{\gamma}{2\pi} - \frac{k_1 + k_2 3\omega}{2} \right| \simeq 1.175787\ 10^{-9}$$

for $k_1 = 5$, $k_2 = -2$, so the resonance is of the type $5 : -2$. This suggests the transition to a normally hyperbolic torus whose subbundles have index $\frac{-2}{2} = -1$, so they are orientable. We can cross the resonance and compute the unstable and stable subbundles of the normally hyperbolic torus using the rotating transformation R_2 (see section 2.2.6).

The results of the full continuation are displayed in Figure 14 (Lyapunov multipliers and rotation number), Figure 15 (θx projection of the 3-periodic torus), and Figure 16 (dynamics of the projective cocycle).

- From $\varepsilon = 0$ to $\varepsilon \simeq 0.061$ the torus is elliptic (outside of small resonance intervals that are not detected numerically with the continuation step 0.001), and the dynamics of the projective cocycle is ergodic. See Figure 16 for $\varepsilon = 0.050$.
- From $\varepsilon \simeq 0.061$ to $\varepsilon \simeq 0.117$ the torus is normally hyperbolic and the invariant subbundles have index -1. See Figure 16 for $\varepsilon = 0.100$. Moreover, the eigenvalues are negative. For instance, for $\varepsilon = 0.100$, the eigenvalues are

$$-0.990944333,\ -1.0091384215.$$

- From $\varepsilon \simeq 0.117$ to $\varepsilon \simeq 0.173$ the torus is again elliptic. See Figure 16 for $\varepsilon = 0.150$.
- From $\varepsilon \simeq 0.173$ to $\varepsilon \simeq 0.350$ the torus is normally hyperbolic and the invariant subbundles again have index -1. See Figure 16 for $\varepsilon = 0.200, 0.250, 0.300, 0.350$. The eigenvalues are again negative. For instance, for $\varepsilon = 0.200$, the eigenvalues are

$$-0.963855575,\ -1.0374998344.$$

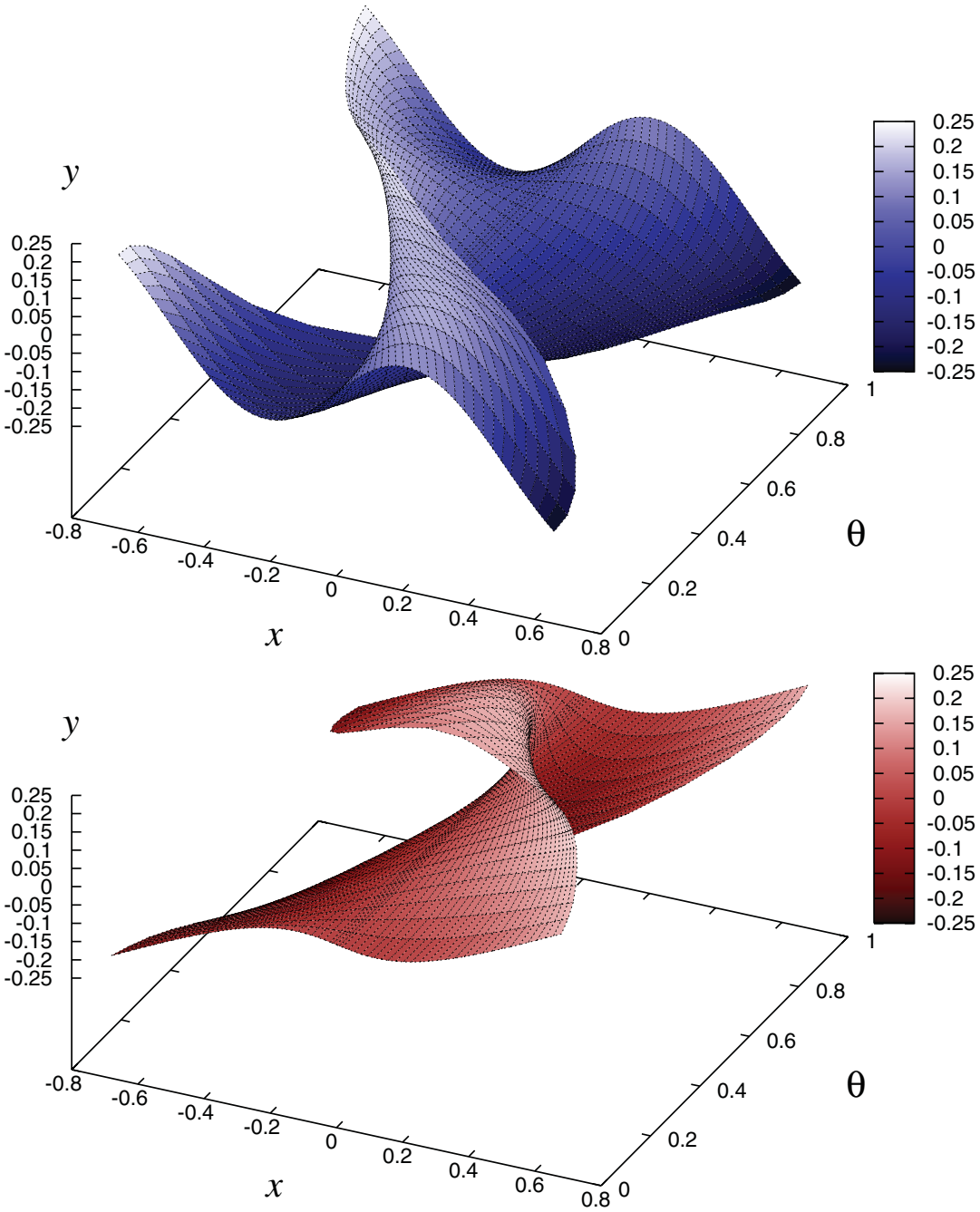


Figure 13. *3D views of the stable (blue) and unstable (red) manifolds of the curve* $\{x = 0, y = 0\}$ *for the rotating standard map with* $\kappa = 0.2$ *and* $\varepsilon = 0.5$. *Both manifolds are nonorientable.*

We observe also a similar phenomena to that in section 3.2. After the second strong resonance (that we can continue with reducibility because we adapted the coordinates to the topology already in the first resonance, and both of them are $5 : -2$), the torus is a saddle-type

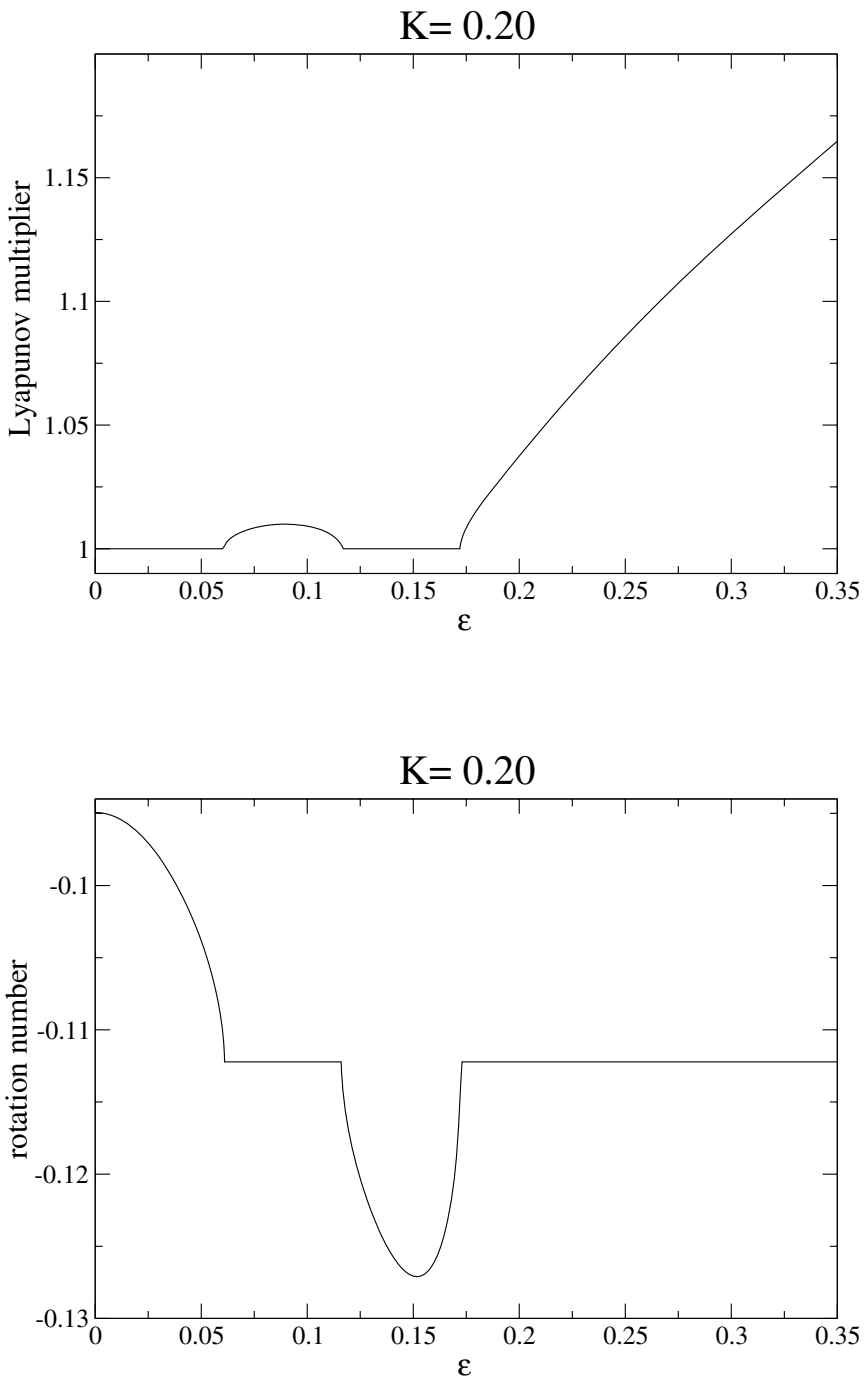


Figure 14. *Lyapunov multiplier and rotation number of the cocycle associated to a 3-periodic torus for the rotating standard map, with* $K = 0.2$.

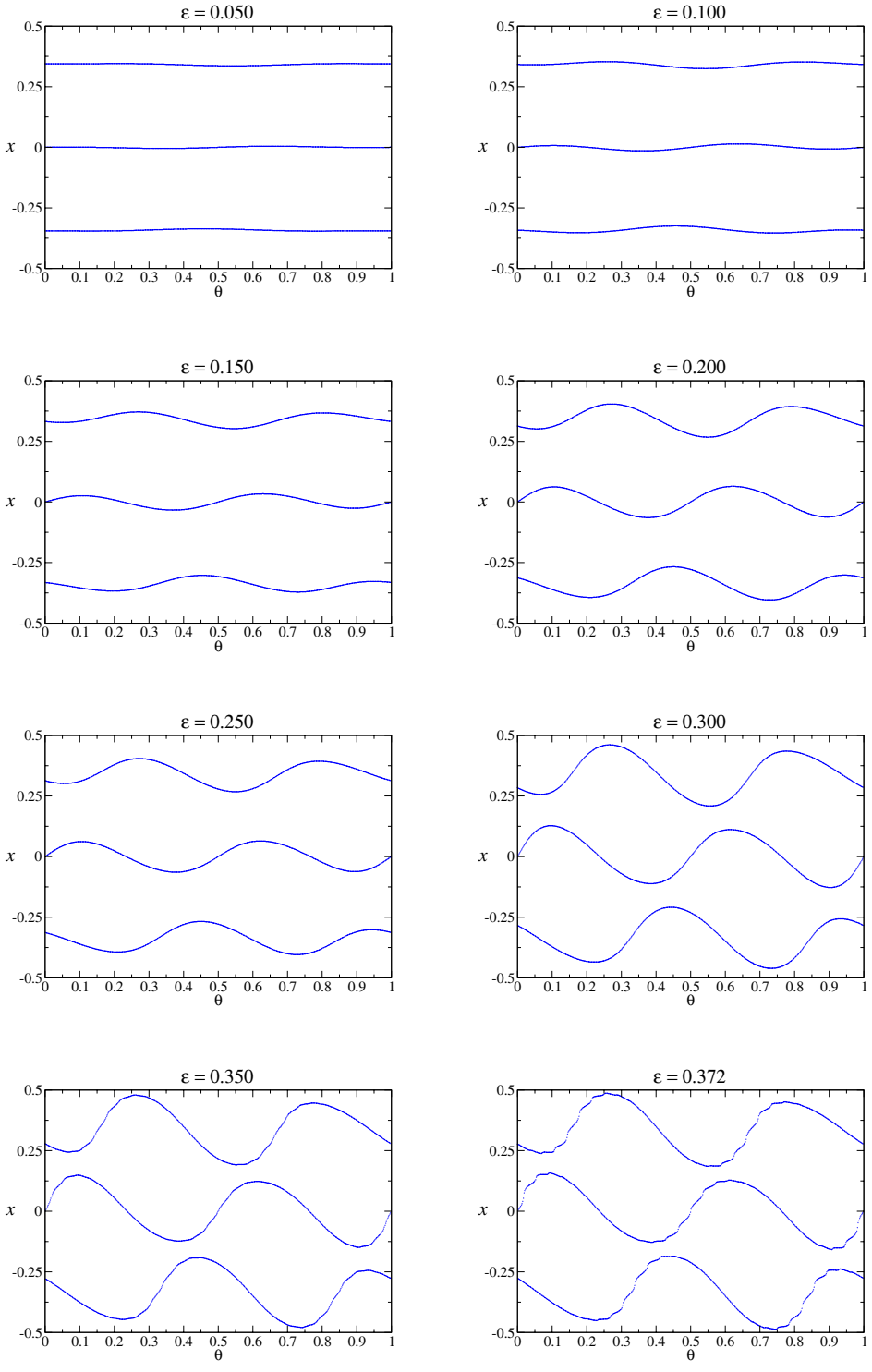


Figure 15. *θx projection of a 3-periodic curve of the rotating standard map, $K = 0.2$.*

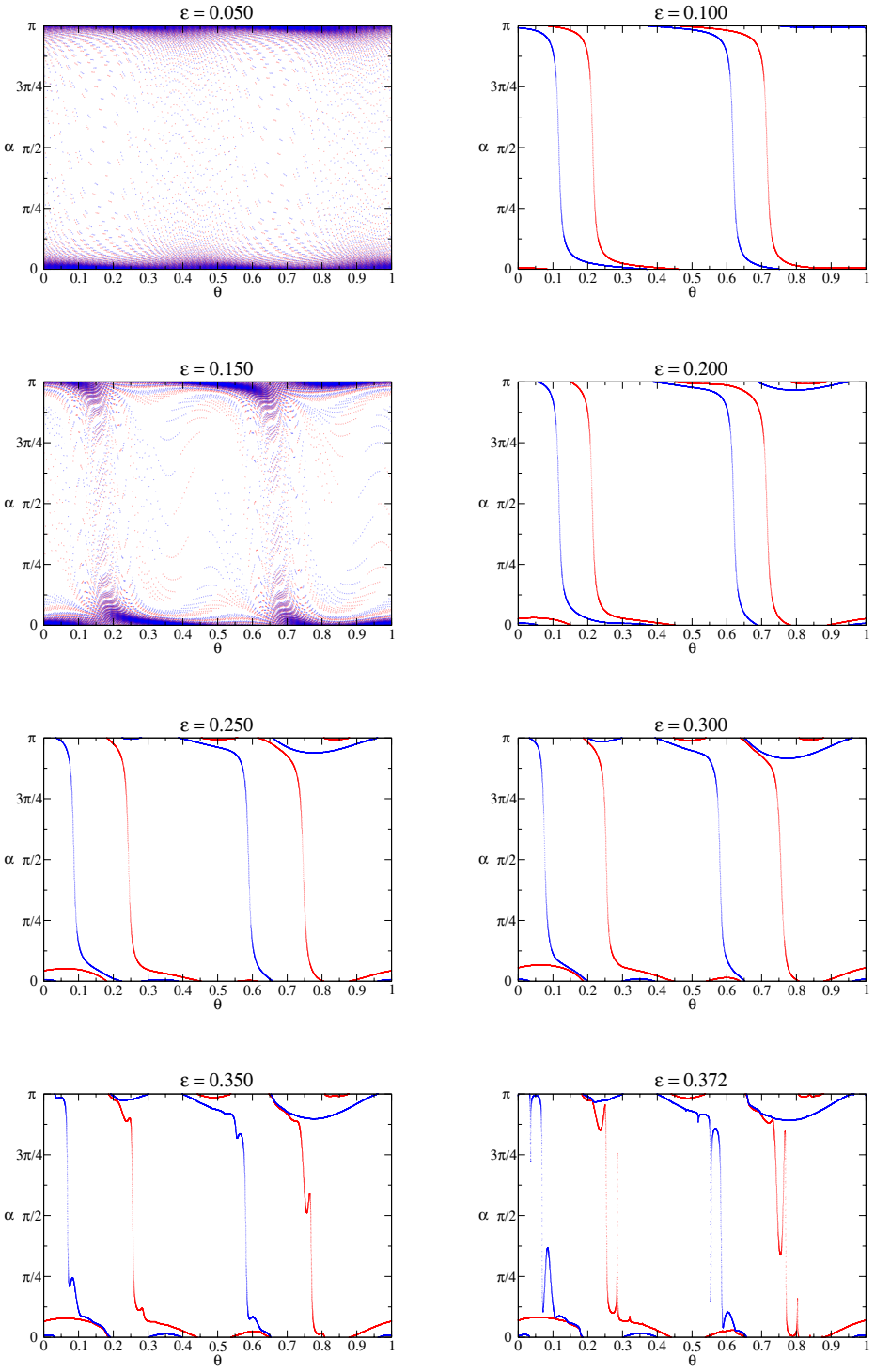


Figure 16. *Attractor (red) and repellor (blue) of the projective cocycle associated to a 3-periodic torus of the rotating standard map,* $K = 0.2$.

circle. But the unstable and the stable subbundles get closer and closer when increasing the parameter ε.

We have been able to continue the torus and its reducibility until $\varepsilon = 0.339$. Since the bundles start to fold a lot, they have to be approximated by Fourier expansions with about 10000 harmonics. Notice, however, that it suffices to use 250 harmonics to approximate the invariant torus.

So, in this regime we have changed to use the full-matrix Newton method. We have continued the torus until $\varepsilon = 0.370$, and the distance between the attractor and repellor of the projective cocycle is of the order of thousandths (see Figure 16). However, the Lyapunov multipliers remain far away from 1.

We can again presume that the stable and unstable bundles touch at a critical value near $\varepsilon = 0.370$, implying loss of reducibility. See Figure 15 for $\varepsilon = 0.370$.

The difference from the example in section 3.2 is that in this case, the spectrum at the collapse of the bundles is a fat annulus that contains the unit circle. So, certainly, the torus is not normally hyperbolic. The theory of [Mn78] suggests that we should expect that the torus breaks down. Certainly, we have found it impossible to continue the Newton method.

So, the continuation reaches $\varepsilon = 0.377950$, where the discretization of the torus has 1200 Fourier harmonics, and the invariance equation is solved up to an error $\simeq 10^{-9}$ (notice that the large matrix in the Newton step is about 264 Mbytes). The torus looks rather irregular and like it is about to break. See Figure 23. Moreover, the stable and unstable bundles are extremely close. A more detailed study of the phenomenon pointed out here will be undertaken in section 6.

In this example, since the torus seems to be smooth and reducible all the way to breakdown, it is possible to make reliable computations reasonably close to breakdown. Recall that a saddle torus is as smooth as the system [HdlL06b], and in this case it is analytic.

Notice, however, that the reducibility method fails because the bundles become so folded that their approximations require very high order Fourier expansions. In a sense, reducibility is much more sensitive to breakdown than the torus itself.

5. Example 3: A Lipschitz rotating standard map. In this section we present an example in which the use of the reducibility method or the projection method is mandatory to compute invariant tori with a high precision.

The example is a quasi-periodic perturbation of the standard map which is Lipschitz but not C^1. From the numerical point of view, the interest of this map is that the Fourier coefficients of the parameterization of the torus decrease rather slowly with the index, and it is, therefore, a good stress test for the efficiency of the methods. It can be considered as a stand-in for the computation of higher dimensional invariant tori (these computations are actively pursued).

The model we consider is similar to the rotating standard map (27), but the perturbation is given by the 1-periodic and even function defined by

$$c(\theta) = \begin{cases} 1 + 4\theta \text{ if } -\frac{1}{2} \le \theta \le 0, \\ 1 - 4\theta \text{ if } 0 \le \theta \le \frac{1}{2} \end{cases}$$

in the interval $[-\frac{1}{2}, \frac{1}{2}]$. The Fourier series of this Lipschitz function is

$$c(\theta) = \frac{8}{\pi^2} \sum_{m \geq 0} \frac{1}{(2m+1)^2} \cos(2\pi(2m+1)\theta).$$

Notice that the harmonic $k = 2m+1$ decreases as $\frac{1}{k^2}$. Henceforth, to approximate $c = c(\theta)$ by its Fourier series with an error smaller than, say, 10^{-8}, we need an expansion of at least 10^4 harmonics. This means that to compute the invariant tori with this error, we need, at least, to compute $4 \cdot 10^4$ coefficients.

The Newton method involves then solving linear systems of such a size at each step. Using quadruple precision, the size of the real variables is 12 bytes, and a matrix of such a size is about 18 Gbytes. So, the direct application of the Newton method is not feasible with a personal computer from 2004.

The reducibility method involves small divisor equations, whose justification needs some regularity to be solved at each step. Nevertheless, we can empirically use the numerical algorithm of reducibility even in regions where the mathematical justification is not yet available. We will obtain approximations of the stable and unstable subbundles that can be validated by the rigorous results in [HdlL06b]. This is one example of a situation where these validation results can be used to justify a posteriori numerical results obtained by algorithms that, by themselves, do not have proofs of convergence.

We have again continued the saddle-type 2-periodic point of the standard map with $K = 0.2$ into a saddle-type invariant torus, when coupling with θ. The results are displayed in Table 5 and Figure 17. In the computations we have used Fourier expansions with 5000 harmonics.

In order to check how the Fourier coefficients of the curve decrease, we have refined the torus corresponding to $\varepsilon = 0.2$, computing the first 100000 harmonics. The harmonics decrease of course very slowly, as can be checked in Figure 18. In fact, they decrease like $\frac{1}{k^2}$, as should be expected. Notice also the fine structure of the Fourier spectrum and its organization in curved stripes. We have no explanation for this curious phenomenon, nor any guess on its relevance.

6. The bundle merging scenario for breakdown of hyperbolicity and exponential dichotomies.

6.1. The bundle merging scenario and its consequences. In our previous studies of the rotating Hénon map (see section 3.2) and the rotating standard map (see section 4.5) we have encountered a possible scenario for the disappearance of exponential dichotomies and, in particular, for the disappearance of normal hyperbolicity. In this section, we will investigate the bundle merging in more detail and report some numerical regularities observed. In particular, we will report that there are scaling laws with universal exponents. These scaling relations were already reported in [HdlL06a]. We also mention the paper [JO05] which qualitatively observes situations where the invariant bundles approach each other.

In the bundle merging scenario, the Lyapunov multipliers of the cocycle of the torus remain different, but the bundles corresponding to them approach each other. One can conjecture that, at a critical value, the separation between the invariant bundles disappears and the collision is produced in a dense set of points.

Table 5

Continuation of a saddle 2-periodic torus for the C^0 rotating standard map, using the reducibility method. The curves are discretized in Fourier series with 5000 harmonics. The whole computation took around 3 hours on a Pentium III at 900 MHz.

C^0 rotating standard map: $K = 0.200$

ε	Eigenvalues	R	S	T	NF	tail
0.000	8.18934863578967502551e-01, 1.22109833696629880250e+00	2.7e-20	0.0e+00	0.0e+00	5000	0.0e+00
0.100	8.09541495085049592261e-01, 1.23526713092692232068e+00	8.6e-16	1.7e-10	2.3e-10	5000	2.2e-09
0.200	7.87128971548113161825e-01, 1.27043983406329795101e+00	1.8e-12	1.7e-08	2.3e-08	5000	3.4e-09
0.300	7.60884403244821979273e-01, 1.31426008436427392957e+00	2.3e-15	2.3e-12	2.8e-12	5000	4.1e-09
0.400	7.34505268365048588587e-01, 1.36146062263986463749e+00	4.0e-12	2.5e-07	3.7e-07	5000	4.6e-09
0.500	7.08169034538036472771e-01, 1.41209224242957117126e+00	1.4e-15	1.3e-10	1.8e-10	5000	5.2e-09
0.600	6.81232590662857064208e-01, 1.46792742112793467594e+00	1.7e-14	5.1e-09	7.8e-09	5000	5.8e-09
0.700	6.53248421318634461060e-01, 1.53081120040292833898e+00	2.5e-14	1.0e-08	1.7e-08	5000	6.6e-09
0.800	6.24238876424479131302e-01, 1.60195085209659427794e+00	3.3e-15	1.6e-09	2.5e-09	5000	7.3e-09
0.900	5.94620947767326503343e-01, 1.68174364484598861693e+00	1.6e-15	2.7e-11	3.0e-11	5000	8.0e-09
1.000	5.64982000445246894469e-01, 1.76996789138764902181e+00	3.0e-14	2.5e-08	4.5e-08	5000	8.6e-09
1.100	5.35877608687003032602e-01, 1.86609775029602792222e+00	2.4e-12	8.4e-08	1.6e-07	5000	9.1e-09
1.200	5.07729974230162079091e-01, 1.96955084543952656391e+00	6.1e-14	4.6e-09	8.9e-09	5000	9.5e-09
1.300	4.80812167097835163306e-01, 2.07981425685816574429e+00	1.3e-13	1.6e-07	3.3e-07	5000	9.7e-09
1.400	4.55273722019797196588e-01, 2.19648082380760632424e+00	1.5e-14	1.4e-11	9.0e-12	5000	9.9e-09
1.500	4.31175411189466501046e-01, 2.31924171473818288482e+00	2.7e-14	1.2e-11	9.2e-12	5000	1.0e-08
1.600	4.08519219409532413135e-01, 2.44786524718563925458e+00	4.2e-14	8.8e-12	6.7e-12	5000	1.0e-08
1.700	3.87270327364771600764e-01, 2.58217562601457775533e+00	5.7e-14	5.7e-12	3.8e-12	5000	9.9e-09
1.800	3.67372076410609758748e-01, 2.72203595267895417649e+00	6.7e-14	3.9e-12	2.1e-12	5000	9.8e-09
1.900	3.48755792337401190764e-01, 2.86733588939210017619e+00	2.8e-13	1.5e-07	4.3e-07	5000	9.7e-09
2.000	3.31347116084920405906e-01, 3.01798311033053172901e+00	2.2e-13	6.7e-08	2.0e-07	5000	9.6e-09

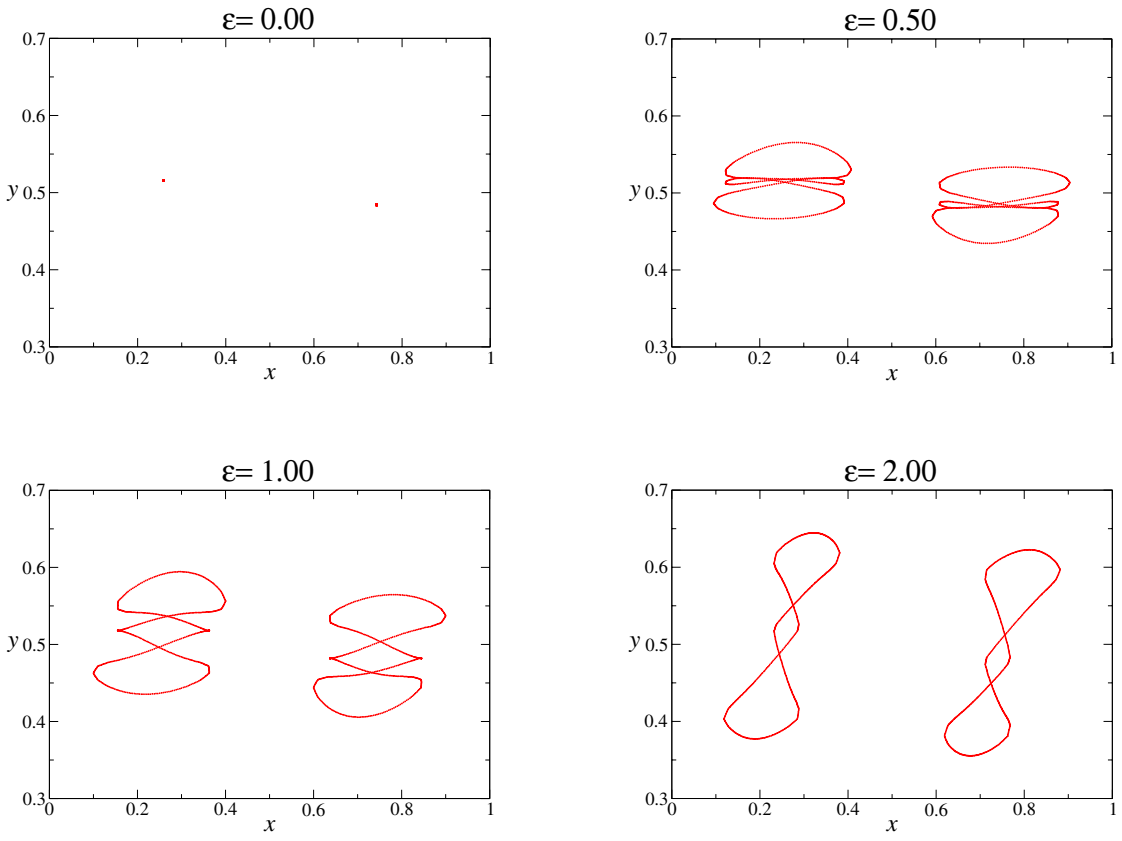


Figure 17. *Continuation of the 2-periodic saddle curve for the C^0 rotating standard map, $K = 0.20$. Projection on the xy plane.*

Note that the bundle merging scenario implies several consequences.

- Up to the critical value we have the following:
  - The maximal and minimal Lyapunov multipliers are different from each other and from 1.
  - The spectrum consists of two circles, whose radii are the Lyapunov multipliers.
  - The bundles associated to these circles are 1D because we are working with 2D maps.
  - The bundles are as smooth as the cocycle; in this case they are analytic.
  - As a consequence, for the examples considered (in which the external rotation ω is Diophantine), linearization is reducible to constants.
- As the parameter approaches the critical value, the following occur:
  - The Lyapunov multipliers remain bounded away from each other and from 1.
  - The two bundles get closer, and their minimum distance approaches zero.
  - As a consequence, the bundles become more and more oscillatory. Even if they remain analytic, the first derivatives become unbounded.

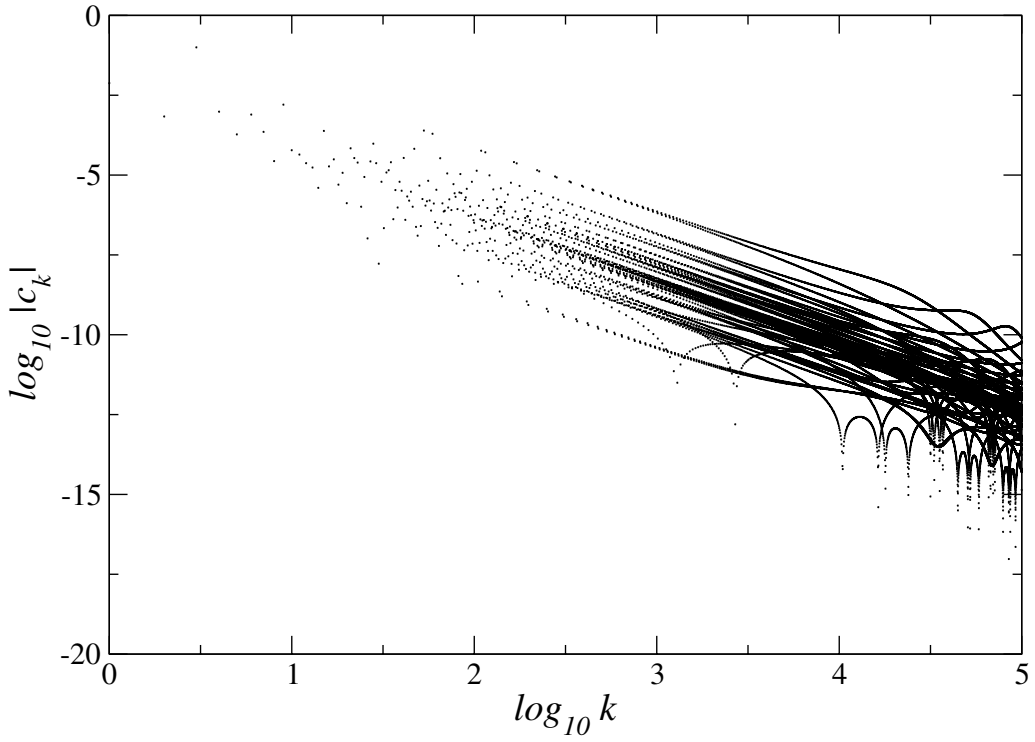


Figure 18. *log-log table of the* 100000 *Fourier modes of the x-component of the saddle-type* 2*-periodic curve for the* C^0 *rotating standard map, with* $K = 0.2$, $\varepsilon = 0.2$. *The errors are* $R = 3.9e\text{-}14$, $S = 1.5e\text{-}11$, $T = 1.3e\text{-}11$.

  - Even if the spectrum of the transfer operator remains being circles that are far apart, the spectral projections over the corresponding spectral subspaces have norms that become unbounded. (Note that the norm of the spectral projections is closely related to the minimum distance between the bundles.)
- At the breakdown and afterward, it is impossible to discern any continuous invariant bundle.
  - As a consequence, the spectrum of the transfer operator at the critical value (and afterward) consists of a filled annulus. The inner and outer boundaries are circles whose radii are the Lyapunov multipliers.
  - The invariant bundles are not continuous. The Oseledec theorem [Ose68] asserts that the bundles are defined in a set of full measure, and they are measurable.

Remark 6.1. At first sight it may seem paradoxical that the bundles approach each other at some points but the Lyapunov multipliers remain different.

It is certainly true that, when the bundles are very close at some point, the Lyapunov multipliers at a fixed time become very close. Nevertheless, remember that the Lyapunov multipliers are obtained by taking the limit of the number of iterations tending to ∞ so that

the proximity for finite time does not imply the proximity of the limits. (In other words, we cannot exchange the limits of the parameters or the $n \to \infty$ in the expresion of the time n Lyapunov exponents.)

Note that the formula (16) is affected little by intervals of small length in which the integrand takes the "wrong" value.

It is true that the computation of the Lyapunov exponent requires longer and longer times to achieve a certain accuracy as we approach the critical value.

Remark 6.2. At the critical value of the parameter, the spectrum of the transfer operator experiences a sudden growth and it is discontinuous as a set. It is well known that the spectrum cannot suddenly decrease [Kat76], but in the same reference, we can find other examples with sudden growth. This phenomenon of sudden growth of the spectrum is closely related to the fact that the spectral projections become unbounded.

Remark 6.3. Notice also that, even if Oseledec bundles are not continuous at a critical value, some kind of regularity could be expected (e.g., Whitney differentiability). This is perhaps what is observed in Figures 5 and 6 (see also Figures 20 and 21). For studies on the regularity of Oseledec bundles, see [Pes77], and for regularity of invariant graphs on forced systems (that here correspond to the projectivization of invariant bundles), see [Sta99].

6.2. Quantitative regularities of bundle collapse. The observables we consider to measure the bundle collapse are the minimum distance between the bundles (Δ), which goes to zero at the collapse for $\varepsilon = \varepsilon_c$, and the upper and lower Lyapunov multipliers ($\Lambda^{\pm}$), which go to critical values $\Lambda^{\pm}$ when ε goes to ε_c.

The most remarkable fact about this transition is that there seem to be some scaling relations of these observables with exponents that are not affected by changes in the map.

More precisely, we formulate the following assertion.

Assertion 6.4. *For an open set of families f_ε, we have the following.*

The observables Δ and $\Lambda^{\pm}$ satisfy the asymptotics

$$\Delta_\varepsilon \sim \alpha(\varepsilon_c - \varepsilon)^\beta, \quad \Lambda^{\pm}_\varepsilon \sim \Lambda^{\pm}_c + A^{\pm}(\varepsilon_c - \varepsilon)^B, \qquad \varepsilon \lesssim \varepsilon_c, \tag{33}$$

where $\alpha, \beta, A^{\pm}, B$ (and $\Lambda^{\pm}_c$) are fitting parameters.

We distinguish two cases:

(a) *The Lyapunov multipliers do not straddle* 1*: $\beta = 1$, $B = 0.5$.*

(b) *The Lyapunov multipliers straddle* 1*: $\beta = 1$, $B = 1$.*

We emphasize that in case (b) above the torus is (generically) destroyed, because the stable and unstable bundles collapse. In case (a), if the torus is attracting, it survives after the collapse of the fast and slow stable bundles in the critical value ε_c. Then, we can also make estimates of the observables after this critical value. Our numerical computations suggest the following.

Assertion 6.5. *For an open set of families f_ε in case* (a) *above, we also have the approximations*

$$\Delta_\varepsilon \approx 0, \quad \Lambda^{\pm}_\varepsilon \approx \Lambda^{\pm}_c + \bar{A}^{\pm}(\varepsilon - \varepsilon_c)^{\bar{B}}, \qquad \varepsilon \gtrsim \varepsilon_c, \tag{34}$$

where $\bar{B} = 1$.

Remark 6.6. For comparison, we recall that known bifurcations—which have analogues for quasi-periodic tori—have scaling relations. See [BHTB90]. The saddle node has $\beta = 0.5$, $B = 0.5$, $\Lambda_c^{\pm} = 1$, and the transcritical has $\beta = 1$, $B = 1$, $\Lambda_c^{\pm} = 1$.

Notice, however, that, in the bifurcations considered in [BHTB90], the bundles remain smooth. This is very different from what happens in the problems considered here.

Remark 6.7. We already noted that if the Lyapunov multipliers, whose associated bundles collapse, do not straddle 1, then the torus continues being normally hyperbolic and survives after the collapse. (In the example presented here, the torus is attracting. See section 3.)

If the Lyapunov multipliers do straddle 1, the torus is not normally hyperbolic at collapse, and a generic perturbation destroys the torus [Mn78]. (See section 4.5.)

Hence, there is an important difference between both cases, which explains the different scalings found.

Notice, however, that one can create an example in which a saddle torus survives after the bundle collapse because the perturbation vanishes on the torus. (See section 4.3.) In this case we found the scalings to be those in (a) of Assertion 6.4.

Remark 6.8. While Assertion 6.4 conjectures asymptotic formulas for the observables, Assertion 6.5 gives only approximations. We also note that the numerics of Assertion 6.4 can be validated using the well-developed theory of normally hyperbolic systems.

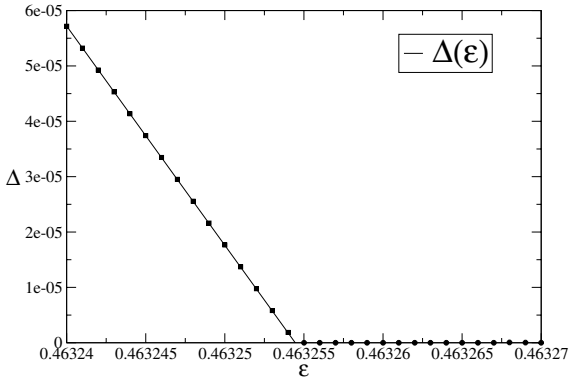
Notice, however, that it is not easy to assess the reliability of the calculations of Assertion 6.5, since the exponential dichotomies have been broken. It seems quite possible that there are gaps of parameter values where Δ, even smaller than the precision of the present computation, is nevertheless positive. A "big" gap $[\varepsilon_{\mathtt{c}}, \varepsilon_{\mathtt{d}}]$ has been already observed in section 3.1. See section 6.3.4 for other "small" gaps. One can conjecture that the fact that increasing the precision reveals more structure will continue. This is consistent with the mathematical work in [AK03].

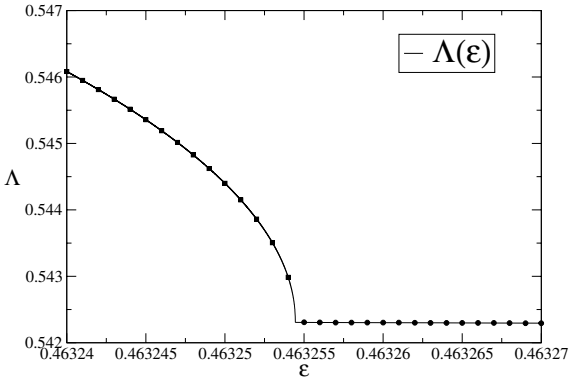
6.2.1. Results for the rotating Hénon map. As evidence for Assertions 6.4 and 6.5, we present some more detailed computations done in the quasi-periodic Hénon map (25), with $a = 0.68$, $b = 0.1$, and $\omega = (\sqrt{5} - 1)/2$. Similar computations have been carried out for some modifications of the map (changing parameters, adding terms, and even taking other rotations—for example, $\omega = e/4$).

Below is a description of the results of detailed calculations at the bifurcation `(b)` described in section 3.1. Similar calculations have been performed for the bifurcations `(c)` and `(d)` in section 3.1.

We have computed the observables Δ—minimum distance between the bundles—and Λ—maximum Lyapunov multiplier—of the attracting circle for values of ε in $[0.463240, 0.463270]$, and fit parameters in (33), (34). We note that, in this case, since the circle is an attractor, the computation can be done without using Fourier methods. It suffices to iterate the map to compute the attractor up to round-off error precision. Once we have the attractor, we iterate the projectivized version of the linearization (15) both forward and backward as indicated in section 2.2.1.

We note that, in this case, the torus, being a uniform attractor, remains very smooth (in fact, as smooth as the system, which in our case implies that the torus is analytic). The calculation is done always in the situation where the torus is smooth and the calculation is quite reliable. Then, the values are extrapolated.





$\varepsilon_c = 0.46325447112 \pm 1 \cdot 10^{-11}$
$\alpha = 3.94933 \pm 7 \cdot 10^{-5}$
$\beta = 0.999979 \pm 2 \cdot 10^{-6}$

$\Lambda_c = 0.5423122 \pm 5 \cdot 10^{-7}$
$A = 1.015 \pm 1 \cdot 10^{-3}$
$B = 0.5020 \pm 1 \cdot 10^{-4}$
$\bar{A} = -0.7409 \pm 6 \cdot 10^{-4}$
$\bar{B} = 1.00035 \pm 8 \cdot 10^{-5}$

Figure 19. *Fits of* Δ *and* Λ *to* (33), (34) *for the transition* $\mathtt{b}$ *that is produced between* $\varepsilon = 0.463254$ *and* $\varepsilon = 0.463255$.

Table 6

Maximal Lyapunov multiplier of the attractor of the rotating Hénon map, and estimate of the error, for different values of ε *close to the critical value* $\varepsilon_c = \varepsilon_{\mathtt{b}}$. *The computations have been done by direct iteration of the cocycle on a vector* 701408733 *times, after a transient time of* 100000 *steps.*

ε	Λ	Error
0.4632544710	0.5423194946536124	1.0e-13
0.4632544711	0.5423149288182757	1.0e-13
0.4632544712	0.5423065545078773	5.2e-10
0.4632544713	0.5423065555707195	5.8e-10
0.4632544714	0.5423065550474250	1.0e-09

The results of the fits are depicted in Figure 19. Notice that we can extrapolate the value of the transition: $\varepsilon_{\mathtt{b}} \simeq 0.4632544711$. See Table 6 for further numerical evidence, that is, the change of the accuracy in the computation of the Lyapunov multipliers. The attractors and their bundles are represented in Figure 20. A magnification of the invariant bundles right before collapse is depicted in Figure 21. Note that in Figure 21 we find a significant difference by changing the parameter value of ε by just 10^{-6}.

Among many other verifications of universality we have carried out, we present in Figure 22 an analogue of Figure 19 when $\omega = e/4$. Note that even if $e/4$ is numerically close to $(\sqrt{5}-1)/2$, its number theoretical properties are very different.

6.2.2. Results for the rotating standard map. We now discuss the phenomena that are observed for the rotating standard map (27). The situation is somewhat different from the situation discussed previously. One obvious difference is that the map is area preserving. Also, since at collapse the spectrum of the linearization is a fat annulus that includes 1, we expect

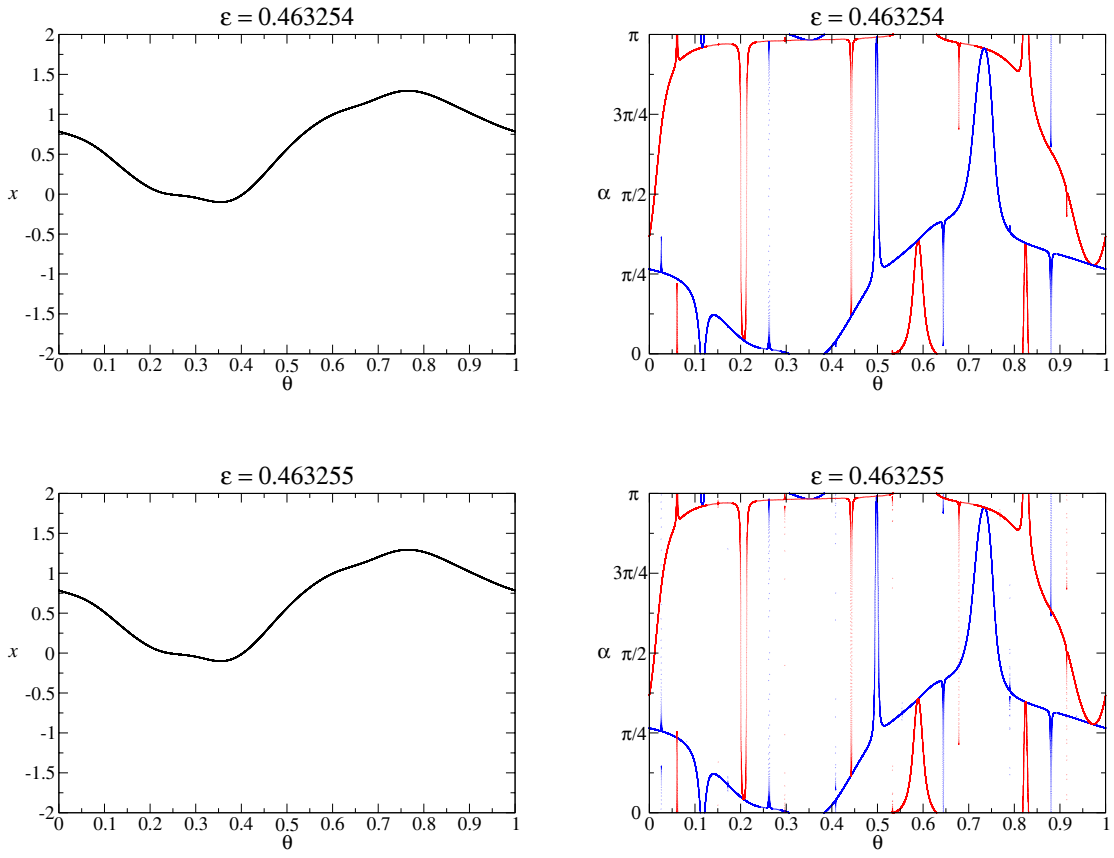


Figure 20. *x-curve of an attracting torus and α-curves of its slow and fast directions, before and after their collapse.*

that the invariant circle disappears at collapse.

In section 4.5 we have continued with respect to ε a 3-periodic torus. The torus becomes hyperbolic at $\varepsilon \simeq 0.173$ and remains hyperbolic for larger values of ε.

The continuation reaches $\varepsilon = 0.377950$, where the torus looks rather irregular and about to break, and the stable and unstable directions are extremely close. See Figure 23. Hence, we conjecture that this is another case of bundle collapse. Of course, the irregularity of the torus is in marked contrast with the situation for the rotating Hénon map discussed before.

We can explore the torus for values of ε close to this value. The results of fitting the scaling laws in Assertion 6.4 are depicted in Figure 24. Notice that we produce an estimate of the critical value in which the torus is destroyed: $\varepsilon_c \simeq 0.379696$.

Compared with the phenomena studied in the rotating Hénon map, we find several differences. The most notable one is that the torus does not seem to persist (at least as a smooth curve). Hence, there is no scaling relation after the bundle merging. Also we note that the torus becomes more irregular.

From the computational point of view, since the projection and reducibility methods

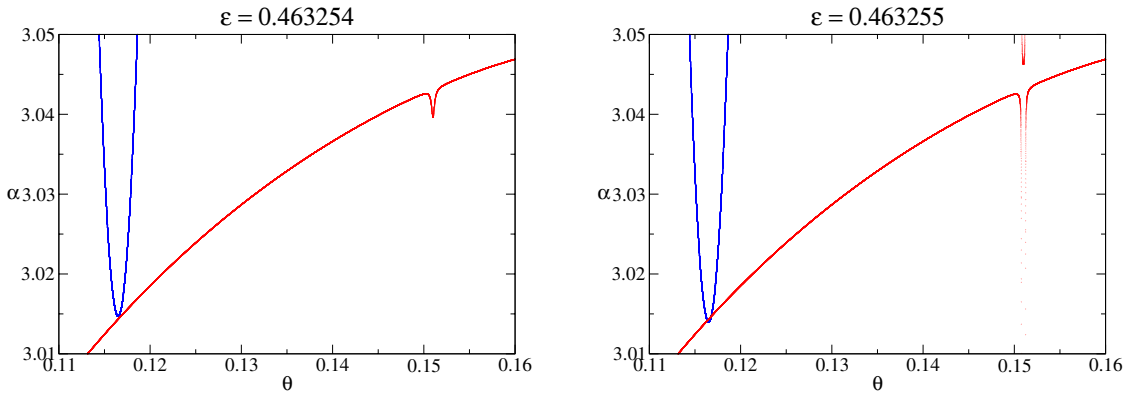
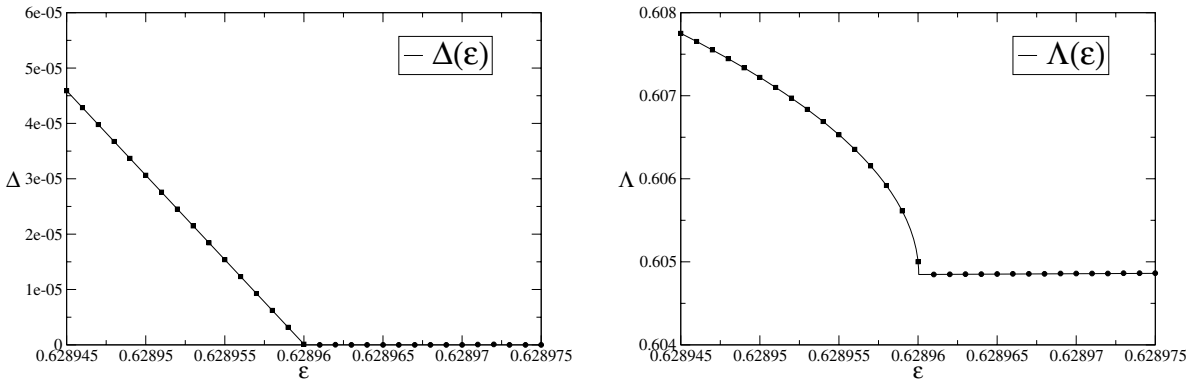


Figure 21. *Zooms of the α-curves of Figure 20.*



$\varepsilon_c = 0.62896004195 \pm 5 \cdot 10^{-11}$
$\alpha = 3.0546 \pm 3 \cdot 10^{-4}$
$\beta = 1.000098 \pm 9 \cdot 10^{-6}$

$\Lambda_c = 0.60484735 \pm 1 \cdot 10^{-8}$
$A = 0.7428 \pm 3 \cdot 10^{-4}$
$B = 0.4993 \pm 4 \cdot 10^{-5}$
$\bar{A} = 0.93 \pm 9 \cdot 10^{-2}$
$\bar{B} = 0.983 \pm 8 \cdot 10^{-3}$

Figure 22. *Fits of Δ and Λ to (33), (34) for a bundle merging bifurcation in an attracting torus of the rotating Hénon map with frequency $\omega = e/4$, and $a = 0.68$, $b = 0.1$. The transition is produced between $\varepsilon = 0.628960$ and $\varepsilon = 0.628961$.*

become more problematic as we approach the critical value, we have to use the full-matrix Newton method. Simple iteration is not possible because the torus is not attracting. As a consequence, we can only explore values of ε which are farther apart from the critical value than those allowed in the rotating Hénon map.

6.3. Some open questions. We conclude this article by mentioning several open questions that arise from the computations here.

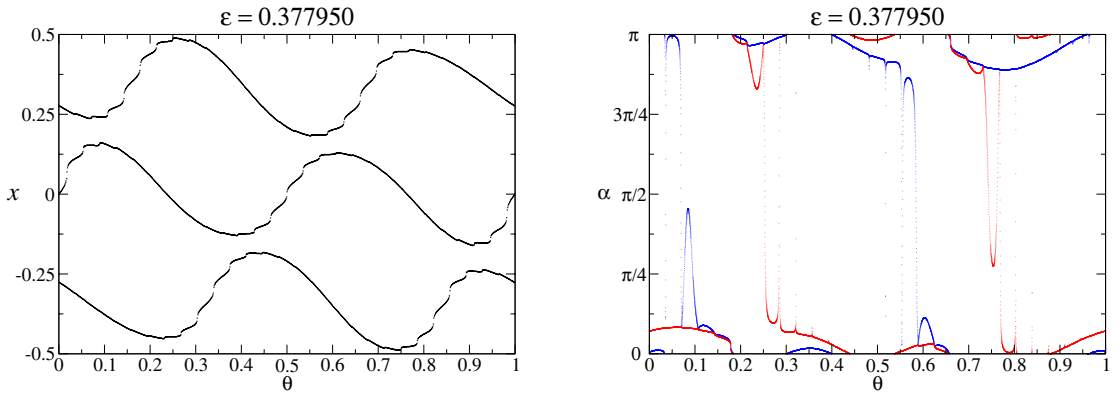
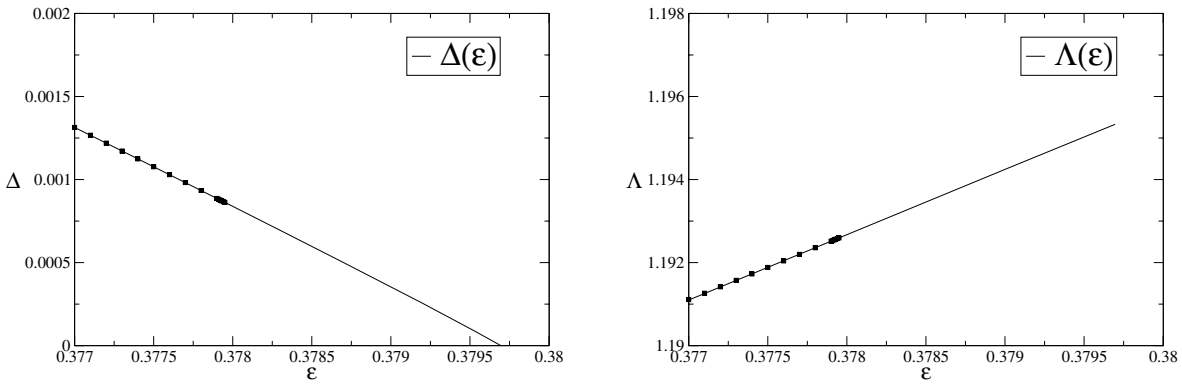


Figure 23. *x-curves of a 3-torus close to breakdown, and the α-curves of its unstable and stable directions.*



$\varepsilon_c = 0.3796965 \pm 8 \cdot 10^{-7}$
$\alpha = 0.4063 \pm 7 \cdot 10^{-4}$
$\beta = 0.9693 \pm 4 \cdot 10^{-4}$

$\Lambda_c = 1.19533 \pm 9 \cdot 10^{-5}$
$A = -1.6 \pm 2 \cdot 10^{-1}$
$B = 1.00 \pm 3 \cdot 10^{-2}$

Figure 24. *Fits of Δ and Λ to (33) near breakdown.*

6.3.1. A renormalization group description. The existence of the scaling properties above suggests that there could be a renormalization group explanation for the phenomena happening. Unfortunately, we have not been able to find a renormalization group transformation which has a fixed point which explains the behavior observed. Nevertheless we point out that renormalization of quasi-periodic cocycles has been considered in the literature in somewhat different contexts. For example, renormalization of Schödinger quasi-periodic cocyles appears already in [Hof76]. Renormalization of quasi-periodic maps has been considered in [OM03, MO00, MO04]. Renormalization of general cocycles of matrices and its relation with reducibility was considered in [Ryc92]. Renormalization procedures for 2D cocycles play an important role in [Kri99c, Kri01]. Nevertheless, the conclusions of later renormalizations are

not formulated in terms of fixed points of an operator and they seem to be in regimes different from the ones considered in this paper.

We think that developing a renormalization group explanation of the phenomena reported here is a very interesting problem. We note, however, that in contrast with other renormalization groups studied, the phenomenon seems to be independent of the rotation number.

6.3.2. Scaling exponents of the observables. The fact that the exponents found numerically are close to being simple numbers makes us hopeful that perhaps there is a relatively simple explanation. Nevertheless, since one has to deal with bundles that are becoming not very smooth it is not clear that a simple normal form explanation works.

We think that exploring whether there is a simple explanation of the scaling behavior reported here is a very interesting problem.

Remark 6.9. The square root behavior of the Lyapunov multiplier has also been observed in a growth of spectrum phenomena similar to that described above for a rotating logistic map that is a quasi-periodically forced noninvertible 1D map. In this case, the mechanism involves the loss of invertibility of the cocycle of the attractor, and it is possible to do a normal form analysis [JT05].

Besides the scaling properties reported in Assertions 6.4 and 6.5, it seems that there is a host of other scaling properties to be observed in this scenario.

We are actively pursuing these numerical explorations, but it is clear that much more needs to be done.

Of course, besides finding more observables that scale, an interesting problem is to find the domain of universality of these relations.

6.3.3. Explicit models of cocycles exhibiting the bundle merging scenarios. The bundle merging scenario is a property of the projective cocycle. Since the projective cocycle is independent of multiplication of the linear cocycle by a constant, it superficially seems that the distinction between the cocycle straddling 1 or not is meaningless.

On the other hand, we note that when the cocycle straddles 1, we expect that the torus disappears and that, therefore, as we approach the critical value, the torus becomes more oscillatory. Hence, the difference between the bundles straddling 1 or not is really a difference between the projective cocycles becoming more oscillatory or not.

We think that it would be very interesting to develop explicit models of linear cocycles that exhibit bundle merging without any reference to the underlying invariant torus. Given the above observations, one can expect that the bundle merging with the spectrum straddling 1 will require cocycles with rather irregular matrices.

6.3.4. Abundance of gaps with exponential dichotomy. In the study of the rotating Hénon map, we have encountered regions of the parameter space in which there are slow and fast stable smooth invariant bundles ($\Delta > 0$, reducible case) and others in which they are not smooth ($\Delta = 0$, nonreducible case). We have also found scaling relations in the transition from one to the other case (see Assertions 6.4 and 6.5).

Notice that $\Delta > 0$ holds in open sets of the parameter space. Even if we present numerical evidence for $\Delta \approx 0$ in an open set of the parameter space (see Assertion 6.5), the theory in [AK03] suggests that in fact $\Delta > 0$ in small gaps of the parameter space that are missed by

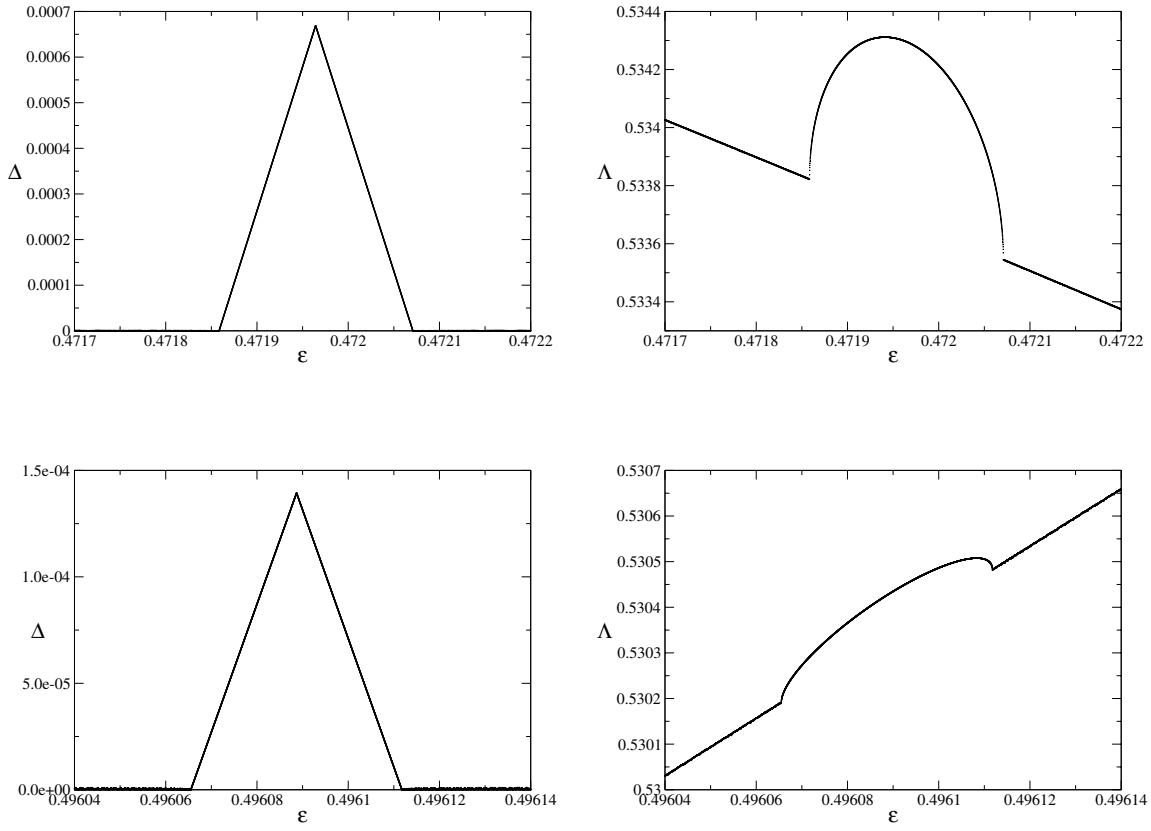


Figure 25. *Some small gaps of reducibility. The observables Δ (left) and Λ (right) in two intervals of the parameter ε for the Hénon map.*

the numerical computation. Some of these gaps for the example in section 6.2.1 are shown in Figure 25.

So, it would be very interesting to quantify the abundance of reducibility and nonreducibility in this context. It is quite plausible that reducibility holds in an open and dense subset of parameter space, but its structure can be very complicated. On the other side, nonreducibility could happen in a closed set of positive measure. (See [HP05] for a model in which this picture holds.)

6.3.5. Existence of strange nonchaotic attractors. The global phenomena that happen at e in section 3.1 are not well understood and they deserve a more detailed study. After the transition e the attractor looks like a strange chaotic attractor (SCA). It seems that this is related to the very poorly understood phenomena of SNAs, and in particular to the so-called *fractalization mechanism* or route "torus→SNA→SCA." (See [PNR01] for a review of experiments in this field. See also [GFPS00, OWGF01].) There are some precise definitions of SNA in the literature. See, for example, [Kan94, Kel96, WY01, WY02, BDV04].

The theory in [HdlL06b] and the numerical experiments described here suggest that the fractalization mechanism is in fact a route "torus→SCA" and that in this transition SNAs are not created. Further detailed numerical experiments of this transition confirm this description. (See [BSV05, JT05] for similar studies on the rotating logistic map that is 1D and noninvertible, [HdlL05b] for the rotating Hénon map, and [HS05] for a comprehensive study of both cases.) But, so far, there are not rigorous explanations of the fractalization mechanism.

We also note that, at the critical value of bundle merging, the attractors and repellors for the projectivization look like SNAs, as in the *collision mechanism* described in the literature [PNR01]. We have obtained evidence in several examples that the collapses are really produced based on topological arguments. Our interpretation is that the SNAs are the projectivization of the Oseledec bundles that are measurable but not continuous after the collapse. It would be quite interesting to decide whether the objects formed at the critical value of the bundle-merging bifurcation fit the rigorous definitions of SNAs. (See [HP05] for the implication of this mechanism in the existence of SNAs in Harper maps.)

Notice also that, in the experiments performed here, both mechanisms seem to be related. That is, the collision mechanisms that happen at the level of the linearization of the dynamics around the torus have a strong influence on the geometric behavior of the torus itself.

6.3.6. Extensions to other dynamical systems. We emphasize that the bundle merging scenario is a dynamical/geometrical/functional mechanism that can be generalized to higher dimensions, to other base dynamics, etc., that is, to general bundle maps (and their invariant sections).

The linearization around an invariant section is a vector bundle map in a suitable vector bundle (the normal bundle to the invariant section).

From the dynamical point of view, the mechanism is the progressive deterioration of the exponential dichotomies in the linearization (the uniform constant C in (20) tends to infinity in the collapse).

From the geometric point of view, it consists in the collapse of the invariant bundles in a nonsmooth way.

From the functional analysis point of view, it consists in a sudden growth of the spectrum of the transfer operator. This functional analysis behavior has different geometric implications depending on whether the rates involved in the collapse straddle 1 or not.

It would be interesting to obtain scaling laws in these transitions for cases other than rotations.

Extensions to fully coupled systems will be also very interesting.

After this paper was submitted, we became aware of [CLR06, CLR]. These papers contain a study of some regions where the Hénon family loses hyperbolicity. They show that the hyperbolicity may be lost even if the Lyapunov exponents of all invariant measures are bounded away from zero. The reason is there are points where the invariant bundles become close to each other. This is reminiscent of the bundle merging scenario discussed here. It would be quite interesting to study whether the rigorous analysis of the above papers applies to the examples here. Conversely, it would be quite interesting to find out whether the quantitative scaling laws found here apply to the internal tangencies. We also became aware of [Bjerklov05, Bjerklov06], which study linear cocycles, find phenomena reminiscent of the

bundle merging scenario, and are related to [HP05].

Acknowledgments. We thank R. Fabbri, À. Jorba, W. Liu, S. Luzzatto, J. Puig, and C. Simó for several discussions and relevant references. We thank the referee for a very thorough reading which improved the exposition.

REFERENCES

[ACS91] R. Artuso, G. Casati, and D. L. Shepelyansky, *Breakdown of universality in renormalization dynamics for critical invariant torus*, Europhys. Lett., 15 (1991), pp. 381–386.

[ACS92] R. Artuso, G. Casati, and D. L. Shepelyansky, *Break-up of the spiral mean torus in a volume-preserving map*, Chaos Solitons Fractals, 2 (1992), pp. 181–190.

[AK03] A. Avila and R. Krikorian, *Reducibility or non-uniform hyperbolicity for quasiperiodic Schrödinger operators*, Ann. of Math. (2), 164 (2006), pp. 911–940.

[AS83] J. Avron and B. Simon, *Almost periodic Schrödinger operators.* II. *The integrated density of states*, Duke Math. J., 50 (1983), pp. 369–391.

[AS82] J. Avron and B. Simon, *Almost periodic Schrödinger operators.* I. *Limit periodic potentials*, Comm. Math. Phys., 82 (1981/82), pp. 101–120.

[Bjerklov05] K. Bjerklöv, *Positive Lyapunov exponent and minimality for a class of one-dimensional quasi-periodic Schrödinger equations*, Ergodic Theory Dynam. Systems, 25 (2005), pp. 1015–1045.

[Bjerklov06] K. Bjerklöv, *Dynamics of the quasi-periodic Schrödinger cocycle at the lowest energy in the spectrum*, preprint, 2006.

[BDV04] C. Bonatti, L. Diaz, and M. Viana, *Dynamics beyond Uniform Hyperbolicity: A Global Geometric and Probabilistic Perspective*, Springer-Verlag, New York, 2004.

[BHJ+03] H. Broer, H. Hanssmann, À. Jorba, J. Villanueva, and F. Wagener, *Normal-internal resonances in quasi-periodically forced oscillators: A conservative approach*, Nonlinearity, 16 (2003), pp. 1751–1791.

[BHTB90] H. W. Broer, G. B. Huitema, F. Takens, and B. L. J. Braaksma, *Unfoldings and bifurcations of quasi-periodic tori*, Mem. Amer. Math. Soc., 83 (421) (1990).

[BJ02] J. Bourgain and S. Jitomirskaya, *Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential. Dedicated to David Ruelle and Yasha Sinai on the occasion of their* 65*th birthdays*, J. Statist. Phys., 108 (2002), pp. 1203–1218.

[BLW91] D. Beigie, A. Leonard, and S. Wiggins, *Chaotic transport in the homoclinic and heteroclinic tangle regions of quasiperiodically forced two-dimensional dynamical systems*, Nonlinearity, 4 (1991), pp. 775–819.

[BOV97] H. W. Broer, H. M. Osinga, and G. Vegter, *Algorithms for computing normally hyperbolic invariant manifolds*, Z. Angew. Math. Phys., 48 (1997), pp. 480–524.

[BS82] J. Béllissard and B. Simon, *Cantor spectrum for the almost Mathieu equation*, J. Funct. Anal., 48 (1982), pp. 408–419.

[BS98] H. Broer and C. Simó, *Hill's equation with quasi-periodic forcing: Resonance tongues, instability pockets and global phenomena*, Bol. Soc. Brasil. Mat. (N.S.), 29 (1998), pp. 253–293.

[BSV05] H. Broer, C. Simó, and R. Vitolo, *Chaos and Quasi-Periodicity in Diffeomorphisms of the Solid Torus*, MP_ARC # 05-107, 2005.

[CEY90] M. D. Choi, G. A. Elliott, and N. Yui, *Gauss polynomials and the rotation algebra*, Invent. Math., 99 (1990), pp. 225–246.

[CFdlL03a] X. Cabré, E. Fontich, and R. de la Llave, *The parameterization method for invariant manifolds.* I. *Manifolds associated to non-resonant subspaces*, Indiana Univ. Math. J., 52 (2003), pp. 283–328.

[CFdlL03b] X. Cabré, E. Fontich, and R. de la Llave, *The parameterization method for invariant manifolds.* II. *Regularity with respect to parameters*, Indiana Univ. Math. J., 52 (2003), pp. 329–360.

[CFdlL05] X. Cabré, E. Fontich, and R. de la Llave, *The parameterization method for invariant manifolds.* III. *Overview and applications*, J. Differential Equations, 218 (2005), pp. 444–515.

[Chi79] B. V. Chirikov, *A universal instability of many-dimensional oscillator systems*, Phys. Rep., 52 (1979), pp. 264–379.

[CJ00] E. Castellà and À. Jorba, *On the vertical families of two-dimensional tori near the triangular points of the bicircular problem*, Celestial Mech. Dynam. Astronom., 76 (2000), pp. 35–54.

[CL99] C. Chicone and Y. Latushkin, *Evolution Semigroups in Dynamical Systems and Differential Equations*, AMS, Providence, RI, 1999.

[CL00] C. Chicone and W. Liu, *On the continuation of an invariant torus in a family with rapid oscillations*, SIAM J. Math. Anal., 31 (2000), pp. 386–415.

[CLR] Y. Cao, S. Luzzatto, and I. Rios, *The Boundary of Hyperbolicity for Henon-Like Families*, http://arxiv.org/PS_cache/math/pdf/0502/0502235.pdf.

[CLR06] Y. Cao, S. Luzzatto, and I. Rios, *Some non-hyperbolic systems with strictly non-zero Lyapunov exponents for all invariant measures: Horseshoes with internal tangencies*, Discrete Contin. Dyn. Syst., 15 (2006), pp. 61–71.

[DL95] L. Dieci and J. Lorenz, *Computation of invariant tori by the method of characteristics*, SIAM J. Numer. Anal., 32 (1995), pp. 1436–1474.

[dlL97] R. de la Llave, *Invariant manifolds associated to nonresonant spectral subspaces*, J. Statist. Phys., 87 (1997), pp. 211–249.

[dlLGJV05] R. de la Llave, A. González, À. Jorba, and J. Villanueva, *KAM theory without action-angle variables*, Nonlinearity, 18 (2005), pp. 855–895.

[dlLT94] R. de la Llave and S. Tompaidis, *Computation of domains of analyticity for some perturbative expansions of mechanics*, Phys. D, 71 (1994), pp. 55–81.

[DLR91] L. Dieci, J. Lorenz, and R. D. Russell, *Numerical calculation of invariant tori*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 607–647.

[DS83a] P. Deift and B. Simon, *Almost periodic Schrödinger operators.* III. *The absolutely continuous spectrum in one dimension*, Comm. Math. Phys., 90 (1983), pp. 389–411.

[DS83b] F. Delyon and B. Souillard, *The rotation number for finite difference operators and its properties*, Comm. Math. Phys., 89 (1983), pp. 415–426.

[Eli88] L. H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 15 (1988), pp. 115–147 (1989).

[Eli01] L. H. Eliasson, *Almost reducibility of linear quasi-periodic systems*, in Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), Proc. Sympos. Pure Math. 69, AMS, Providence, RI, 2001, pp. 679–705.

[ERS00] K. D. Edoh, R. D. Russell, and W. Sun, *Computation of invariant tori by orthogonal collocation*, Appl. Numer. Math., 32 (2000), pp. 273–289.

[ESA] ESA, *Soho - ESA Space Mission*, http://esapub.esrin.esa.it/bulletin/bullet88/vande88.htm.

[FdlL92] C. Falcolini and R. de la Llave, *Numerical calculation of domains of analyticity for perturbation theories in the presence of small divisors*, J. Statist. Phys., 67 (1992), pp. 645–666.

[Fen72] N. Fenichel, *Persistence and smoothness of invariant manifolds for flows*, Indiana Univ. Math. J., 21 (1971/1972), pp. 193–226.

[FJN03a] R. Fabbri, R. Johnson, and C. Núñez, *Rotation number for non-autonomous linear Hamiltonian systems.* I. *Basic properties*, Z. Angew. Math. Phys., 54 (2003), pp. 484–502.

[FJN03b] R. Fabbri, R. Johnson, and C. Núñez, *Rotation number for non-autonomous linear Hamiltonian systems.* II. *The Floquet coefficients*, Z. Angew. Math. Phys., 54 (2003), pp. 652–676.

[GC91] A. Griewank and G. F. Corliss, eds., *Automatic Differentiation of Algorithms. Theory, Implementation, and Application*, SIAM, Philadelphia, 1991.

[GFPS00] P. Glendinning, U. Feudel, A. S. Pikovsky, and J. Stark, *The structure of mode-locked regions in quasi-periodically forced circle maps*, Phys. D, 140 (2000), pp. 227–243.

[GJ04] F. Gabern and À. Jorba, *Generalizing the restricted three-body problem. The bianular and tricircular coherent problems*, Astronomy and Astrophysics, 420 (2004), pp. 751–762.

[GJMS91a] G. Gómez, A. Jorba, J. Masdemont, and C. Simó, *A dynamical system approach for the analysis of Soho mission*, in Proceedings of the 3rd International Symposium on Spacecraft Flight Dynamics, ESTEC, ESA Publications Division, Noordwijt, Holland, 1991, pp. 449–454.

[GJMS91b] G. Gómez, A. Jorba, J. Masdemont, and C. Simó, *A quasi-periodic solution as a substitute of L_4 in the Earth-Moon system*, in Proceedings of the 3rd International Symposium on Spacecraft Flight Dynamics, ESTEC, ESA Publications Division, Noordwijt, Holland, 1991, pp. 35–41.

[Gre79] J. M. Greene, *A method for determining a stochastic transition*, J. Math. Phys., 20 (1979), pp. 1183–1201.

[Har98] À. Haro, *The Primitive Function of an Exact Symplectomorphism*, Ph.D. thesis, Universitat de Barcelona, Barcelona, Spain, 1998, http://www.maia.ub.es/dsg/1998/index.html.

[Har99] A. Haro, *Center and center-(un)stable manifolds of elliptic-hyperbolic fixed points of 4D-symplectic maps. An example: The Froeschlé map*, in Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995), Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999, pp. 403–407.

[Har02] À. Haro, *An algorithm to generate canonical transformations: Application to normal forms*, Phys. D, 167 (2002), pp. 197–217.

[Har05] À. Haro, *The Breakdown of Invariant Tori in Quasi-Periodic Area Preserving Maps. An Example: The Rotating Standard Map*, work in progress, 2005.

[HdlL03a] A. Haro and R. de la Llave, *Spectral Theory of Transfer Operators* (I): *General Results*, preprint, 2003.

[HdlL03b] A. Haro and R. de la Llave, *Spectral Theory of Transfer Operators* (II): *Vector Bundle Maps over Rotations*, preprint, 2003.

[HdlL05a] A. Haro and R. de la Llave, *A parameterization method for the computation of whiskers in quasi periodic maps: Numerical algorithms*, Discrete Contin. Dyn. Syst. Ser. B, 6 (2006), pp. 1261–1300.

[HdlL05b] A. Haro and R. de la Llave, *Invariant manifolds in quasi-periodic systems: Theory, algorithms and computations*, in SIAM Conference on Applications of Dynamical Systems, Snowbird, UT, 2005.

[HdlL05c] A. Haro and R. de la Llave, *Manifolds on the verge of a hyperbolicity breakdown*, in SIAM Conference on Applications of Dynamical Systems, Snowbird, UT, 2005.

[HdlL06a] A. Haro and R. de la Llave, *Manifolds at the verge of a hyperbolicity breakdown*, Chaos, 16 (2006), 012130.

[HdlL06b] A. Haro and R. de la Llave, *A parameterization method for the computation of invariant tori and their whiskers in quasi periodic maps: Rigorous results*, J. Differential Equations, 228 (2006), pp. 530–579.

[Her83] M.-R. Herman, *Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnol'd et de Moser sur le tore de dimension* 2, Comment. Math. Helv., 58 (1983), pp. 453–502.

[Hof76] D. R. Hofstadter, *Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields*, Phys. Rev. B (3), 14 (1976), pp. 2239–2245.

[HP05] À. Haro and J. Puig, *Strange Nonchaotic Attractors in Harper Maps*, http://www.arXiv.org/abs/nlin.CD/0510073, 2005.

[HPS77] M. W. Hirsch, C. C. Pugh, and M. Shub, *Invariant Manifolds*, Lecture Notes in Math. 583, Springer-Verlag, Berlin, 1977.

[HS05] À. Haro and C. Simó, *To Be or Not to Be an SNA: That Is the Question*, preprint, 2005-17, University of Barcelona, Barcelona, Spain, 2005.

[JdlLZ99] À. Jorba, R. de la Llave, and M. Zou, *Lindstedt series for lower-dimensional tori*, in Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995), Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999, pp. 151–167.

[Jit99] S. Ya. Jitomirskaya, *Metal-insulator transition for the almost Mathieu operator*, Ann. of Math. (2), 150 (1999), pp. 1159–1175.

[JO05] A. Yu. Jalnine and A. H. Osbaldestin, *Smooth and nonsmooth dependence of Lyapunov vectors upon the angle variable on a torus in the context of torus-doubling transitions in the quasiperiodically forced Hénon map*, Phys. Rev. E (3), 71 (2005), 016206.

[Jor99] À. Jorba, *A methodology for the numerical computation of normal forms, centre manifolds and first integrals of Hamiltonian systems*, Experiment. Math., 8 (1999), pp. 155–195.

[Jor01] A. Jorba, *Numerical computation of the normal behaviour of invariant curves of n-dimensional maps*, Nonlinearity, 14 (2001), pp. 943–976.

[JS81] R. A. Johnson and G. R. Sell, *Smoothness of spectral subbundles and reducibility of quasiperiodic linear differential systems*, J. Differential Equations, 41 (1981), pp. 262–288.

[JS92] À. Jorba and C. Simó, *On the reducibility of linear differential equations with quasiperiodic coefficients*, J. Differential Equations, 98 (1992), pp. 111–124.

[JS96] À. Jorba and C. Simó, *On quasi-periodic perturbations of elliptic equilibrium points*, SIAM J. Math. Anal., 27 (1996), pp. 1704–1737.

[JT05] A. Jorba and J. C. Tatjer, *A Mechanism for the Fractalization of Invariant Curves in Quasi-Periodically Forced 1-D Maps*, preprint 2005-12, University of Barcelona, Barcelona, Spain, 2005.

[JV97] À. Jorba and J. Villanueva, *On the normal behaviour of partially elliptic lower-dimensional tori of Hamiltonian systems*, Nonlinearity, 10 (1997), pp. 783–822.

[Kan84] K. Kaneko, *Fractalization of torus*, Progr. Theoret. Phys., 71 (1984), pp. 1112–1115.

[Kan94] I. Kan, *Open sets of diffeomorphisms having two attractors, each with an everywhere dense basin*, Bull. Amer. Math. Soc. (N.S.), 31 (1994), pp. 68–74.

[Kat76] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed., Grundlehren Math. Wiss. 132, Springer-Verlag, Berlin, 1976.

[Kel96] G. Keller, *A note on strange nonchaotic attractors*, Fund. Math., 151 (1996), pp. 139–148.

[Knu97] D. E. Knuth, *The Art of Computer Programming. Vol. 2: Seminumerical Algorithms*, 3rd rev. ed., Addison-Wesley, Reading, MA, London, Don Mills, ON, 1997.

[KO98] B. Krauskopf and H. Osinga, *Growing 1D and quasi-2D unstable manifolds of maps*, J. Comput. Phys., 146 (1998), pp. 404–419.

[Koc99] H. Koch, *A renormalization group for Hamiltonians, with applications to KAM tori*, Ergodic Theory Dynam. Systems, 19 (1999), pp. 1–47.

[Kri99a] R. Krikorian, *C^0-densité globale des systèmes produits-croisés sur le cercle réductibles*, Ergodic Theory Dynam. Systems, 19 (1999), pp. 61–100.

[Kri99b] R. Krikorian, *Réductibilité des systèmes produits-croisés à valeurs dans des groupes compacts*, Astérisque, (259) (1999).

[Kri99c] R. Krikorian, *Réductibilité presque partout des flots fibrés quasi-périodiques à valeurs dans des groupes compacts*, Ann. Sci. École Norm. Sup. (4), 32 (1999), pp. 187–240.

[Kri01] R. Krikorian, *Global density of reducible quasi-periodic cocycles on $\mathbf{T}^1 \times \mathrm{SU}(2)$*, Ann. of Math. (2), 154 (2001), pp. 269–326.

[KSW96] H. Koch, A. Schenkel, and P. Wittwer, *Computer-assisted proofs in analysis and programming in logic: A case study*, SIAM Rev., 38 (1996), pp. 565–604.

[Las94] Y. Last, *Zero measure spectrum for the almost Mathieu operator*, Comm. Math. Phys., 164 (1994), pp. 421–432.

[LM00] H. E. Lomelí and J. D. Meiss, *Heteroclinic primary intersections and codimension one Melnikov method for volume-preserving maps*, in Chaotic Kinetics and Transport (New York, 1998), Chaos, 10 (1) (2000), pp. 109–121.

[LM03] H. E. Lomelí and J. D. Meiss, *Heteroclinic intersections between invariant circles of volume-preserving maps*, Nonlinearity, 16 (2003), pp. 1573–1595.

[Mat68] J. N. Mather, *Characterization of Anosov diffeomorphisms*, Nederl. Akad. Wetensch. Proc. Ser. A 71 = Indag. Math., 30 (1968), pp. 479–483.

[Mei92] J. D. Meiss, *Symplectic maps, variational principles, and transport*, Rev. Modern Phys., 64 (1992), pp. 795–848.

[MKM97] H. Mingyou, T. Küpper, and N. Masbaum, *Computation of invariant tori by the Fourier methods*, SIAM J. Sci. Comput., 18 (1997), pp. 918–942.

[Mn78] R. Mañé, *Persistent manifolds are normally hyperbolic*, Trans. Amer. Math. Soc., 246 (1978), pp. 261–283.

[MO00] B. D. Mestel and A. H. Osbaldestin, *Periodic orbits of renormalisation for the correlations of strange nonchaotic attractors*, Math. Phys. Electron. J., 6 (2000), paper 5.

[MO04] B. D. Mestel and A. H. Osbaldestin, *Renormalization of correlations in a quasiperiodically forced two-level system: Quadratic irrationals*, J. Phys. A, 37 (2004), pp. 8969–8989.

[MS91] K. R. Meyer and D. S. Schmidt, eds., *Computer Aided Proofs in Analysis*, IMA Vol. Math. Appl. 28, Springer-Verlag, New York, 1991.

[NAS] NASA, *Genesis - NASA Space Mission*, http://spacelink.msfc.nasa.gov/NASA.Projects.

[NK96] T. Nishikawa and K. Kaneko, *Fractalization of a torus as a strange nonchaotic attractor*, Phys. Rev. E (3), 54 (1996), pp. 6114–6124.

[NNO98] S. Novo, C. Núñez, and R. Obaya, *Ergodic properties and rotation number for linear Hamiltonian systems*, J. Differential Equations, 148 (1998), pp. 148–185.

[OF00] H. M. Osinga and U. Feudel, *Boundary crisis in quasiperiodically forced systems*, Phys. D, 141 (2000), pp. 54–64.

[OM03] A. H. Osbaldestin and B. D. Mestel, *Renormalization in quasiperiodically forced systems*, Fluct. Noise Lett., 3 (2003), pp. L251–L258.

[OS87] A. Olvera and C. Simó, *An obstruction method for the destruction of invariant curves*, Phys. D, 26 (1987), pp. 181–192.

[Ose68] V. I. Oseledec, *A multiplicative ergodic theorem. Characteristic Lyapunov, exponents of dynamical systems*, Trudy Moskov. Mat. Obšč., 19 (1968), pp. 179–210.

[OWGF01] H. Osinga, J. Wiersig, P. Glendinning, and U. Feudel, *Multistability and nonsmooth bifurcations in the quasiperiodically forced circle map*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 11 (2001), pp. 3085–3105.

[Pes77] Ja. B. Pesin, *Characteristic Lyapunov exponents, and smooth ergodic theory*, Russian Math. Surveys, 32 (1977), pp. 55–114.

[PNR01] A. Prasad, S. S. Negi, and R. Ramaswamy, *Strange nonchaotic attractors*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 11 (2001), pp. 291–309.

[Pös89] J. Pöschel, *On elliptic lower-dimensional tori in Hamiltonian systems*, Math. Z., 202 (1989), pp. 559–608.

[Pui04] J. Puig, *Cantor spectrum for the almost Mathieu operator*, Comm. Math. Phys., 244 (2004), pp. 297–309.

[Pui05] J. Puig, *A nonperturbative Eliasson's reducibility theorem*, Nonlinearity, 19 (2006), pp. 335–376.

[RJB83] R. R. Ricklefs, W. Jefferys, and R. Broucke, *A general precompiler for algebraic manipulation*, Celest. Mech., 29 (1983), pp. 179–190.

[Rue85] D. Ruelle, *Rotation numbers for diffeomorphisms and flows*, Ann. Inst. H. Poincaré Phys. Théor., 42 (1985), pp. 109–115.

[Ryc92] M. Rychlik, *Renormalization of cocycles and linear ODE with almost-periodic coefficients*, Invent. Math., 110 (1992), pp. 173–206.

[Sac78] R. J. Sacker, *Existence of dichotomies and invariant splittings for linear differential systems.* IV, J. Differential Equations, 27 (1978), pp. 106–137.

[SFKA96] O. Sosnovtseva, U. Feudel, J. Kurths, and A. Pikovsky, *Multiband strange attractors in quasiperiodically forced systems*, Phys. Lett. A, 218 (1996), pp. 255–272.

[Sim79] C. Simó, *On the Hénon-Pomeau attractor*, J. Statist. Phys., 21 (1979), pp. 465–494.

[Sim98] C. Simó, *Effective computations in celestial mechanics and astrodynamics*, in Modern Methods of Analytical Mechanics and Their Applications (Udine, 1997), V. V. Rumynantsev and A. V. Karapatian, eds., Springer-Verlag, Vienna, 1998, pp. 55–102.

[Sim99] C. Simó, *Dynamical systems methods for space missions on a vicinity of collinear libration points*, in Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995), Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999, pp. 223–241.

[Sta99] J. Stark, *Regularity of invariant graphs for forced systems*, Ergodic Theory Dynam. Systems, 19 (1999), pp. 155–199.

[Thi97] Ph. Thieullen, *Ergodic reduction of random products of two-by-two matrices*, J. Anal. Math., 73 (1997), pp. 19–64.

[Tom96] S. Tompaidis, *Numerical study of invariant sets of a quasiperiodic perturbation of a symplectic map*, Experiment. Math., 5 (1996), pp. 211–230.

[Tru00] M. R. Trummer, *Spectral methods in computing invariant tori*, in Auckland Numerical Ordinary Differential Equations (Auckland, 1998), Appl. Numer. Math., 34 (2–3) (2000), pp. 275–292.

[WR87] R. L. Warnock and R. D. Ruth, *Invariant tori through direct solution of the Hamilton-Jacobi equation*, Phys. D, 26 (1987), pp. 1–36.

[WY01] Q. Wang and L.-S. Young, *Strange attractors with one direction of instability*, Comm. Math. Phys., 218 (2001), pp. 1–97.

[WY02] Q. Wang and L.-S. Young, *From invariant curves to strange attractors*, Comm. Math. Phys., 225 (2002), pp. 275–304.

SIAM J. APPLIED DYNAMICAL SYSTEMS
Vol. 6, No. 1, pp. 208–235

Some Canonical Bifurcations in the Swift–Hohenberg Equation*

L. A. Peletier† and J. F. Williams‡

Abstract. We study the nature and stability of stationary solutions $u(x)$ of the fourth order Swift–Hohenberg equation on a bounded domain $(0, L)$ with boundary conditions $u = 0$ and $u'' = 0$ at $x = 0$ and $x = L$. It is well known that as L increases, the set of stationary solutions becomes increasingly complex. Numerical studies have exhibited two interesting types of structures in the bifurcation diagram for (L, u). In this paper we demonstrate through a center manifold analysis how these structures arise naturally near certain bifurcation points, and that there are no others. We also analyze their stability properties.

Key words. Swift–Hohenberg equation, higher order parabolic equations, bifurcation, nonlinear eigenvalue problems

AMS subject classifications. 34B15, 34C23, 35J40, 37G10

DOI. 10.1137/050647232

1. Introduction.

In this paper we consider solutions of the *Swift–Hohenberg equation*

$$\frac{\partial u}{\partial t} = \alpha\, u - \left(1 + \frac{\partial^2}{\partial x^2}\right)^2 u - u^3, \qquad \alpha \in \mathbf{R}, \tag{1.1}$$

on the cylindrical domain $Q = (0, L) \times \mathbf{R}^+$ subject to the boundary conditions

$$u = 0 \quad \text{and} \quad \frac{\partial^2 u}{\partial x^2} = 0 \qquad \text{at } x = 0,\, L. \tag{1.2}$$

It is well known (see [1], [2], and [18, Chapter 9]) that (1.1) has an increasing number of stationary solutions which satisfy the boundary conditions (1.2) as the length L of the domain increases. The objective of this paper is to study the set of stationary solutions as it depends on α and L. In particular, we focus on understanding certain complex but recurrent structures in the bifurcation diagram, which were first observed in numerical studies.

In studies of pattern formation, the Swift–Hohenberg equation plays a central role. Proposed in 1977 by Swift and Hohenberg [21] in connection with Rayleigh–Bénard convection, it has since featured in a variety of problems, such as Taylor–Couette flow [14], [19] and the study of lasers [15]. For further references we mention the surveys given in [8] and [9] and the recent review [3].

*Received by the editors December 11, 2005; accepted for publication (in revised form) by B. Fiedler September 6, 2006; published electronically March 30, 2007. This research was supported by RTN network HPRN-CT-2002-00274.
http://www.siam.org/journals/siads/6-1/64723.html

†Mathematical Institute, Leiden University, PB 9512, 2300 RA, Leiden, The Netherlands, and Centrum voor Wiskunde en Informatica, Postbus 94079, 1090 GB, Amsterdam, The Netherlands (peletier@math.leidenuniv.nl).

‡Department of Mathematics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada (jfw@math.sfu.ca).

We view the Swift–Hohenberg equation as a model equation for a large class of higher order parabolic model equations arising in a wide range of applications, such as the extended Fisher–Kolmogorov equation in statistical mechanics [10], [22], and a sixth order equation introduced by Caginalp and Fife in phase field models [4], [11]. Many of the phenomena observed here for the Swift–Hohenberg equation are found in the dynamics of such equations as well [2], [5], [6], [7] and [18].

The Swift–Hohenberg equation is interesting from the point of view of pattern formation, because of its many qualitatively different stable equilibrium solutions. This raises the question, which of these equilibrium solutions will be selected as time tends to infinity, and how this selection depends on the parameters α and L involved in the problem.

We study the Cauchy–Dirichlet problem for the Swift–Hohenberg equation on the cylinder $Q = (0, L) \times \mathbf{R}^+$ with boundary conditions (1.2). Thus, writing (1.1) in a more conventional form, we consider the problem

$$\text{(1.3)} \qquad \begin{cases} u_t = -u_{xxxx} - 2u_{xx} - (1-\alpha)u - f(u) & \text{for} \quad 0 < x < L, \quad t > 0, \\ u = 0 \quad \text{and} \quad u_{xx} = 0 & \text{at} \quad x = 0, L, \quad t > 0, \\ u(x, 0) = u_0(x) & \text{for} \quad 0 < x < L, \end{cases}$$

where we consider nonlinearities $f(u)$ which are either stabilizing or destabilizing. We shall assume that f is a smooth function, and that $f(s) = o(s)$ as $s \to 0$. As typical examples of such functions we choose

$$\text{(1.4)} \qquad f(s) = s^3 \text{ (stabilizing)} \qquad \text{or} \qquad f(s) = -s^3 + s^5 \text{ (destabilizing).}$$

The boundary conditions have been chosen so that solutions can be extended as periodic functions on $\mathbf{R}$. The initial function u_0 is a smooth function that vanishes at $x = 0$ and $x = L$. In most of this article we will assume that u_0 is symmetric with respect to the center of the domain; i.e.,

$$\text{(1.5)} \qquad u_0(L - x) = u_0(x) \qquad \text{for} \qquad 0 < x < L.$$

Note that this assumption implies that the solution remains symmetric for all t. It is motivated by the fact that it increases the set of lengths L for which the trivial solution is globally stable [16], [17].

Problem (1.3) is a *gradient system* with corresponding *Lyapunov functional*

$$\text{(1.6)} \qquad J(u; L) = \frac{1}{L} \int_0^L \left\{ \frac{1}{2}(u'')^2 - (u')^2 + \frac{1-\alpha}{2} u^2 + F(u) \right\} dx,$$

where $F'(s) = f(s)$ and $F(0) = 0$. This means (cf. [13]) that if the stationary solutions of problem (1.3) are isolated, then $u(x, t)$ tends to one of these solutions as $t \to \infty$. In other words, for every $x \in (0, L)$,

$$\text{(1.7)} \qquad u(x, t) \to v(x) \qquad \text{as} \qquad t \to \infty,$$

where $v(x)$ is a solution of the two-point boundary value problem

$$\begin{cases} v^{iv} + 2v'' + (1-\alpha)v + f(v) = 0 & \text{for} \quad 0 < x < L, \\ v = 0 \quad \text{and} \quad v'' = 0 & \text{at} \quad x = 0,\ L. \end{cases} \tag{1.8}$$

For the stabilizing case it is known that if $\alpha \leq 0$, then problem (1.3) has the trivial solution only (cf. Chapter 9 of [18]), and hence, for every $x \in (0, L)$,

$$u(x,t) \to 0 \qquad \text{as} \qquad t \to \infty.$$

For both cases the situation is much more complex when $\alpha > 0$, and, depending on the value of α and L, there may be very many stationary profiles to choose from. In this article we make a start with answering the question of which of these profiles is chosen.

In [16], [17] a series of numerical simulations was carried out on the Swift–Hohenberg equation in order to gain insight into the different types of limiting behaviors and the effect of the parameters α and L on the final profile that is selected. One of the striking observations was that for values of α smaller than, say, $\frac{1}{4}$, critical lengths L_1 and L_2 could be identified such that the solution converged to one type of limit for $L_1 = \pi/\sqrt{1+\sqrt{\alpha}} < L < L_2 = \pi/\sqrt{1-\sqrt{\alpha}}$ and to another when $3L_1 < L < 3L_2$, provided that the intervals $[L_1, L_2]$, $[3L_1, 3L_2]$, and $[5L_1, 5L_2]$ are disjunct. In between, for $L \in [L_2, 3L_1]$ and for $L \in [3L_2, 5L_1]$ the trivial solution proved globally stable.

When the above intervals overlapped, the simulations were not very revealing. This motivated us to carry out a numerical study of solution branches of the stationary problem (1.8). In Figure 1 we present a result of such a study for $\alpha = 0.65$. It reveals interesting qualitatively new structures in the intersections of the intervals, which are quite robust under changes of α. We see that loop-type branches bifurcate from the trivial solution at $L = nL_1$ and $L = nL_2$, $n = 1, 3, 5, \ldots$. But in addition we see the following:

(A) comparable—nontrivial—structures in the intersection of the intervals $[L_1, L_2]$ and $[3L_1, 3L_2]$, and in the intersection of the intervals $[3L_1, 3L_2]$ and $[9L_1, 9L_2]$,

(B) pairs of branches which connect the loop-type branch bifurcating from $5L_1$ to the ones from $3L_1$ and from $7L_1$.

A dynamic exploration of this behavior is presented in the movie linked from Figure 2, which traces out the stable behavior of (1.1)–(1.2) for evolving L.

In this paper we analyze these structures by means of a center manifold analysis carried out when, for instance, $3L_1 \approx L_2$ and the intervals $[L_1, L_2]$ and $[3L_1, 3L_2]$ slightly overlap.

We shall show that when $nL_1 \approx mL_2$ and the intervals (mL_1, mL_2) and (nL_1, nL_2) slightly overlap for certain odd integers m and n, then two types of structures may be distinguished,

(i) when $n = 3m$ we obtain the structure shown in Figure 3(left), and

(ii) when $n \neq 3m$ we obtain the structure shown in Figure 3(right),

and these are the only two possible structures. In addition, only if $m = 1$ and $n = 3$ is the dimension of the unstable manifold X^u equal to zero. In all other cases it is positive.

The plan of the paper is the following. First, in section 2, we recall some results about the stability properties of the linearized Swift–Hohenberg equation. Then, in section 3, we discuss stationary solutions in the limit as $\alpha \to 0$. In section 4, we turn to the first characteristic

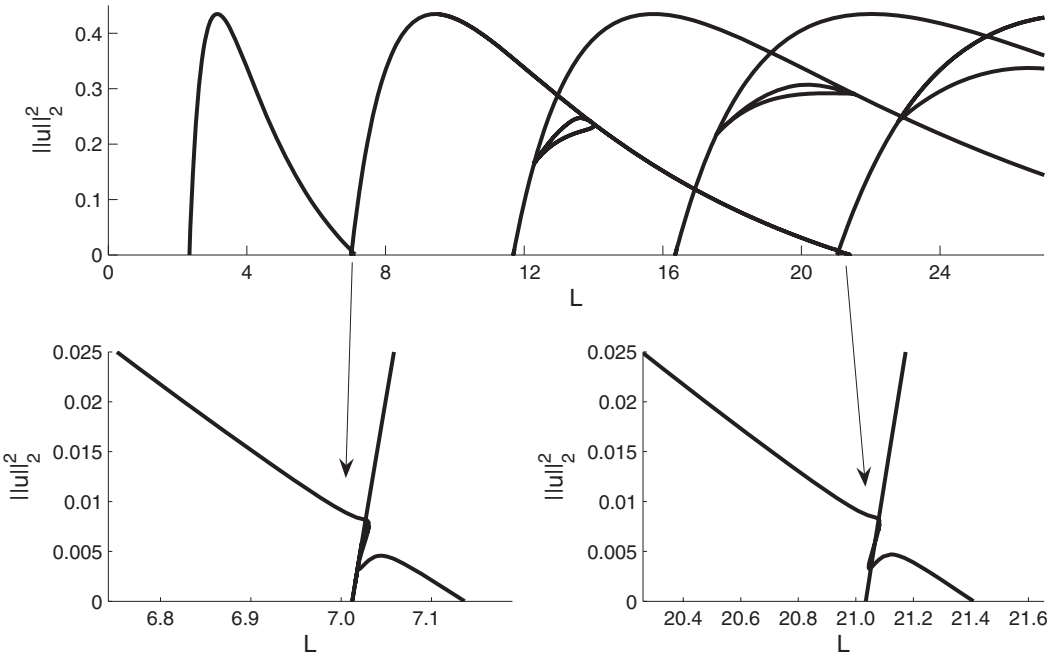


Figure 1. *Global bifurcation diagram when* $\alpha = 0.65$.

solution set when $3L_1 \approx L_2$, and in section 5, we generalize the analysis to the situation when $nL_1 \approx mL_2$. In section 6, we compute the Lyapunov functions for the solutions found in sections 4 and 5 and examine which of the solutions is the global minimizer, given a value of L. Finally, in section 7, we apply the techniques used in the previous sections to the Swift–Hohenberg equation with the destabilizing quintic nonlinearity: $f(s) = -s^3 + s^5$.

2. The linearized Swift–Hohenberg equation. In this section we recall some properties of the linear problem associated with problem (1.3):

$$\begin{cases} u_t = -u_{xxxx} - 2u_{xx} - (1-\alpha)u & \text{for} \quad 0 < x < L, \quad t > 0, \\ u = 0 \quad \text{and} \quad u_{xx} = 0 & \text{at} \quad x = 0, L, \quad t > 0, \\ u(x,0) = u_0(x) & \text{for} \quad 0 < x < L. \end{cases} \tag{2.1}$$

The stability of the zero solution to (2.1) is determined by examining the ansatz $u(x,t) = \varphi(x)e^{-\lambda t}$. This leads to the eigenvalue problem

$$\begin{cases} \varphi^{(iv)} + 2\varphi'' + (1-\alpha)\varphi = \lambda\varphi & \text{for} \quad 0 < x < L, \\ \varphi = 0 \quad \text{and} \quad \varphi'' = 0 & \text{at} \quad x = 0, \ L. \end{cases} \tag{2.2}$$

The eigenvalues λ_n and the eigenfunctions φ_n are given by

$$\lambda_n = \lambda_n(L) = P\left(\frac{n\pi}{L}\right) \qquad \text{and} \qquad \varphi_n(x) = \sqrt{2}\sin\left(\frac{n\pi x}{L}\right), \tag{2.3}$$

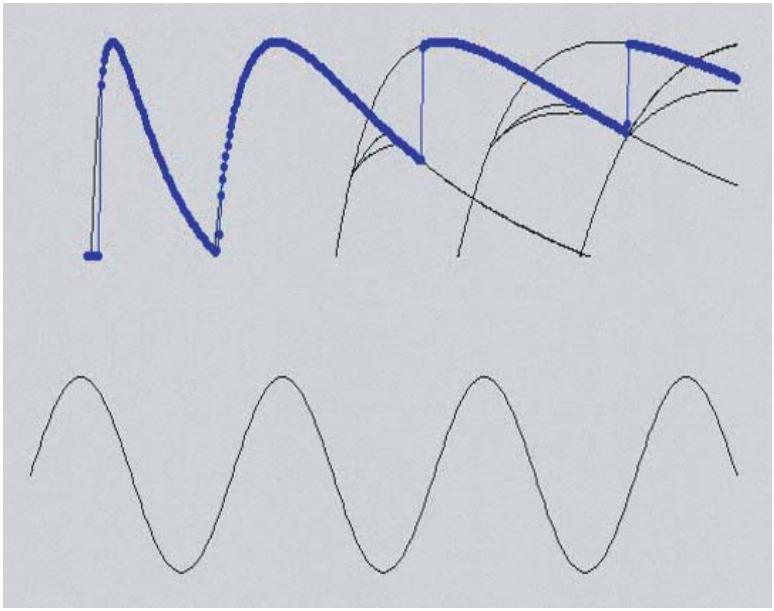


Figure 2. *Animation of a PDE simulation with time-dependent length* $L = 2 + 0.01t$ *and* $\alpha = 0.65$. *Top: The solid line is the bifurcation diagram from Figure* 1*; the dotted line is sampled from the PDE computation. Bottom: The solution corresponding to the* $\times$ *above. Clicking on the above image displays the accompanying animation* (64723_01.mpg [1.81MB]).

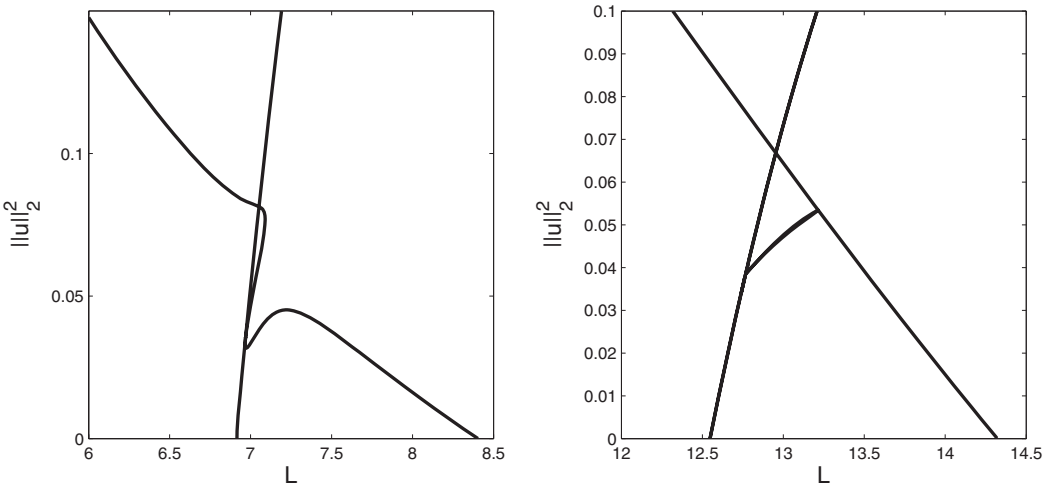


Figure 3. *Detail of bifurcation diagrams when* $[L_1, L_2]$ *and* $[3L_1, 3L_2]$ *overlap and* $\alpha = 0.75$ *(left), and when* $[3L_1, 3L_2]$ *and* $[5L_1, 5L_2]$ *overlap and* $\alpha = 0.32$ *(right).*

where $P(\xi)$ is the symbol of the operator on the left-hand side of (2.2),

$$P(\xi) = (\xi^2 - 1)^2 - \alpha,$$

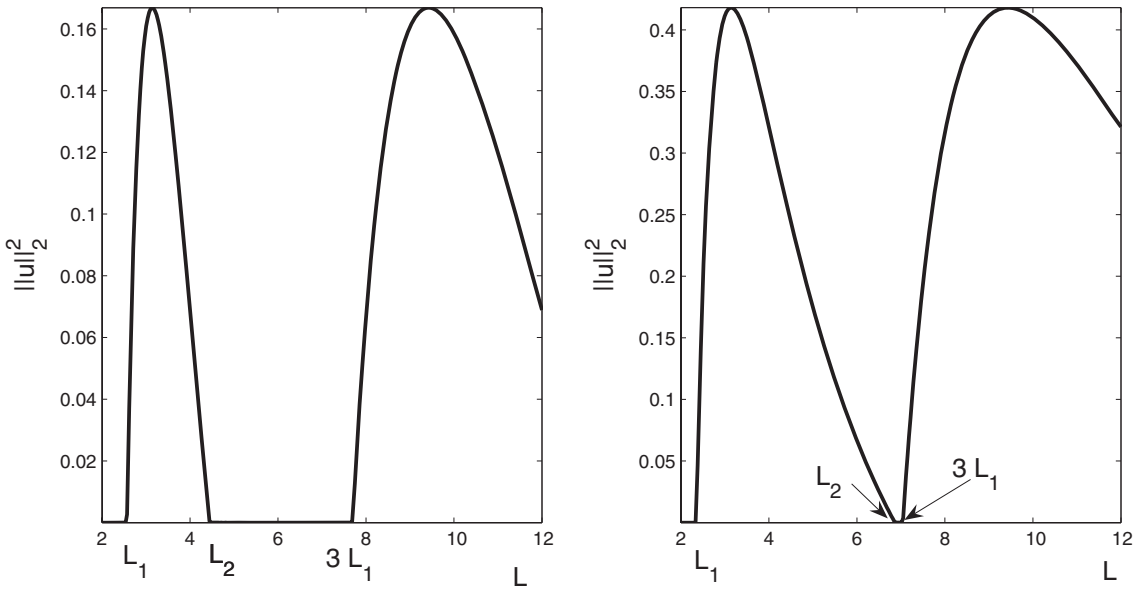


Figure 4. *L^2-norms of stable solutions as a function of L. In the left figure $\alpha = 1/4$, while in the right $\alpha = 5/8$. Notice that as α increases further, the two solution branches come together at the base line. They meet when $\alpha = \alpha_1^* = 16/25$.*

and the eigenfunctions have been normalized so that

$$(\varphi_m, \varphi_n) \stackrel{\text{def}}{=} \frac{1}{L}\int_0^L \varphi_m(x)\varphi_n(x)\,dx = 1 \qquad \text{if } m = n.$$

What is interesting about this equation is that, given $\alpha \in (0,1)$ small enough, there are domain lengths for which the zero solution is stable and others for which it is not. Let $\alpha \in (0,1)$. Then $P(\xi)$ has two zeros $\xi_\pm$ $(0 < \xi_- < \xi_+)$, so that $P < 0$ on (ξ_-, ξ_+) and $P > 0$ on $(0, \xi_-) \cup (\xi_+, \infty)$. This implies that

$$\lambda_n(L) < 0 \qquad \text{when } L \in (nL_1, nL_2),$$

where $L_1 = \pi/\xi_+$ and $L_2 = \pi/\xi_-$; i.e.,

$$L_1 = \frac{\pi}{\sqrt{1+\sqrt{\alpha}}} \qquad \text{and} \qquad L_2 = \frac{\pi}{\sqrt{1-\sqrt{\alpha}}}.$$

This phenomenon is exhibited in Figure 4. Here we have solved the Cauchy–Dirichlet problem (1.3) numerically for $\alpha = 1/4$ and $\alpha = 5/8$ and randomly chosen initial data. The L^2-norm of the limiting solution $v(x)$ is plotted against the domain length, clearly showing regions when the zero solution is or is not stable. Details of the numerical simulations are discussed in Appendix B.

In order to analyze the solution branches on the (L,u)-plane, we vary α so that two critical domain lengths coalesce:

$$\text{Case I: } L_2 - L_1 \searrow 0 \qquad \text{and} \qquad \text{Case II: } 3L_1 - L_2 \searrow 0.$$

The first case arises when $\alpha \to 0$. Here we will formally construct the solutions in this limit to describe the branch of solutions connecting L_1 and L_2 as depicted in Figure 4.

The second case arises when

$$\alpha \searrow \alpha_1^* \stackrel{\text{def}}{=} \frac{16}{25} \qquad \text{and} \qquad L_1^* = 3L_1(\alpha_1^*) = L_2(\alpha_1^*) = \pi\sqrt{5}.$$

In [16], [17] the local shape near $L = L_1$ and $L = L_2$ of this diagram was determined. Specifically, the following results were established.

Theorem 2.1. *Let $\alpha > 0$. There exists a unique branch in the (L, u)-plane of nontrivial solutions of problem (1.3), which emanates from the trivial solution at L_1. Its local behavior is given by*

$$\|u\|^2 \sim \frac{8}{3\pi}\sqrt{\alpha}(1+\sqrt{\alpha})^{3/2}(L - L_1) \qquad \text{as } L \searrow L_1.$$

Theorem 2.2. *Let $0 < \alpha < 1$. There exists a unique branch in the (L, u)-plane of nontrivial solutions of problem (1.3), which emanates from the trivial solution at L_2. Its local behavior is given by*

$$\|u\|^2 \sim \frac{8}{3\pi}\sqrt{\alpha}(1-\sqrt{\alpha})^{3/2}(L_2 - L) \qquad \text{as } L \nearrow L_2.$$

Here $\|u\|^2 = (1/L)\int_0^L u^2(x)\,dx$.

It is our present interest to understand the full time-dependent problem near those values of α and L where these bifurcation points approximately coalesce.

We have the following center manifold theorem.

Theorem 2.3. *Suppose that $\alpha = \alpha_1^*$ and so $3L_1 = L_2$. Then if $L = 3L_1$,*

(a) *problem (1.3) has a two-dimensional center manifold X^c about the trivial solution which is spanned by φ_1 and φ_3;*

(b) *the dimension of the unstable manifold about the zero solution, X^u, is zero.*

Proof. If we extend problem (1.3) to include $\alpha_t = 0$, and linearize about $u(x,t) = 0$ and $\alpha = \alpha_1^*$, we obtain the system

$$\text{(2.4)} \qquad \begin{cases} u_t = -u_{xxxx} - 2u_{xx} - (1-\alpha)u & \text{for} \quad 0 < x < L,\ t > 0, \\ u = 0 \quad \text{and} \quad u_{xx} = 0 & \text{at} \quad x = 0, L,\ t > 0, \\ \alpha_t = 0 & \text{for} \quad t > 0. \end{cases}$$

The eigenfunctions and eigenvalues of problem (2.4) are given by (2.3). Since α has been chosen such that $3L_1 = L_2$ and $L = 3L_1$, it follows that

$$\lambda_1(L) = 0 \qquad \text{and} \qquad \lambda_3(L) = 0.$$

A simple calculation shows that

$$\lambda_n = \left(\left(\frac{n}{\sqrt{5}}\right)^2 - 1\right)^2 - \frac{16}{25} = \frac{n^2}{5}\left(\frac{n^2}{25} - 2\right) + \frac{9}{25} > 0 \qquad \text{for } n = 5, 7, \ldots$$

so that all the eigenvalues, starting from λ_5, are positive. Therefore, (i) there exists a two-dimensional center manifold for this value $\alpha = \alpha_1^*$ which is spanned by φ_1 and φ_3, and (ii) the unstable manifold has dimension zero.

By smoothness this manifold persists for small changes in α and L [20]. ∎

3. Stationary solutions in the limit as $\alpha \to 0^+$. In this section we construct asymptotic expressions for nontrivial stationary solutions of problem (1.3) in the limit as $\alpha \to 0^+$ when the nonlinearity is cubic, i.e., when it is given by $f(s) = s^3$. Thus, we seek to describe nontrivial solutions of the problem

$$\begin{cases} v^{iv} + 2v'' + (1-\alpha)v + v^3 = 0 & \text{for} \quad 0 < x < L, \\ v = 0 \quad \text{and} \quad v'' = 0 & \text{at} \quad x = 0,\ L. \end{cases} \tag{3.1}$$

We choose $L \in (nL_1, nL_2)$, where n is an arbitrary positive integer, so that a nontrivial solution is known to exist, and we choose α so small that the intervals $[nL_1, nL_2]$ and $[(n+2)L_1, (n+2)L_2]$ are disjoint. This is the case when $\alpha < \alpha_n^*$, where

$$\alpha_n^* = \frac{4(n+1)^2}{(n^2+2n+2)^2}.$$

It will be convenient to rescale the spatial variable so that the domain $(0, L)$ maps onto the fixed domain $(0, 1)$. Thus, we set

$$y = Lx \qquad \text{and} \qquad w(y) = v(x).$$

We then find that problem (3.1) becomes

$$\begin{cases} w^{iv} + 2L^2 w'' + L^4\left((1-\alpha)w + w^3\right) = 0 & \text{for} \quad 0 < x < 1, \\ v = 0, \quad v'' = 0 & \text{at} \quad x = 0,\ 1. \end{cases} \tag{3.2}$$

We are interested in the situation that $\alpha \to 0^+$. The existence of branches bifurcating from L_1 and L_2 is established in Theorems 2.1 and 2.2. These local results show that as $L \to L_1^+$ and $L \to L_2^-$ the solution scales like $\sqrt{\alpha}$. We therefore seek an expansion of $w(y)$ in a series of the form

$$w(y) \sim \alpha^{1/2} w_1(y) + \alpha w_2(y) + \alpha^{3/2} w_3(y) + \alpha^2 w_4(y) + \cdots. \tag{3.3}$$

We introduce a scaling factor $\delta \in [0, 1]$, write

$$L = n\pi L_1 + n\pi\delta(L_2 - L_1) = n\pi\{L_1 + \delta L_{\text{gap}}\}, \qquad 0 \le \delta \le 1,$$

and expand L in the limit as $\alpha \to 0^+$. A simple computation shows that

$$L_1(\alpha) = \frac{\pi}{\sqrt{1+\sqrt{\alpha}}} = \pi\left(1 - \frac{1}{2}\alpha^{1/2} + \frac{3}{8}\alpha - \frac{5}{16}\alpha^{3/2} + \cdots\right)$$

and

$$L_2(\alpha) = \frac{\pi}{\sqrt{1-\sqrt{\alpha}}} = \pi\left(1 + \frac{1}{2}\alpha^{1/2} + \frac{3}{8}\alpha + \frac{5}{16}\alpha^{3/2} + \cdots\right),$$

so that

$$L_{\text{gap}}(\alpha) \stackrel{\text{def}}{=} L_2(\alpha) - L_1(\alpha) \sim \pi\left(\alpha^{1/2} + \frac{5}{8}\alpha^{3/2} + \cdots\right) \qquad \text{as} \quad \alpha \to 0^+.$$

Therefore as $\alpha \to 0^+$ we have

$$L(\alpha) = n\pi \left\{ 1 + \frac{1}{2}(2\delta - 1)\alpha^{1/2} + \frac{3}{8}\alpha + \frac{5}{16}(2\delta - 1)\alpha^{3/2} + \cdots \right\}. \tag{3.4}$$

We now expand (3.2), using (3.3) and (3.4), in terms of powers of $\alpha^{1/2}$, and equate the coefficients of $\alpha^{1/2}$, α, $\alpha^{3/2}$, and α^2 equal to zero. We then obtain the equations

$$\begin{aligned}
\mathcal{O}(\alpha^{1/2}): &\quad \mathcal{L}_1 w_1 = 0,\\
\mathcal{O}(\alpha): &\quad \mathcal{L}_1 w_2 + \mathcal{L}_2 w_1 = 0,\\
\mathcal{O}(\alpha^{3/2}): &\quad \mathcal{L}_1 w_3 + \mathcal{L}_2 w_2 + \mathcal{L}_3(w_1) + n^4\pi^4 w_1^3 = 0,\\
\mathcal{O}(\alpha^2): &\quad \mathcal{L}_1 w_4 + \mathcal{L}_2 w_3 + \mathcal{L}_3 w_2 + \mathcal{L}_4 w_1 + n^4\pi^4\{3w_1^2 w_2 + 2(2\delta - 1)w_1^3\} = 0,
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{L}_1 z &\stackrel{\text{def}}{=} z^{iv} + 2n^2\pi^2 z'' + n^4\pi^4 z,\\
\mathcal{L}_2 z &\stackrel{\text{def}}{=} 2n^2\pi^2(2\delta - 1)(z'' + n^2\pi^2 z),\\
\mathcal{L}_3 z &\stackrel{\text{def}}{=} 2n^2\pi^2\{1 - \delta(1 - \delta)\}z'' + n^4\pi^4\{2 - 6\delta(1 - \delta)\}z,\\
\mathcal{L}_4 z &\stackrel{\text{def}}{=} 2n^2\pi^2(2\delta - 1)\left(z'' + n^2\pi^2\{2 - \delta(1 - \delta)\}z\right).
\end{aligned}$$

The functions $w_j(y)$, $j = 1, 2, 3, 4, \ldots$, all need to satisfy the boundary conditions of problem (3.2):

$$w_j(y) = 0 \quad \text{and} \quad w_j''(y) = 0 \qquad \text{at} \qquad y = 0,\ 1.$$

This means that

$$w_1(y) = \gamma_1 \sin(n\pi y),$$

where γ_1 still needs to be determined. Since $\mathcal{L}_2 w_1 = 0$, the equation for w_2 reduces to $\mathcal{L}_1 w_2 = 0$, so that

$$w_2(y) = \gamma_2 \sin(n\pi y)$$

as well, and γ_2 must also be determined. To fix γ_1 we enforce the solvability condition for the equation for w_3:

$$(w_1, \mathcal{L}_1 w_3) = -(w_1, \mathcal{L}_2 w_2 + \mathcal{L}_3 w_1 + n^4\pi^4 w_1^3) = 0,$$

where $(\cdot, \cdot)$ is the standard inner product in $L^2(0, 1)$. Since $\mathcal{L}_2$ is self-adjoint and $\mathcal{L}_2 w_1 = 0$, this condition reduces to

$$(w_1, \mathcal{L}_3 w_1 + n^4\pi^4 w_1^3) = 0,$$

which yields

$$\gamma_1(\delta) = \frac{4}{\sqrt{3}}\sqrt{\delta(1 - \delta)} \qquad \text{with} \quad \delta \in [0, 1]. \tag{3.5}$$

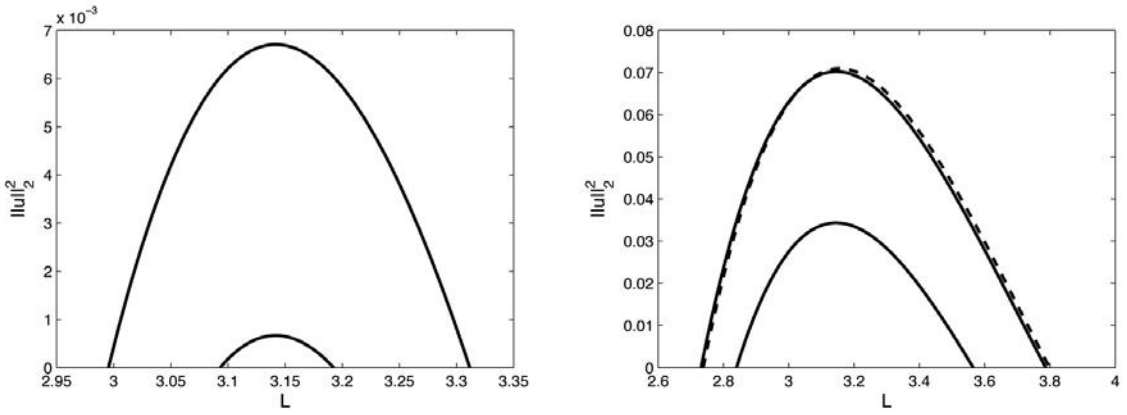


Figure 5. *Comparison of the numerical and asymptotic structures of the branches for* $\alpha = 10^{-3}, 10^{-2}$ *on the left and* $\alpha = 5 \cdot 10^{-2}$ *and* 10^{-1} *on the right.*

To determine γ_2 we use the solvability condition for the problem for the w_4 term in the expansion. This yields the condition

$$(w_1, \mathcal{L}_3 w_2 + \mathcal{L}_4 w_1 + n^4\pi^4\{3w_1^2 w_2 + 2(2\delta - 1)w_1^3\}) = 0,$$

from which we derive γ_2. The first two terms of the expansion are thus found to be

$$w(y;\delta) = \sqrt{\delta(1-\delta)}\left(\frac{4}{\sqrt{3}}\sqrt{\alpha} + \alpha\sqrt{3}\,(1-2\delta)\right)\sin(n\pi y) + \cdots . \tag{3.6}$$

A comparison of the branch predicted by this formula with numerical computations is shown in Figure 5.

The stability of the solutions on this branch is easily computed by setting

$$u(x,t) \simeq c(t)\sqrt{\alpha}v_1(x) + \cdots , \qquad v_1(x) = \gamma_1 \sin\left(\frac{n\pi x}{L}\right).$$

Clearly one asymptotic solution of the full time-dependent PDE is $c(t) = \gamma_1$. Projecting this solution onto the local center manifold $X^c = \text{span}\{\sin(n\pi x/L)\}$ and expanding to leading order in α leads to

$$\dot{c} = -\frac{1}{8}c\left\{3c^2 - 16\delta(1-\delta)\right\}. \tag{3.7}$$

Linearizing this equation about $c(t) = \gamma_1 + d(t)$ for small d yields

$$\dot{d} = -4\delta(1-\delta)d,$$

from which we conclude that this solution is stable for admissible δ. This comes as no surprise, as it is precisely in this interval that the zero solution is unstable.

The construction described above involves an interval (nL_1, nL_2), which is disjunct from $((n+2)L_1, (n+2)L_2)$. In sections 4 and 5 we consider what happens when these two intervals overlap, albeit slightly, and we investigate the dynamics on a two-dimensional center manifold.

4. Stationary solutions in the limit as $\alpha \to \alpha_1^*$. In this section we focus on the situation when the branch which connects the branchpoints at L_1 and L_2 overlaps with the branch which connects $3L_1$ and $3L_2$. As we saw in the introduction, this will be the case when $\alpha > \alpha_1^*$. Numerical studies suggest that the set of stationary solutions then can become quite complex. In this section we analyze this set when the bifurcation points at L_2 and $3L_1$ are close together, i.e., when $\alpha = \alpha_1^* + \varepsilon$ and ε is positive and small.

As we have seen in Theorem 2.3, when $\alpha = \alpha_1^*$ then problem (1.7) has a two-dimensional center manifold X^c about the trivial solution

$$X^c = \operatorname{span}\{\varphi_1, \varphi_3\} = \{\sin(x\pi/L), \sin(3x\pi/L)\}.$$

In our analysis we shall fix α close to α_1^* and use L as a bifurcation parameter taking values close to $L_1^* = L_2(\alpha_1^*)$.

To determine the dynamics on the space X^c we write

$$u(x,t) = a(t)\varphi_1(x) + b(t)\varphi_3(x)$$

and project the differential equation onto X^c:

$$\frac{1}{L}\int_0^L \{u_t + u_{xxxx} + 2u_{xx} + (1-\alpha)u + u^3\}\varphi_1\, dx = 0,$$

$$\frac{1}{L}\int_0^L \{u_t + u_{xxxx} + 2u_{xx} + (1-\alpha)u + u^3\}\varphi_3\, dx = 0.$$

This yields a pair of differential equations for $a(t)$ and $b(t)$:

$$\begin{cases} \dot a = -P\left(\dfrac{\pi}{L}\right)a - \dfrac{3}{2}a^3 + \dfrac{3}{2}a^2 b - 3ab^2, \\ \dot b = -P\left(\dfrac{3\pi}{L}\right)b + \dfrac{1}{2}a^3 - 3a^2 b - \dfrac{3}{2}b^3, \end{cases} \tag{4.1}$$

where dots denote differentiation with respect to t.

Because the original equation is a gradient system, the reduced system is one as well. The corresponding Lyapunov functional (the projection of (1.6) onto X^c) is given by

$$V(a,b) = \frac{1}{2}P\left(\frac{\pi}{L}\right)a^2 + \frac{1}{2}P\left(\frac{3\pi}{L}\right)b^2 + \frac{3}{8}a^4 - \frac{1}{2}a^3 b + \frac{3}{2}a^2 b^2 + \frac{3}{8}b^4. \tag{4.2}$$

Plainly,

$$\dot a = -V_a(a,b) \qquad \text{and} \qquad \dot b = -V_b(a,b),$$

where subscripts denote partial differentiation, and

$$\frac{dV}{dt} = V_a\,\dot a + V_b\,\dot b = -(V_a^2 + V_b^2) \le 0.$$

Thus, the function $V(a,b)$ decreases along orbits.

We will consider the structure and stability of the set of stationary solutions of the system (4.1). They are defined by the pair of equations

$$V_a(a,b)=0 \qquad \text{and} \qquad V_b(a,b)=0.$$

This solution set contains at least three distinct branches: (i) the trivial state, (ii) a branch bifurcating from the bifurcation point at $3L_1$, and (iii) a branch emanating from L_2. The existence of these branches was established in [17], but there the structure was only described in the limit as $\|u\| \to 0$. However, because of the center manifold structure near $L_1^* = 3L_1 = L_2$ at α_1^*, we see that these branches persist under small changes in α and L near the first bifurcation point $(u, L) = (0, L_1^*)$. It is their local structure and stability that we now wish to investigate.

We fix

$$\alpha = \alpha_1^* + \varepsilon, \qquad \varepsilon > 0.$$

Then $3L_1 < L_2$ and we write L, which serves as a bifurcation parameter, as

$$L = 3L_1 + \delta L_{\text{gap}}, \qquad \text{where} \qquad L_{\text{gap}} = L_2 - 3L_1, \qquad \delta \in \mathbf{R}.$$

Note that $L_{\text{gap}} > 0$ since $3L_1 < L_2$ in this range of α. The parameter δ positions the length L with respect to the interval $(3L_1, L_2)$ so that $\delta = 0$ corresponds with $3L_1$, and $\delta = 1$ corresponds with L_2. In what follows we shall describe the set of stationary solutions of (4.1) in the (δ, u)-space, where $u = (a, b)$, and discuss their stability properties.

Plainly, $L_{\text{gap}}(\varepsilon) \to 0$ as $\varepsilon \to 0$. To obtain a more precise estimate, we expand $3L_1$ and L_2 in powers of ε. This yields

$$3L_1 = \frac{3\pi}{\sqrt{1+\sqrt{\alpha_1^*+\varepsilon}}} = \pi\sqrt{5}\left\{1 - \frac{25}{144}\varepsilon + O(\varepsilon^2)\right\} \qquad \text{as } \varepsilon \to 0,$$

and

$$L_2 = \frac{\pi}{\sqrt{1-\sqrt{\alpha_1^*+\varepsilon}}} = \pi\sqrt{5}\left\{1 + \frac{25}{16}\varepsilon + O(\varepsilon^2)\right\} \qquad \text{as } \varepsilon \to 0.$$

Therefore

$$L_{\text{gap}}(\varepsilon) = \frac{125}{72}\pi\sqrt{5}\,\varepsilon + O(\varepsilon^2) \qquad \text{as } \varepsilon \to 0.$$

Based on the natural balance of terms in (4.1) we rescale the variables according to

$$(a,b) \mapsto \sqrt{\varepsilon}(a,b), \qquad t \mapsto \frac{1}{\varepsilon}t. \tag{4.3}$$

Using this rescaling and expanding (4.1) to $O(\varepsilon)$, we obtain the leading order problem

$$\begin{cases} \dot{a} = g_1(a,b) \stackrel{\text{def}}{=} a\left\{\dfrac{10}{9}(1-\delta) - \dfrac{3}{2}a^2 + \dfrac{3}{2}ab - 3b^2\right\}, \\ \dot{b} = g_2(a,b) \stackrel{\text{def}}{=} 10\delta b + \dfrac{1}{2}a^3 - 3a^2b - \dfrac{3}{2}b^3. \end{cases} \tag{4.4}$$

Any solution of the system (4.4) recovers a solution to (4.1) to $O(\varepsilon^2)$.

4.1. Stationary solutions. The stationary solutions of the system (4.4) are the points where the null clines Γ_1 and Γ_2, defined by

$$\Gamma_i = \{(a,b) : g_i(a,b) = 0\} \qquad \text{for } i = 1,2, \tag{4.5}$$

intersect. Plainly, Γ_1 consists of two components:

$$\Gamma_1^{(1)} = \{(a,b) : a = 0\} \qquad \text{and} \qquad \Gamma_1^{(2)} = \left\{(a,b) : \frac{3}{2}a^2 - \frac{3}{2}ab + 3b^2 = \frac{10}{9}(1-\delta)\right\},$$

and

$$\Gamma_2 = \left\{(a,b) : \frac{1}{2}a^3 - 3a^2b - \frac{3}{2}b^3 + 10\delta b = 0\right\}.$$

Note that both Γ_1 and Γ_2 are invariant under the transformation $(a,b) \to (-a,-b)$; i.e., they are symmetric with respect to the origin. In what follows we shall at times discuss only the null clines in the *first* and the *second* quadrant, since the structure there will be duplicated in the third and the fourth quadrant.

We immediately see that the intersection of $\Gamma_1^{(1)}$ and Γ_2 consists of two points in the (a,b)-plane:

$$O = (0,0) \quad \text{for } \delta \in \mathbf{R} \qquad \text{and} \qquad P = \left(0, \sqrt{\frac{20\delta}{3}}\right) \quad \text{for } \delta \geq 0. \tag{4.6}$$

The first point yields the branch of trivial solutions, and the second point yields the branch of stationary solutions of the system (4.4) which emanates from $3L_1$, as described in [16] and [17]. It undergoes no bifurcation as α increases through α_1^*.

It is the third branch, the one which emanates from L_2, for which a is not identically zero, and which corresponds to the points of intersection of $\Gamma_1^{(2)}$ and Γ_2, that we now concentrate on. We see that for any $\delta < 1$ the null cline $\Gamma_1^{(2)}$ is an ellipse in the (a,b)-plane around the origin, and that Γ_2 defines a continuous curve connecting $(0,0)$ to the point at infinity. Thus, for any $\delta < 1$ there is always at least one stationary solution of the system (4.4). In Figure 6 we show the null clines Γ_1 and Γ_2 for $\delta = 0.025$ (left), $\delta = 0.1$ (middle), and $\delta = 0.2$ (right).

The graphs of $\Gamma_1^{(2)}$ and Γ_2 suggest that there are three critical values of δ: $\delta_1 < \delta_2 < \delta_3$ in the interval $(0,1)$ such that

- if $0 < \delta < \delta_1$, then $\Gamma_1^{(2)} \cap \Gamma_2$ consists of *one* point in the first quadrant, which we denote by $A(\delta)$, and *one* point, $-A(\delta)$, in the third quadrant.

When δ increases and passes through δ_1, two additional points of intersection appear in the *second* quadrant, and two appear in the fourth quadrant:

- if $\delta_1 < \delta < \delta_2$, then $\Gamma_1^{(2)} \cap \Gamma_2$ consists of *two* points in the second quadrant, which we denote by $B(\delta)$ and $C(\delta)$, and *one* point, $A(\delta)$, in the first quadrant. Symmetrically located with respect to the origin, there are points $-A(\delta)$, $-B(\delta)$, and $-C(\delta)$ in the lower half-plane.

When δ increases, $B(\delta)$ moves to the b-axis and crosses it at δ_2, so that

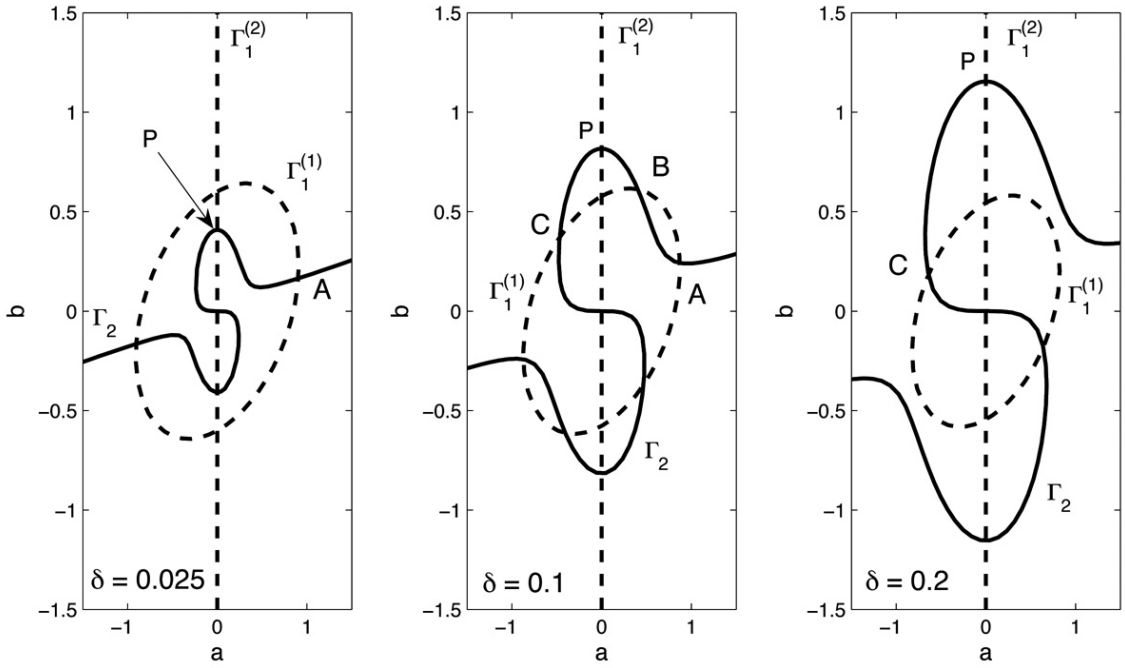


Figure 6. *Plots of* $\Gamma_1^{(1)}$ *and* $\Gamma_1^{(2)}$ *(dashed) and* Γ_2 *(solid) for* $\delta = 0.025$, $\delta = 0.1$, *and* $\delta = 0.2$.

- if $\delta_2 < \delta < \delta_3$, then $\Gamma_1^{(2)} \cap \Gamma_2$ consists of the origin, *one* point $C(\delta)$ in the second quadrant, and *two* points, $A(\delta)$ and $B(\delta)$, in the first quadrant. Symmetrically located with respect to the origin, there are points $-A(\delta)$, $-B(\delta)$, and $-C(\delta)$ in the lower half-plane.

When $\delta = \delta_3$, the points $A(\delta)$ and $B(\delta)$ coalesce to a point in the *first* quadrant and subsequently disappear, so that

- if $\delta_3 < \delta < 1$, then $\Gamma_1^{(2)} \cap \Gamma_2$ consists of *one* point, $C(\delta)$, in the second quadrant and *one* point, $-C(\delta)$, in the fourth quadrant.

Remark. The points $B(\delta)$ and $C(\delta)$ move together as $\delta \to \delta_1^+$, and the points $B(\delta)$ and $A(\delta)$ move together as $\delta \to \delta_2^-$.

In Appendix A we prove all these statements. There we find that

$$\delta_1 = 0.050785\ldots, \qquad \delta_2 = 0.052631\ldots, \qquad \text{and} \qquad \delta_3 = 0.15191\ldots.$$

4.2. Stability. The stability of the stationary solutions obtained above can easily be established by means of an analysis of the vector field of the system (4.4).

As a first observation, we note that $\Gamma_1^{(1)} = \{(a, b) : a = 0\}$ is an invariant set. This set contains three stationary points: the points P and $-P$ as well as the origin O. Within this set, the origin is a repellor and the points $\pm P$ are both attractors. Thus, the origin is always *unstable*.

To discuss the nontrivial stationary solutions, we consider the four ranges of δ-values: $(0, \delta_1)$, (δ_1, δ_2), (δ_2, δ_3), and $(\delta_3, 1)$ in succession. Graphs of orbits in the ranges $(0, \delta_1)$, (δ_2, δ_3), and $(\delta_3, 1)$ are shown in Figure 7.

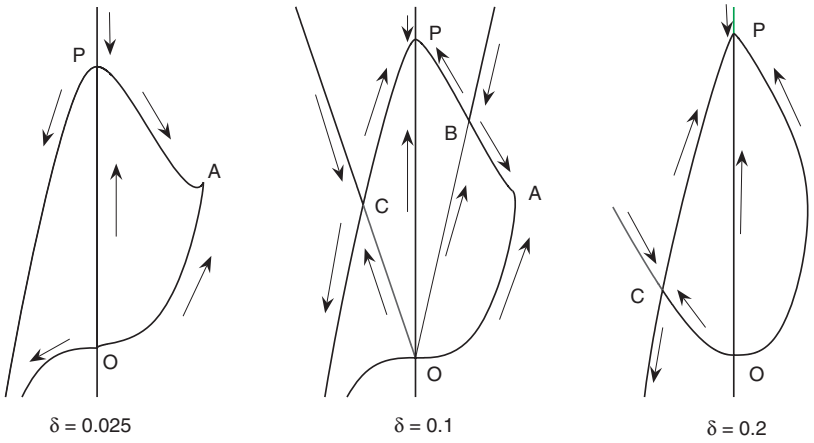


Figure 7. *Connecting orbits in the phase plane when* $\delta = 0.025 \in (0, \delta_1)$, $\delta = 0.1 \in (\delta_2, \delta_3)$, *and* $\delta = 0.2 \in (\delta_3, 1)$.

The range $0 < \delta < \delta_1$. We have four nontrivial stationary points: $\pm P$ and $\pm A$. By symmetry the characters of $\pm P$ and $\pm A$ are the same:

- Both P and $-P$ are saddles: their stable manifold is $\Gamma_1^{(1)}$, and their unstable manifold is perpendicular to the b-axis. Therefore, both P and $-P$ are unstable.
- The points A and $-A$ are both stable nodes.

The range $\delta_1 < \delta < \delta_2$. We have four nontrivial stationary points, P, A, B, and C in the upper half-plane, and their symmetric images $-P$, $-A$, $-B$, and $-C$ in the lower half-plane.

- The points $\pm A$ and $\pm B$ are stable nodes.
- The points $\pm P$ and $\pm C$ are saddles and hence unstable.

The range $\delta_2 < \delta < \delta_3$. We have nontrivial stationary points, P, A, B, and C in the upper half-plane and their symmetric images $-P$, $-A$, $-B$, and $-C$ in the lower half-plane.

- The points $\pm P$ and $\pm A$ and their symmetric images are stable nodes.
- The points B and C and their symmetric images are saddles and hence unstable.

The range $\delta_3 < \delta < 1$. We have four nontrivial stationary points: $\pm P$ and $\pm C$.

- The points $\pm P$ are stable nodes.
- The points $\pm C$ are saddles.

When we translate these results to the solution branches in the (L, u)-plane, we obtain the following description:

• *The branch emanating from the point* $(3L_1, 0)$ *starts at* $\delta = 0$ *and corresponds to the point* P *in the phase plane.*

- *For* $0 < \delta < \delta_2$ *the branch is unstable.*
- *For* $\delta_2 < \delta < 1$ *it is stable.*

• *The branch emanating from the point* $(L_2, 0)$ *starts at* $\delta = 1$ *and corresponds successively to the points* C, B, *and* A *in the phase plane.*

• *For* $\delta_1 < \delta < 1$ *the branch corresponds to* C, *and it is unstable. At* δ_1 *the branch bends back and passes through a saddle-node bifurcation (note that the local maximum near* $a = 0$

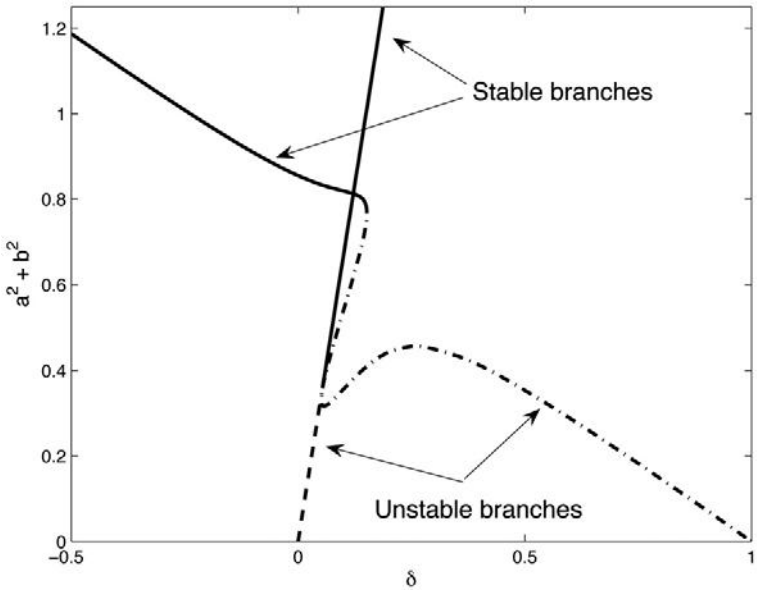


Figure 8. *Stability of solution branches.*

of Γ_2 *occurs for* $a < 0$*). This part of the branch corresponds to the point* B *and is stable for* $\delta_1 < \delta < \delta_2$*, i.e., until it intersects with the branch which emanates from* $(3L_1, 0)$*. Continuing beyond* δ_2*, it becomes unstable, until at* δ_3 *it bends back again through a saddle-node bifurcation. Then, for* $0 < \delta < \delta_3$ *the branch corresponds to the point* A*, and it is stable again.*

The stability properties of the different pieces of the branches are indicated in Figure 8.

5. Further bifurcation points. In this section we study structures of bifurcation curves in the (L, u)-plane when higher order bifurcation points (nearly) coincide, i.e., when there exists an $\alpha^* \in (0, 1)$ and positive integers m and n such that

$$nL_1(\alpha^*) = mL_2(\alpha^*). \tag{5.1}$$

Since $L_1 < L_2$ it follows that $m < n$. Throughout we consider only *odd* integers. It is readily verified that (5.1) occurs when

$$\alpha^* = \alpha^*_{m,n} \stackrel{\text{def}}{=} \left(\frac{n^2 - m^2}{n^2 + m^2}\right)^2. \tag{5.2}$$

As a first observation we note that when $n = 3m$, (5.1) becomes $3L_1(\alpha^*) = L_2(\alpha^*)$, so that $\alpha^* = \alpha^*_1$, and we return to the structure described in section 4. In Figure 9, in which $5L_1 \approx 3L_2$, we see a different type of structure. This type of connecting branch will be the main focus in this section.

Let us generalize the center manifold theorem, Theorem 2.3, from section 2 to this situation.

Theorem 5.1. *Suppose that* (5.1) *holds for some* $\alpha^* \in (0, 1)$ *and for odd integers* m *and* n*. Then, if* $L = nL^*_1$*, the following hold:*

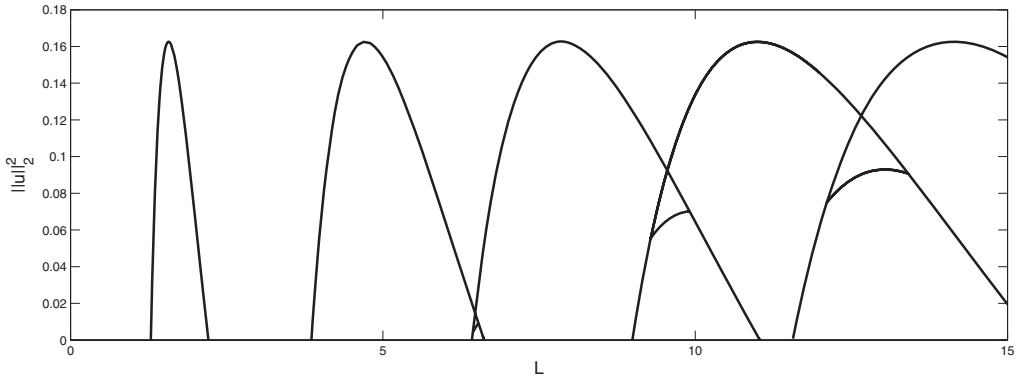


Figure 9. *Global bifurcation diagram when* $\alpha = 0.24$ *and* $5L_1 \approx 3L_2$.

(a) *There exists a two-dimensional center manifold* X^c *spanned by the eigenfunctions* φ_m *and* φ_n.
(b) *If* $n = m + 2\ell$ *for* $\ell \geq 1$, *then the dimension of the unstable manifold* X^u *is equal to* $\ell - 1$.

Proof. Since we have chosen L such that $L = nL_1(\alpha^*) = mL_2(\alpha^*)$, it follows that

$$\lambda_m(L) = 0 \qquad \text{and} \qquad \lambda_n(L) = 0.$$

Remembering that $\lambda_k(L) = P(k\pi/L)$, we therefore have

$$P(\xi) = 0 \qquad \text{for} \qquad \xi = \xi_m \quad \text{and} \quad \xi = \xi_n, \quad \text{where} \quad \xi_k = \frac{k\pi}{L}.$$

Given the properties of the function $P(\xi)$, we see immediately that

$$P(\xi) > 0 \qquad \text{for } \xi \in (0, \xi_m) \cup (\xi_n, \infty)$$

and

$$P(\xi) < 0 \qquad \text{for } \xi \in (\xi_m, \xi_n).$$

Since $\xi_n = \xi_m + (2l\pi/L)$, this implies that

$$\lambda_k(L) = P(\xi_k) < 0 \qquad \text{for} \qquad k = 1, 2, \ldots, \ell - 1,$$

so that the dimension of the unstable manifold is equal to $\ell - 1$. ∎

Remark. We note that the dimension of X^u is positive, precisely when there is a bifurcation point inside an interval (nL_1, nL_2). Thus while we can construct the local bifurcation diagram for any bifurcation point, only those which occur at α_1^* are related to stable dynamics.

We now extend the calculations of section 3 to the general case. Projecting the Swift–Hohenberg equation onto the space $X^c = \text{span}\{\varphi_m, \varphi_n\}$ and writing

$$u(x,t) = a(t)\varphi_m(x) + b(t)\varphi_n(x),$$

we obtain the system

$$\begin{cases} \dot{a} = -P\left(\dfrac{m\pi}{L}\right) a - \dfrac{3}{2}a^3 - 3ab^2 + \dfrac{3}{2}a^2 b(\varphi_{3m}, \varphi_n), \\ \dot{b} = -P\left(\dfrac{n\pi}{L}\right) b - 3a^2 b - \dfrac{3}{2}b^3 + \dfrac{1}{2}a^3(\varphi_{3m}, \varphi_n), \end{cases} \tag{5.3}$$

where we have used the fact that

$$(\varphi_m^3, \varphi_n) = -\frac{1}{2}(\varphi_{3m}, \varphi_n) \qquad \text{and} \qquad (\varphi_m^2, \varphi_n^2) = 1.$$

We note that when $3m = n$ we retrieve the system (4.1) in section 4, and we obtain the same structure, albeit scaled by a factor m.

If $3m \neq n$, the inner products in (5.3) vanish and we obtain the new system

$$\begin{cases} \dot{a} = -P\left(\dfrac{m\pi}{L}\right) a - \dfrac{3}{2}a^3 - 3ab^2, \\ \dot{b} = -P\left(\dfrac{n\pi}{L}\right) b - 3a^2 b - \dfrac{3}{2}b^3. \end{cases} \tag{5.4}$$

Next, we increase α by ε; i.e., we set $\alpha = \alpha_{m,n}^* + \varepsilon$ ($\varepsilon > 0$). Then the bifurcation points $nL_1(\alpha)$ and $mL_2(\alpha)$ move apart. As in section 4, we write

$$L = L_{m,n}^* + \delta L_{\text{gap}},$$

where now

$$L_{m,n}^* = nL_1(\alpha^*) = \pi\sqrt{\frac{m^2+n^2}{2}} \qquad \text{and} \qquad L_{\text{gap}} = \frac{\pi}{8\sqrt{2}}\frac{(n^2+m^2)^{7/2}}{n^2m^2(n^2-m^2)}\varepsilon + \mathcal{O}(\varepsilon^2),$$

and

$$L = L_{m,n}^* + \delta L_{\text{gap}} = L_{m,n}^*\left\{1 + \frac{\varepsilon}{8}\frac{(n^2+m^2)^2}{n^2(n^2-m^2)}\left(\frac{m^2+n^2}{m^2}\delta - 1\right) + \mathcal{O}(\varepsilon^2)\right\}.$$

Rescaling the variables as in (4.3) and expanding $P\left(\frac{m\pi}{L}\right)$ and $P\left(\frac{n\pi}{L}\right)$ into powers of ε as before then leads to the system

$$\begin{cases} \dot{a} = a\left\{\dfrac{n^2+m^2}{n^2}(1-\delta) - \dfrac{3}{2}a^2 - 3b^2\right\}, \\ \dot{b} = b\left\{\dfrac{n^2+m^2}{m^2}\delta - 3a^2 - \dfrac{3}{2}b^2\right\}. \end{cases} \tag{5.5}$$

5.1. Stationary solutions. The null clines of the system (5.3) are two ellipses, as well as the two axes:

$$\Gamma_1 = \Gamma_1^{(1)} \cup \Gamma_1^{(2)}, \qquad \Gamma_2 = \Gamma_2^{(1)} \cup \Gamma_2^{(2)},$$

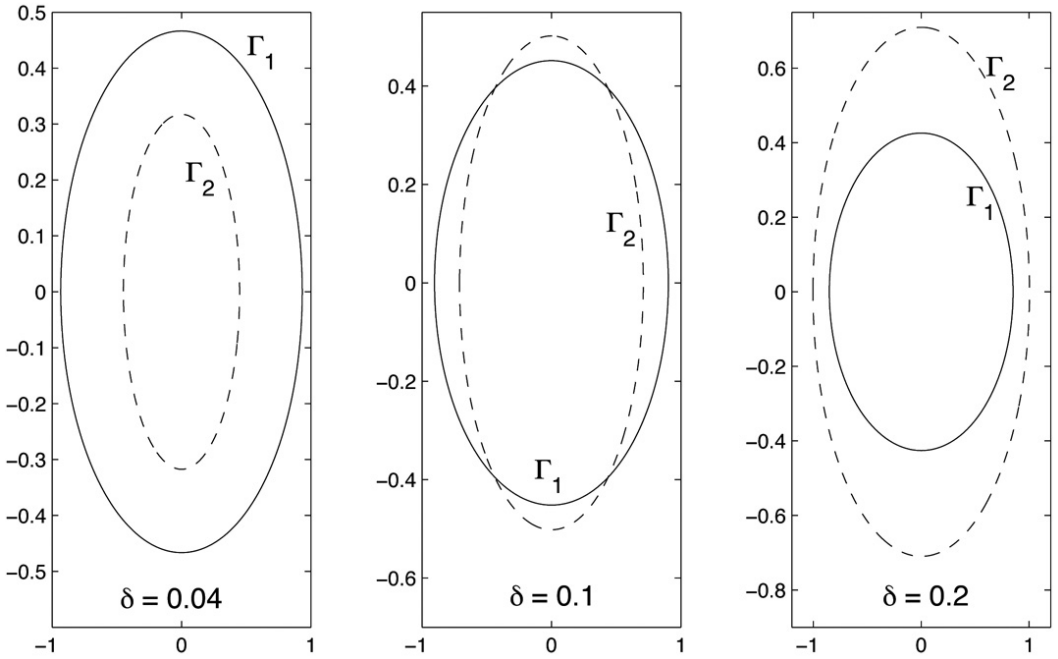


Figure 10. *Plots of* $\Gamma_1^{(2)}$ *and* $\Gamma_2^{(2)}$ *for* $\delta = 0.04$, $\delta = 0.1$, *and* $\delta = 0.2$.

where

$$\Gamma_1^{(1)} = \{(a,b) : a = 0\} \qquad \text{and} \qquad \Gamma_1^{(2)} = \left\{(a,b) : a^2 + 2b^2 = \frac{2}{3}\frac{n^2+m^2}{n^2}(1-\delta)\right\}$$

and

$$\Gamma_2^{(1)} = \{(a,b) : b = 0\} \qquad \text{and} \qquad \Gamma_2^{(2)} = \left\{(a,b) : 2a^2 + b^2 = \frac{2}{3}\frac{n^2+m^2}{m^2}\delta\right\}.$$

They are shown in Figure 10.

We see that for all $\delta \in (0,1)$ the set $\Gamma_1 \cap \Gamma_2$ contains the origin as well as the points P and $-P$ and Q and $-Q$, where

$$P(\delta) = \left(0, \sqrt{\frac{2(n^2+m^2)}{3m^2}\delta}\right) \qquad \text{and} \qquad Q(\delta) = \left(\sqrt{\frac{2(n^2+m^2)}{3n^2}(1-\delta)}, 0\right).$$

The points $\pm P$ correspond to the branch which bifurcates at nL_1, and the points $\pm Q$ correspond to the branch which bifurcates at mL_2. These solutions are the only ones for $0 < \delta < \delta_1$ and for $\delta_2 < \delta < 1$, where

$$\delta_1 = \frac{m^2}{m^2+2n^2} \qquad \text{and} \qquad \delta_2 = \frac{2m^2}{2m^2+n^2}.$$

At δ_1 and δ_2 branches of solutions bifurcate from $\pm P$ and $\pm Q$, and for $\delta_1 < \delta < \delta_2$ we have four additional solutions. We denote them by $\pm A$ and $\pm B$, where

$$A(\delta) = K\left(\sqrt{(m^2+2n^2)\delta - m^2}, \sqrt{2m^2 - (2m^2+n^2)\delta}\right),$$

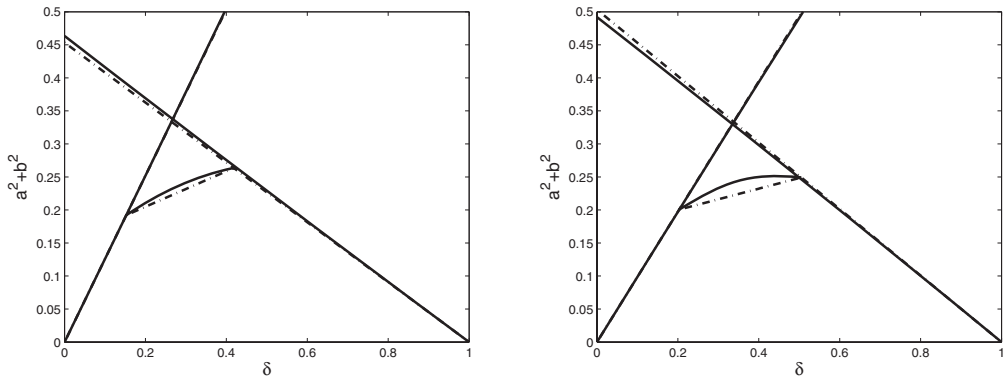


Figure 11. *Comparison of the asymptotic and numerical bifurcation diagrams. Dashed lines are asymptotic solutions, solid are numerical. Left: $m = 3$ and $n = 5$. Right: $m = 5$ and $n = 7$.*

where $K = (1/(3mn))\sqrt{2(m^2+n^2)}$ and

$$B(\delta) = K\left(-\sqrt{(m^2+2n^2)\delta - m^2}, \sqrt{2m^2 - (2m^2+n^2)\delta}\right).$$

5.2. Stability. A stability analysis of these solutions such as presented in section 4 shows the following:

If $0 < \delta < \delta_1$, then
- P and $-P$ are saddles and hence unstable.
- Q and $-Q$ are stable nodes.

If $\delta_1 < \delta < \delta_2$, then
- P and $-P$ and Q and $-Q$ are all stable nodes.
- A and $-A$ and B and $-B$ are saddles and unstable.

If $\delta_2 < \delta < 1$, then
- P and $-P$ are stable nodes.
- Q and $-Q$ are saddles and unstable.

Let us finally translate these results to the bifurcation picture in the $(L, \|u\|^2)$-diagram. For convenience we denote the branch which bifurcates at nL_1 by $\mathcal{C}_1$ and the branch which bifurcates at mL_2 by $\mathcal{C}_2$.

- *The branch $\mathcal{C}_1$ corresponds to the point P; it is unstable for $0 < \delta < \delta_1$ and stable for $\delta_1 < \delta < 1$.*
- *The branch $\mathcal{C}_2$ corresponds to the point Q; it is unstable for $\delta_2 < \delta < 1$ and stable for $0 < \delta < \delta_2$.*
- *There are secondary bifurcations from $\mathcal{C}_1$ and $\mathcal{C}_2$ at, respectively, δ_1 and δ_2. These branch points are connected by a branch $\mathcal{C}_3$ which spans (δ_1, δ_2). The branch $\mathcal{C}_3$ is unstable.*

A comparison of these structures with numerical computations is presented in Figure 11. Note that

$$\|u\|^2 = a^2 + b^2,$$

so that the branch $\mathcal{C}_3$ becomes

$$\|u\|^2 = \frac{2}{9}\frac{m^2+n^2}{m^2n^2}\left\{m^2 + (n^2-m^2)\delta\right\}, \qquad \delta_1 \le \delta \le \delta_2;$$

i.e., $\mathcal{C}_3$ is a straight line in the $(L, \|u\|^2)$-plane (see Figure 11).

6. Global minimizers. In sections 4 and 5 we have computed solutions of problem (1.3) in the center manifold X^c of the trivial solutions at certain critical lengths, whenever $mL_1 \approx nL_2$ and $mL_1 < nL_2$ for some odd integers m and n. The same analysis also yielded the stability properties of these solutions in X^c.

Since problem (1.3) is a gradient system, it is interesting to know which of these solutions is the global minimizer. In this section we determine which of the solutions that we obtained in X^c is the global minimizer in X^c, given any value of $\delta \in [0,1]$. To this end we compute the value of the Lyapunov functional $V(a,b)$ along the solution branches in the two structures discussed in the previous sections, the one we find when $n = 3m$ and the one we find when $n \neq 3m$. We discuss the two structures in succession.

Case I: $n = 3m$. We recall from section 4 that in this case the Lyapunov functional $V(a,b)$ associated with problem (4.1) is given by

$$V(a,b) = \frac{1}{2}P\left(\frac{\pi}{L}\right)a^2 + \frac{1}{2}P\left(\frac{3\pi}{L}\right)b^2 + \frac{3}{8}a^4 - \frac{1}{2}a^3b + \frac{3}{2}a^2b^2 + \frac{3}{8}b^4. \tag{6.1}$$

We set $\alpha = \alpha_1^* + \varepsilon$, carry out the scaling

$$(a,b) \mapsto \sqrt{\varepsilon}(a,b), \qquad t \to \frac{1}{\varepsilon}t, \qquad \text{and} \qquad V(a,b) \to \frac{1}{\varepsilon}V(a,b),$$

and expand $P\left(\frac{\pi}{L}\right)$ and $P\left(\frac{3\pi}{L}\right)$. This yields for the scaled Lyapunov functional

$$V(a,b) = -\frac{20}{9}(1-\delta)a^2 - 20\delta b^2 + \frac{3}{8}a^4 - \frac{1}{2}a^3b + \frac{3}{2}a^2b^2 + \frac{3}{8}b^4. \tag{6.2}$$

In Figure 12 we show graphs of the Lyapunov functional $J(u;L)$ (numerically computed) and of $V(a,b)$ along the solution branches near $3L_1$, when $3L_1 \approx L_2$.

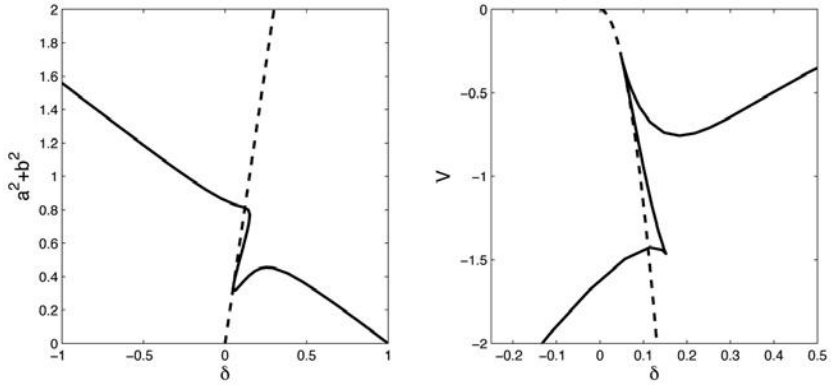


Figure 12. *Solution branches (left) and scaled Lyapunov function V along the branches (right) when $m = 1$ and $n = 3$.*

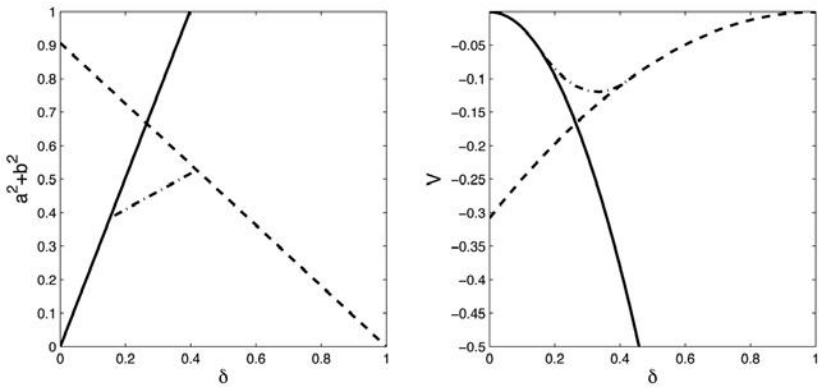


Figure 13. *Solution branches (left) and scaled Lyapunov function V along the branches (right) when $m = 3$ and $n = 5$.*

Case II: $n \neq 3m$. Proceeding as in Case I, we find that the scaled Lyapunov function $V(a, b)$ is now given by

$$V(a,b) = -\frac{n^2+m^2}{2n^2}(1-\delta)a^2 - \frac{n^2+m^2}{2m^2}\delta b^2 + \frac{3}{8}a^4 + \frac{3}{2}a^2b^2 + \frac{3}{8}b^4. \tag{6.3}$$

In Figure 13 we show graphs of the Lyapunov functional $J(u; L)$ (numerically computed) and of $V(a, b)$ along the solution branches near $5L_1$, when $5L_1 \approx 3L_2$.

7. The quintic Swift–Hohenberg equation. In this section we expand the analysis of the previous sections to the Swift–Hohenberg equation in which the cubic term u^3 has been replaced by the quintic polynomial $-u^3 + u^5$. Specifically, we consider the following initial-boundary value problem:

$$\begin{cases} u_t = -u_{xxxx} - 2u_{xx} - (1-\alpha)u + u^3 - u^5 & \text{for} \quad 0 < x < L, \quad t > 0, \\ u = 0 \quad \text{and} \quad u_{xx} = 0 & \text{at} \quad x = 0, L, \quad t > 0, \\ u(x,0) = u_0(x) & \text{for} \quad 0 < x < L. \end{cases} \tag{7.1}$$

As for the Swift–Hohenberg equation, the linearized version of problem (7.1) is problem (2.1), which has been studied in detail in section 2. We find the same critical lengths L_1 and L_2 and the same eigenvalues λ_n and eigenfunctions φ_n. Solution branches bifurcate from the points $(L, u) = (nL_k, 0)$ $(k = 1, 2)$ for every $n \geq 1$. However, whereas in the Swift–Hohenberg equation the bifurcations at $(nL_1, 0)$ are supercritical and subcritical at $(nL_2, 0)$, here they are subcritical at $(nL_1, 0)$ and supercritical and at $(nL_2, 0)$. Proceeding as in [17], we can find the local behavior at $(L_1, 0)$ and $(L_2, 0)$, as follows.

Theorem 7.1. *Let $\alpha > 0$. There exists a unique branch in the (L, u)-plane of nontrivial solutions of problem (7.1) which emanates from the trivial solution at L_1. Its local behavior is given by*

$$\|u\|^2 \sim \frac{8}{3\pi}\sqrt{\alpha}(1+\sqrt{\alpha})^{3/2}(L_1 - L) \qquad \text{as } L \nearrow L_1.$$

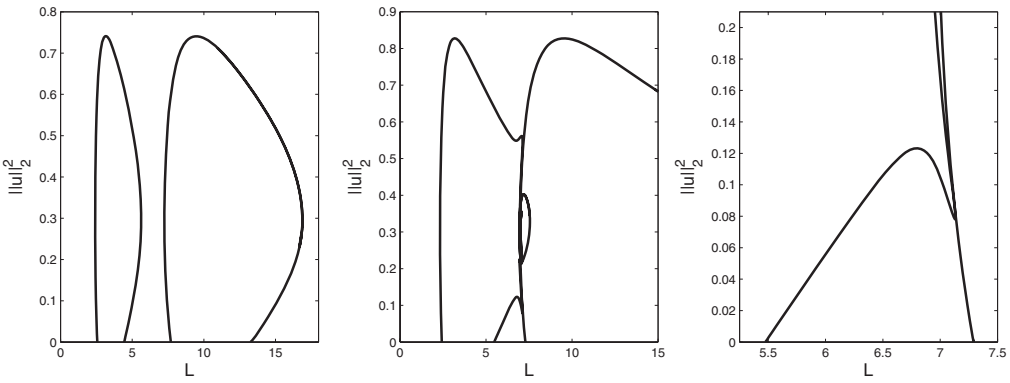


Figure 14. *Sample bifurcation diagrams for destabilizing nonlinearity. Notice that here the branches emanate from L_1 and L_2 in the opposite direction from that in the stabilizing case. Left: $\alpha = .42$. Middle: $\alpha = .49$. Right: Detail of $\alpha = .49$, showing the same local behavior as constructed in section 4.*

Theorem 7.2. *Let $0 < \alpha < 1$. There exists a unique branch in the (L, u)-plane of nontrivial solutions of problem (7.1) which emanates from the trivial solution at L_2. Its local behavior is given by*

$$\|u\|^2 \sim \frac{8}{3\pi}\sqrt{\alpha}(1-\sqrt{\alpha})^{3/2}(L_2 - L) \qquad \text{as } L \searrow L_2.$$

In Figure 14 we show two global bifurcation diagrams, one for $\alpha = 0.42$ and one for $\alpha = 0.49$, in which these properties clearly show up. We also see the characteristic structure when $3L1 \approx L_2$, which we studied in sections 4 and 6. However, here it "flipped over."

Because the solution branches bifurcating from the points $(kL_1, 0)$ and $(kL_2, 0)$ now point in the opposite direction from the one in the stabilizing case we need to fix $\alpha < \alpha^*_{m,n}$ to ensure that $nL_1 > mL_2$. Thus, in order to study the solution set when $3L_1 \approx L_2$, we write

$$\alpha = \alpha_1^* - \varepsilon \qquad \text{for } \varepsilon > 0,$$

and put

$$L = 3L_1 + \delta L_{\text{gap}}, \qquad \text{where} \qquad L_{\text{gap}} = L_2 - 3L_1.$$

Note that now $L_{\text{gap}} < 0$. Rescaling the equations as in sections 4 and 5, we now obtain the $\mathcal{O}(\varepsilon)$ system

$$\begin{cases} \dot{a} = -a\left\{\dfrac{10}{9}(1-\delta) - \dfrac{3}{2}a^2 + \dfrac{3}{2}ab - 3b^2\right\}, \\ \dot{b} = -10\delta b - \dfrac{1}{2}a^3 + 3a^2 b + \dfrac{3}{2}b^3 \end{cases} \tag{7.2}$$

when $n = 3m$. Reversing time and setting $t = -\tau$, we obtain system (4.4), where the dots now denote differentiation with respect to τ.

It is evident that for every $\delta \in [0,1]$ the family of stationary solutions is the same as that of (4.4), but, because L now *decreases* when δ *increases*, the bifurcation picture is the mirror image of the one in section 4.

If $n \neq 3m$, the analysis is similar to that in section 5.

Finally, because all the nontrivial equilibrium solutions of (4.4) are either stable nodes or saddles, the equilibrium solutions of (7.2) are all unstable nodes or saddles, and therefore all unstable. It is easily seen that the origin $(a,b) = (0,0)$ is now a stable node.

8. Conclusions. The bifurcation diagram of the classical Swift–Hohenberg equation (1.1) with parameter $\alpha \in (0,1)$, defined on a bounded domain $(0,L)$ with boundary conditions $u = 0$ and $u'' = 0$, is characterized by an infinite sequence of loop-type branches Γ_m which connect pairs of bifurcation points $\{mL_1, mL_2\}$, $m = 1,2,3,\ldots$, at which they bifurcate from the trivial solution $u = 0$. The critical lengths are given by $L_1 = \frac{1}{\sqrt{1+\sqrt{\alpha}}}$ and $L_2 = \frac{1}{\sqrt{1-\sqrt{\alpha}}}$ and depend on the value of the parameter $\alpha \in (0,1)$.

We have exhibited numerical results for *even* solutions, which show that the solution branches Γ_n are connected through *secondary bifurcations*, involving complex well-defined structures. Only two distinct types of such structures were observed:

1. those which connect Γ_n and Γ_m, where $n = 3m$, and
2. those which connect Γ_n and Γ_m, where $n \neq 3m$.

The qualitative properties of the structures around these secondary bifurcations were found to be very robust under perturbations of the parameters α and L.

By means of a center-manifold analysis these secondary bifurcations were analyzed when the distance between the points $nL_1(\alpha)$ and $mL_2(\alpha)$ is small, for instance, for values of α when

$$3L_1(\alpha) < L_2(\alpha) \qquad \text{and} \qquad L_2(\alpha) - 3L_1(\alpha) \ll L_2(\alpha).$$

These conditions are satisfied if $\alpha = \alpha_1^* + \varepsilon$, where $\alpha_1^* = 16/25$ and $\varepsilon > 0$ is small.

In this center-manifold analysis (i) the two types of secondary bifurcations and their complex structures were exactly retrieved, and (ii) the existence of only two types of such structures could be explained.

The stability properties of the solutions on these structures was also determined by means of the center-manifold analysis. It was found that all connecting structures are unstable, except the one involving Γ_1 and Γ_3, which has stable as well as unstable segments. In particular, it was found that there exists an interval $[\ell_1, \ell_2] \subset (3L_1, L_2)$ such that if $L \in (\ell_1, \ell_2)$, then there exist two *stable* stationary solutions.

It will be interesting to see how robust the structures found in this paper are when we turn to different systems.

Appendix A. Additional algebraic details. In order to prove the results of section 4 it will be convenient to express the null clines $\Gamma_1^{(2)}$ and Γ_2 in terms of polar coordinates, writing

$$a = r\cos(\theta) \qquad \text{and} \qquad b = r\sin(\theta). \tag{A.1}$$

By symmetry this is possible if the ray

$$R_\tau = \{(a,b) : b = \tau a, \quad b > 0\}, \qquad \tau \in \mathbf{R},$$

intersects each null cline at most at one point and is never tangent to one of the null clines. This is clearly the case for Γ_2, which is an ellipse.

In fact, it is also the case for $\Gamma_1^{(2)}$. Suppose to the contrary that for some τ_0 the ray R_{τ_0} is tangent to $\Gamma_1^{(2)}$ at the point (a_0, b_0). Then, at (a_0, b_0) we have (dropping the subscript "0")

$$a^3 - 6a^2b - 3b^3 = 20\delta b \qquad \text{and} \qquad b = \tau a,$$

so that

$$a^2(1 - 6\tau - 3t^3) = 20\delta\tau, \tag{A.2}$$

as well as

$$3a^2 - 12ab - 6a^2\tau - 9b^2\tau = 20\delta\tau.$$

Because $\Gamma_1^{(2)}$ and R_τ are tangent, we also have

$$3a^2(1 - 6\tau - 3\tau^3) = 20\delta\tau. \tag{A.3}$$

It follows that $1 - 6\tau - 3\tau^3 = 0$, so that $\tau = \tau^* \stackrel{\text{def}}{=} 0.16444\ldots$. This would imply that $\delta = 0$, a contradiction.

As a corollary we conclude by continuity that R_τ intersects $\Gamma_1^{(2)}$ at most at one point.

Thus, we can write $\Gamma_1^{(2)}$ as $r_1(\theta)$ for $\theta_0 < \theta < \pi$, where θ_0 will be determined later, and Γ_2 as $r_2(\theta)$ for $0 < \theta < \pi$. Using the change of variable (A.1) in (A.2) and (A.3), we obtain the following equations:

$$20\delta \sin(\theta) + r_1^2 F(\theta) = 0 \qquad \text{and} \qquad \frac{20}{27}(1 - \delta) = r_2^2 G(\theta), \qquad 0 < \theta < \pi, \tag{A.4}$$

where

$$F(\theta) = \cos^3(\theta) - 6\cos^2(\theta)\sin(\theta) - 3\sin^3(\theta)$$

and

$$G(\theta) = \cos^2(\theta) - \cos(\theta)\sin(\theta) + 2\sin^2(\theta).$$

It follows from the equations in (A.4) that θ needs to be so restricted that $F(\theta) < 0$ and $G(\theta) > 0$. When we divide $F(\theta)$ by $\sin^3(\theta)$, the first condition can be expressed as

$$\cot^3(\theta) - 6\cot^2(\theta) - 3 < 0 \qquad \Longrightarrow \qquad -\infty < \cot(\theta) < 1/\tau^*,$$

so that $\theta \in (\theta^*, \pi)$, where $\theta^* = \cot^{-1}(1/\tau^*)$. Because

$$G(\theta) = \left(\cos(\theta) - \frac{1}{2}\sin(\theta)\right)^2 + \frac{7}{4}\sin^2(\theta) > 0 \qquad \text{for all } \theta \in [0, 2\pi),$$

the second condition poses no restriction on θ.

Thus,

$$r_1^2(\theta) = -20\delta\frac{\sin(\theta)}{F(\theta)} \qquad \text{and} \qquad r_2^2(\theta) = \frac{20}{27}(1-\delta)\frac{1}{G(\theta)}. \tag{A.5}$$

At the points where the null clines intersect we have $r_1 = r_2$, so that

$$-20\delta\frac{\sin(\theta)}{F(\theta)} = \frac{20}{27}(1-\delta)\frac{1}{G(\theta)},$$

or

$$\sin(\theta)G(\theta) + \gamma F(\theta) = 0, \qquad \gamma = \frac{1-\delta}{27\delta}. \tag{A.6}$$

If we divide by $\sin^3(\theta)$ and write $x = \cot(\theta)$, we obtain the cubic polynomial equation

$$H(x;\gamma) \stackrel{\text{def}}{=} x^2 - x + 2 + \gamma(x^3 - 6x^2 - 3) = 0, \qquad -\infty < t < 1/\tau^*. \tag{A.7}$$

Plainly,

$$H'(x;\gamma) = 2x - 1 - \gamma(3x^2 - 12x) \qquad \text{and} \qquad H''(x;\gamma) = 2 - \gamma(6x - 12).$$

Thus $H'(x;\gamma) = 0$ if $x \in \{\xi_-, \xi_+\}$, where

$$\xi_\pm = \frac{1}{3\gamma}\left(6\gamma - 1 \pm \sqrt{(1-6\gamma)^2 + 3\gamma}\right).$$

It is easily verified that

$$H''(\xi_-;\gamma) < 0 \qquad \text{and} \qquad H''(\xi_+;\gamma) > 0.$$

Therefore (A.7) will have three zeros if $H(\xi_+;\gamma) < 0 < H(\xi_-;\gamma)$, two zeros if $H(\xi_-;\gamma) = 0$ or $H(\xi_+;\gamma) = 0$, and one zero if either $H(\xi_-;\gamma) < 0$ or $H(\xi_+;\gamma) > 0$. These conditions can easily be translated into conditions on γ. An elementary computation yields the values

$$\gamma_1 = 0.69224\ldots \qquad \to \qquad \delta_1 = 0.050785\ldots$$

and

$$\gamma_3 = 0.206767\ldots \qquad \to \qquad \delta_3 = 0.15191\ldots,$$

and that the null clines $\Gamma_1^{(2)}$ and Γ_2 intersect at three points when $\delta \in (\delta_1, \delta_3)$. At $\delta = \delta_2$, one of these critical point lies on the b-axis, i.e., $\theta = \pi/2$. Then $x = \cot(\theta) = 0$, and we deduce from (A.7) that $\gamma = 2/3$ and hence $\delta_2 = 0.05263158\ldots$.

Appendix B. Numerical approximation. In this paper we have used a continuation package for Matlab called MATCONT [12] for the computation of all bifurcation diagrams. This

solves the algebraic system of equations corresponding to the finite difference approximation to the ODE problem. Continuation is then done via pseudoarclength parametrization.

The fourth derivative is approximated with a five point stencil, and the second derivative with a three point stencil. Because of the symmetry properties of the solutions, the numerical problem was solved on $x \in [0, L/2]$ with the boundary conditions

$$u'(-L/2) = u'''(-L/2) = 0 \qquad \text{and} \qquad u'(L/2) = u'''(L/2) = 0.$$

Derivatives at the boundary are approximated with ghost points.

Because continuation is being done with respect to the interval length the spatial grid spacing changes along the branches. In all cases the number of grid points N was chosen such that the grid spacing h satisfied $h = L/\{2(N-1)\} < 1/10$.

REFERENCES

[1] J. B. van den Berg and R. C. A. M. van der Vorst, *Stable patterns for fourth order parabolic equations*, Duke Math. J., 115 (2002), pp. 513–558.

[2] G. J. B. van den Berg, L. A. Peletier, and W. C. Troy, *Global branches of multi-bump periodic solutions of the Swift-Hohenberg equation*, Arch. Ration. Mech. Anal., 158 (2001), pp. 91–153.

[3] E. Bodenschatz, W. Pesch, and G. Ahlers, *Recent developments in Rayleigh-Bénard convection*, Annu. Rev. Fluid Mech., 32 (2000), pp. 709–778.

[4] G. Caginalp and P. C. Fife, *Higher order phase field models and detailed anisotropy*, Phys. Rev. B (3), 34 (1986), pp. 4940–4943.

[5] J. Chaparova, *Existence and numerical approximation of periodic solutions of semilinear fourth order differential equations*, J. Math. Anal. Appl., 273 (2003), pp. 121–136.

[6] J. Chaparova, L. A. Peletier, and S. A. Tersian, *Existence and nonexistence of nontrivial solutions of semilinear sixth order ordinary differential equations*, Appl. Math. Lett., 17 (2004), pp. 1207–1212.

[7] J. Chaparova, L. A. Peletier, and S. Tersian, *Existence and nonexistence of nontrivial solutions of semilinear fourth- and sixth-order ordinary differential equations*, Adv. Differential Equations, 8 (2003), pp. 1237–1258.

[8] P. Collet and J. P. Eckmann, *Instabilities and Fronts in Extended Systems*, Princeton Ser. Phys., Princeton University Press, Princeton, NJ, 1990.

[9] M. C. Cross and P. C. Hohenberg, *Pattern formation outside of equilibrium*, Rev. Modern Phys., 65 (1993), pp. 851–1112.

[10] G. T. Dee and W. van Saarloos, *Bistable systems with propagating fronts leading to pattern formation*, Phys. Rev. Lett., 60 (1988), pp. 2641–2644.

[11] R. A. Gardner and C. K. R. T Jones, *Traveling waves of a perturbed diffusion equation arising in a phase field model*, Indiana Univ. Math. J., 38 (1989), pp. 1197–1222.

[12] A. Ghooge, W. Govaerts, and Yu. Kuznetsov, *MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs*, ACM Trans. Math. Software, 26 (2003), pp. 141–163.

[13] J. K. Hale, *Asymptotic Behavior of Dissipative Systems*, Math. Surveys Monogr. 25, AMS, Providence, RI, 1988.

[14] P. C. Hohenberg and J. B. Swift, *Effects of additive noise at the onset of Rayleigh-Bénard convection*, Phys. Rev. A (3), 46 (1992), p. 4773–4785.

[15] J. Lega, J. V. Moloney, and A. C. Newell, *Swift-Hohenberg equation for lasers*, Phys. Rev. Lett., 73 (1994), pp. 2978–2981.

[16] L. A. Peletier and V. Rottschäfer, *Large time behaviour of solutions of the Swift-Hohenberg equation*, C. R. Math. Acad. Sci. Paris, 336 (2003), pp. 225–230.

[17] L. A. Peletier and V. Rottschäfer, *Pattern selection of solutions of the Swift-Hohenberg equation*, Phys. D, 194 (2004), pp. 95–126.

[18] L. A. Peletier and W. C. Troy, *Spatial Patterns: Higher Order Models in Physics and Mechanics*, Birkhäuser Boston, Cambridge, MA, 2001.

[19] Y. Pomeau and P. Manneville, *Wave length selection in cellular flows*, Phys. Letters, 75A (1980), pp. 296–298.
[20] G. R. Sell and Y. You, *Dynamics of Evolutionary Equations*, Springer-Verlag, New York, 2002.
[21] J. B. Swift and P. C. Hohenberg, *Hydrodynamic fluctuations at the convective instability*, Phys. Rev. A (3), 15 (1977), pp. 319–328.
[22] W. Zimmerman, *Propagating fronts near a Lifschitz point*, Phys. Rev. Lett., 66 (1991), p. 1546.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 1, pp. 236–262

Coherent Structures Generated by Inhomogeneities in Oscillatory Media*

Richard Kollár† and Arnd Scheel‡

Abstract. We investigate the effect of spatially localized inhomogeneities on a spatially homogeneous oscillation in a reaction-diffusion system. In dimension up to two, we find sources and contact defects, that is, the inhomogeneity may either send out phase waves or act as a weak sink. We show that small inhomogeneities cannot act as sources in more than two space dimensions. We also derive asymptotics for wavenumbers and group velocities in the far field. The results are established rigorously for radially symmetric inhomogeneities in reaction-diffusion systems, and for arbitrary inhomogeneities in a modulation equation approximation.

Key words. reaction-diffusion, phase diffusion, eikonal equation, coherent structures, target patterns

AMS subject classifications. 37L10, 35K57, 34C37

DOI. 10.1137/060666950

1. Oscillatory reaction-diffusion systems.

1.1. Introduction.

We are interested in patterns that arise in dissipative spatially extended systems far from equilibrium. The arguably simplest nonequilibrium pattern in a dynamical system is a periodic orbit. Periodic orbits are ubiquitous in dynamical systems, a fact which is partly justified by their robustness. Indeed, when studying ordinary differential equations or partial differential equations posed on bounded domains, periodic orbits are typically robust: the trivial Floquet multiplier associated with the phase of the oscillation is algebraically simple, and for any small perturbation of the system, one will find a nearby periodic orbit with similar frequency.

Spatially extended large systems of oscillators have attracted attention in the physical and mathematical literature in many contexts. A classical prominent example is the Belousov–Zhabotinsky reaction, a reaction-diffusion system where the chemical concentrations undergo a relaxation-type oscillation which can be sustained for many cycles; see, for example, [25, 9]. Other examples include biological systems such as cardiac tissue [24], neural systems [23], and ecological systems [3].

When studying such large systems, in unbounded or in large domains, two interrelated issues complicate the concept of a *robust oscillation*.

First, robustness turns out to be a delicate issue on a technical level. In large domains, the fixed point problem for the Poincaré map is ill-conditioned due to clusters of eigenvalues

*Received by the editors August 5, 2006; accepted for publication (in revised form) by D. Barkley December 7, 2006; published electronically March 30, 2007.
http://www.siam.org/journals/siads/6-1/66695.html
†University of Michigan, Department of Mathematics, 530 Church Street, Ann Arbor, MI 48109 (kollar@umich.edu).
‡Corresponding author. University of Minnesota, School of Mathematics, 206 Church Street, S.E., Minneapolis, MN 55455 (scheel@umn.edu). This author was partially supported by the NSF through grants DMS-0203301 and DMS-0504271.

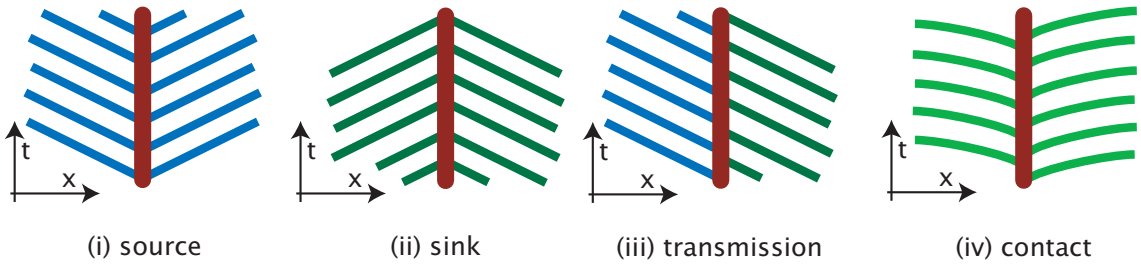


Figure 1. *Schematic space-time plots of the four types of coherent structures that were studied in* [17]. *The vertical brown bar denotes the location of a defect, where the pattern is not necessarily close to the wave trains. Blue and green lines denote lines of constant phase in the oscillation, with in- and outgoing group velocities, respectively. Sources* (i) *are isolated, they occur for discrete sets of asymptotic wavenumbers* $k_\pm$, *and group velocities point away from the center. Sinks* (ii) *come in two-parameter families, group velocities point towards the center, and asymptotic wavenumbers can be prescribed. For transmission defects* (iii), *group velocities point in the same direction,* $k_+k_- > 0$; *only the wavenumber with ingoing group velocity can be prescribed; the wavenumber with outgoing group velocity is selected. For contact defects* (iv), *the group velocities vanish at infinity. Asymptotic wavenumbers are typically equal.*

of the linearization near the neutral phase mode; in unbounded domains, the neutral phase mode is even embedded into a continuum of spectrum; see section 1.3 below. The spatial diffusive coupling is responsible for this lack of separation between *slow* phase modes and the *fast* normal modes near the periodic orbit, since it covers a full band of possible exponential relaxation rates. As a consequence, it is often not obvious whether periodic orbits are robust under changes of system parameters!

Second, periodic orbits come in very different spatial flavors: spatially homogeneous oscillations, plane waves, target patterns, and spiral waves, to name but a few. A perturbation theory for spatially extended systems should be able to distinguish between those different types of periodic solutions.

Wave trains, the simplest nonhomogeneous periodic solutions, are solutions where the phase Φ of the oscillation varies periodically in both time *and* space. Associated with this variation of the phase is the *group velocity* c_g of a wave train: small spatially localized variations of the phase $\Phi(t = 0) = \Phi_0(x)$ are simply advected to leading order, $\Phi(t, x) = \Phi_0(x - c_g t)$. Wave trains come in one-parameter families with parameter k, the wavenumber of the spatial variation of the phase. In [17], more general time-periodic solutions of reaction-diffusion systems posed on $x \in \mathbb{R}$ were classified according to their limiting behavior at $x = \pm\infty$. In addition to periodicity, the crucial assumption was convergence to wave trains in the far field. Such solutions, time-periodic and asymptotic to wave trains, were referred to as *defects* or *coherent structures*. The crucial property of such solutions turned out to be the sign of the group velocities at the wave trains at $\pm\infty$: physically, they determine the direction in which perturbations are transported, away or towards the center of the coherent structure; mathematically, they determine Fredholm indices of the linearized period map close to the trivial multiplier 1, and hence give multiplicity results. In summary, the physically relevant "shape," as determined by the group velocities, determines mathematical robustness and multiplicity properties; see Figure 1 for a short summary of the results, connecting transport and

multiplicity properties.

The simplest scenario, where the different types of coherent structures emerge, is when a spatially extended system with a spatially homogeneous oscillation is perturbed by introducing a spatially localized inhomogeneity. It was noted in [17] that wave trains may nucleate at the small inhomogeneity, leading to sources. We pick up this example in this article and study the case of space dimensions two and higher. Our main results characterize the existence of coherent structures, that is, time-periodic solutions which converge to wave trains as $|x| \to \infty$ for small inhomogeneities. We exhibit a dichotomy similar to the one-dimensional case, between sources, where group velocities point away from the inhomogeneity in the far field, and contact defects, where group velocities point towards the inhomogeneity but converge to zero as $|x| \to \infty$. In space dimensions less than or equal to two, either sources or contact defects bifurcate, and we give expansions for wavenumbers and group velocities in terms of the strength of the inhomogeneity. In space dimensions larger than two, one finds contact defects only for small inhomogeneities. We prove our results for general reaction-diffusion systems in the case of radially symmetric inhomogeneities. We also formally derive a viscous eikonal equation, for which we prove existence and expansions in the general case, without radial symmetry.

In the remainder of this first section, we will set the scene and explain our main results. The basic setup of oscillations, wave trains, and group velocities is given in section 1.2; we define and characterize coherent structures in section 1.3; we give our main results in section 1.4; and we conclude with an outline of the remainder of this article.

1.2. Oscillations, wave trains, and modulations. As a prototype for nonequilibrium spatially extended systems, we consider reaction-diffusion systems

$$u_t = D\triangle u + f(u), \tag{1.1}$$

$u \in \mathbb{R}^N$, $D = \operatorname{diag}(d_j) > 0$, $f \in \mathbb{C}^\infty$, in $x \in \mathbb{R}^n$. We assume the existence of an *asymptotically stable spatially homogeneous oscillation*. The following list of assumptions roughly states that the spectrum of the period map of (1.1), linearized at a periodic solution, is as stable as possible: it is strictly contained in the unit circle up to a curve of spectrum touching the unit circle at $\lambda = 1$ with a quadratic tangency. More precisely, we assume that there exists a solution $u(t,x) = u_*(-\omega_* t) = u_*(2\pi - \omega_* t)$ to (1.1), and we define

$$L_k u = -Dk^2 u + \omega_* u_\tau + f'(u_*(\tau))u, \quad L_k : \mathcal{D}(L) = H^1_{\mathrm{per}}(0,2\pi) \to L^2(0,2\pi). \tag{1.2}$$

For stability, we assume that

(i) $\operatorname{spec} L_k \cap \{\lambda \neq 0 \bmod \mathrm{i}\omega_*, \operatorname{Re}\lambda \geq 0\} = \emptyset$;
(ii) $\operatorname{spec} L_k \cap \{\operatorname{Re}\lambda \geq 0\} = \emptyset$ for $k \neq 0$;
(iii) $\lambda = 0$ is algebraically simple as an eigenvalue of $\operatorname{spec} L_0$;
(iv) the curve of eigenvalues $\lambda(k)$ to L_k with $\lambda(0) = 0$ satisfies $d = -\lambda''(0) > 0$.

Examples of such oscillations include all stable periodic solutions to the pure kinetics $u_t = f(u)$ if the diffusion matrix is the identity (or close to the identity), an assumption that is typically satisfied for reactions in solvent. It also includes the example of small amplitude oscillations whenever the Benjamin–Feir stability criterion on sideband instability is satisfied. The chemical oscillations observed in the Belousov–Zhabotinsky reaction provide a prominent experimental example.

Stable homogeneous oscillations are accompanied by a family of wave trains $u_*(k\cdot x-\omega t;|k|)$, which solve

$$|k|^2Du''+\omega u'+f(u)=0. \tag{1.3}$$

It is not difficult to see that there exists a smooth family of such 2π-periodic wave train solutions $u_*(\cdot;|k|)$, with ω given as a function of the parameter k. We refer to this dependence $\omega=\Omega(k)$ as the *nonlinear dispersion relation*. Note that due to isotropy of our medium, $\Omega(k)$ is merely a function of $|k|$. Also note that we slightly abuse notation, writing $u_*(\cdot)=u_*(\cdot,0)$ for the spatially homogeneous oscillations. We define the group velocity of a wave train $u_*(\cdot;k)$ via

$$c_{\mathrm{g}}(k)=\nabla\Omega(k). \tag{1.4}$$

Again, by isotropy,

$$c_{\mathrm{g}}(k)=|c_{\mathrm{g}}(k)|\frac{k}{|k|}.$$

We emphasize that the existence of wave trains already shows that homogeneous oscillations are not isolated in phase space, and any robustness result needs to take the occurrence of families of solutions into account.

More general solutions can be found by varying the wavenumber of the wave trains on slow spatio-temporal scales. Inserting the ansatz

$$u(t,x)=u_*(-\omega_*t-\Phi(T,X);\varepsilon\Phi_X),\quad X=\varepsilon x,\quad T=\varepsilon^2t, \tag{1.5}$$

into (1.1), we find at order ε^2 a viscous eikonal equation

$$\Phi_T=d\Delta_X\Phi-\frac{1}{2}\Omega''(0)|\Phi_X|^2; \tag{1.6}$$

see, for instance, [5]. In particular, the nonlinear dispersion relation can be interpreted as a nonlinear flux in a transport equation, and c_{g} provides precisely the speed of characteristic transport. The viscosity d is defined in item (iv) of our list of linear stability assumptions.

1.3. Inhomogeneities and coherent structures. Our interest here is in the effect of small inhomogeneities in the medium on the oscillations. We therefore consider

$$u_t=D\triangle u+f(u)+\varepsilon g(x,u) \tag{1.7}$$

for ε small, with $g\in C^\infty$ smooth and localized, $|g(x,u)|=O(|x|^{-2-\beta})$ for $|x|\to\infty$ and some $\beta>0$.

One may be tempted to analyze this perturbation problem in a large ball of radius L and Neumann boundary conditions, for instance, rather than in unbounded domains. In this case, the Poincaré map possesses an asymptotically stable fixed point which can be readily seen to persist for small values of ε. The validity of this *regular perturbation argument* in a bounded

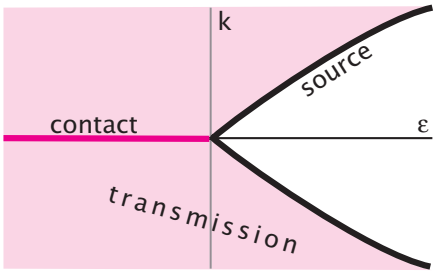


Figure 2. *Bifurcation diagram for small inhomogeneities in one-dimensional media. Sources exist for positive ε with emitted wavenumbers $k_\pm \sim \pm\varepsilon$. Contact defects exist for $\varepsilon < 0$. The sector between sources and contact defects is filled with transmission defects; see section 3.2.*

domain, however, is restricted to $\varepsilon = \mathrm{O}(L^{-2})$ at best, since the trivial Floquet exponent is isolated in the spectrum of the period map by a gap of this size only in a large domain.

Our goal in this article is to describe patterns resulting from the introduction of inhomogeneities using *perturbation theory in the presence of essential spectrum* in unbounded domains. The idealization of the domain as the entire space has the additional advantage of characterizing the resulting patterns by their asymptotic profile in the limit $|x| \to \infty$.

We distinguish solutions to the inhomogeneous reaction-diffusion system in terms of their transport properties at infinity. In one space dimension, this task was carried out in a systematic fashion in [17]. We borrow some of the terminology from there and generalize to the multidimensional case next. In [17], we say a solution $u_\mathrm{c}(x,t)$ to the (possibly inhomogeneous) reaction-diffusion system is a *coherent structure* if u_c is

- *time-periodic*: $u_\mathrm{c}(x, t + \frac{2\pi}{\omega_\mathrm{c}}) = u_\mathrm{c}(x,t)$ for some $\omega_\mathrm{c} > 0$;
- *localized*: there are $k_\pm$ such that $|u_\mathrm{c}(x,t) - u_*(k_\pm x - \omega_\mathrm{c} t - \varphi(x); k_\pm)| \to 0$ for $x \to \pm\infty$, uniformly in t, for some $\varphi(x)$ with $\varphi'(x) \to 0$.

Note that "localized" does not refer to decay, but rather implies convergence as $x \to \pm\infty$ towards wave trains with certain asymptotic wavenumbers $k_\pm$. We can then classify coherent structures according to the group velocity $c_\mathrm{g}(k_\pm)$ at these asymptotic wave trains. We say u_c is a

- *source* if $c_\mathrm{g}(k_-) < 0 < c_\mathrm{g}(k_+)$;
- *sink* if $c_\mathrm{g}(k_-) > 0 > c_\mathrm{g}(k_+)$;
- *contact defect* if $c_\mathrm{g}(k_-) = 0 = c_\mathrm{g}(k_+)$;
- *transmission defect* if $c_\mathrm{g}(k_-) \cdot c_\mathrm{g}(k_+) > 0$.

Note that the reflection symmetry of the equation and the uniqueness of the family of wave trains solving (1.3) imply that $\Omega(k) = \Omega(-k)$. The time periodicity of coherent structures implies $\Omega(k_+) = \Omega(k_-) = \omega_\mathrm{c}$, which typically, for example if $\Omega''(0) \neq 0$, implies that $|k_-| = |k_+|$. In [17, section 6.5], we showed that for small inhomogeneities there is a dichotomy between the existence of sources and contact defects. A schematic bifurcation diagram is shown in Figure 2. The analysis was outlined in [17], but we include a sketch in section 3.2 here for the convenience of the reader.

In case of radially symmetric defects, the classification is readily adapted. We call a radially symmetric solution $u_\mathrm{c}(r,t)$ a *coherent structure* if u_c is

- *time-periodic*: $u_c(t+\frac{2\pi}{\omega_c},r)=u_c(t,r)$ for some $\omega_c>0$;
- *localized*: there exists k such that $|u_c(r,t)-u_*(kr-\omega_c t-\varphi(r);k)|\to 0$ for $r\to\infty$, uniformly in t, and $\varphi'(r)\to 0$.

The classification in this radial multidimensional case is somewhat more restrictive: we say u_c is a

- *source* if $c_g(k)>0$;
- *sink* if $c_g(k)<0$;
- *contact defect* if $c_g(k)=0$.

1.4. Main results: Radially symmetric inhomogeneities. Our main technical results describe coherent structures in spatially homogeneous oscillations generated by radially symmetric inhomogeneities in reaction-diffusion systems. A formal perturbation theory for the periodic solution u_* would isolate the time-derivative u_*' as the kernel in the fixed point equation for the linearized period map (1.1). Similarly, one finds a unique bounded solution $u_{ad}(-\omega_* t)$ to the adjoint linearized kinetics, $u_t=-f'(u_*)^T u$. We assume that u_{ad} is normalized such that $\int(u_{ad}(\tau),u_*'(\tau))\mathrm{d}\tau=1$, where $(\cdot,\cdot)$ denotes the scalar product in $\mathbb{R}^N$. Since our perturbation is dependent on x, we would like to use these eigenfunctions to carry out a perturbation theory in $L^2(\mathbb{R}^n)$ or $BC^0(\mathbb{R}^n)$, say. If the linearized operator were Fredholm, one would proceed with Lyapunov–Schmidt reduction: one evaluates the perturbation on the kernel and projects it back onto the kernel by taking the scalar product with the kernel of the adjoint, and by then normalizing with the scalar product between kernel and cokernel. This procedure fails at the first step: since u_*' is not localized as a function of x, the linearization is not Fredholm; see, for example, [19, Lemma 6.4]. One may still formally continue to derive an expansion for a perturbed periodic solution by projecting the perturbation, evaluated in the periodic solution, onto the kernel, using the adjoint kernel. However, the fact that the space-time $L^2(S^1\times\mathbb{R}^n)$-scalar product between u_{ad} and u_*' diverges indicates strongly that the formal results obtained in this fashion will not be valid.

Nevertheless, we define the Melnikov-type coefficient

(1.8)
$$M=(\mathrm{Vol}(S^{n-1}))^{-1}\int_x\int_\tau(u_{ad}(\tau),g(|x|,u_*(\tau)))\,\mathrm{d}\tau\mathrm{d}x=\int_0^\infty\int_\tau(u_{ad}(\tau),g(r,u_*(\tau)))\,\mathrm{d}\tau r^{n-1}\mathrm{d}r,$$

which precisely represents the projection of the perturbation on the kernel. As we will see, this coefficient is the essential ingredient to our main result. Our result is stated for radially symmetric solutions, and we therefore may consider the space dimension as a continuous parameter, $n\geq 1$. The theorem describes coherent structures close to the homogeneous oscillation, for ε small. *Close* here refers to the existence of an appropriate δ_0, small, and $\phi=\phi(x)$ such that

(1.9)
$$\sup_x|\nabla\phi(x)|<\delta_0,\qquad \sup_x|u(t,x)-u_*(-\omega_* t+\phi(x))|_{H^1_t}<\delta_0,$$

where H^1_t refers to the H^1-norm with respect to the variable $t\in[0,2\pi/\omega_*]$.

Notation. We denote by $\mathrm{O}(y)$ and $\mathrm{o}(y)$ the Landau symbols for functions h which vanish at $y=0$ and which satisfy $|h/y|\leq C$ for some $C>0$ and $\lim_{y\to 0}|h/y|=0$, respectively. We write $\mathrm{o}_y(1)$ for functions which converge to zero as $y\to 0$.

Theorem 1.1. *Consider the reaction-diffusion system* (1.7) *with* ε *small in space dimension* $n \leq 2$.

First assume that $\varepsilon M\Omega''(0) > 0$. *Then there exists a constant* $\hat{c}$ *with* $\hat{c}\Omega''(0) > 0$ *such that for all* $\varepsilon > 0$ *sufficiently small there exists a unique source with emitted wavenumber*

$$\begin{aligned} k(\varepsilon) &= \hat{c}|M\varepsilon|^{\frac{1}{2-n}}(1+\mathrm{o}_\varepsilon(1)) \quad \text{for } n<2,\\ k(\varepsilon) &= \hat{c}\mathrm{e}^{-\frac{2d}{\Omega''M\varepsilon}}(1+\mathrm{o}_\varepsilon(1)) \quad \text{for } n=2. \end{aligned} \tag{1.10}$$

In particular, the group velocity in the far field points outward.

Next assume that $\varepsilon M\Omega''(0) < 0$. *Then there exists a constant* $\hat{c}$ *with* $\hat{c}\Omega''(0) < 0$ *such that for all* ε *sufficiently small there exists a unique contact defect; that is, the asymptotic wavenumber is* $k = 0$. *At large but finite distances from the center, we have the wavenumber asymptotics*

$$k = \frac{1}{r}(\hat{c}+\mathrm{o}_{1/r}(1)), \quad n<2 \text{ and } r \gg \log\varepsilon, \qquad k = \frac{|\varepsilon M|}{|\varepsilon M| r\log r + r}(\hat{c}+\mathrm{o}_{1/r}(1)), \quad n=2.$$

In particular, the group velocity converges to zero in the far field, but it is pointing inward at large finite r. *There also exists a family of sinks with wavenumbers close to* 0 *for all small* ε. *There do not exist sources close to the homogeneous oscillations.*

In space dimension $n > 2$ *there exists a unique branch of contact defects and accompanying sinks, with wavenumber asymptotics*

$$k(r,\varepsilon) = \frac{M\varepsilon}{r^{n-1}}(\hat{c}+\mathrm{o}_{1/r}(1)), \quad n>2,$$

for an appropriate constant $\hat{c} \neq 0$. *There do not exist sources regardless of the sign of* $\varepsilon M\Omega''(0)$.

Remark 1.2. Note that the phase $\phi(r) = \int_0^r k(s)\mathrm{d}s$ diverges for contact defects for $n \leq 2$, logarithmically in one space dimension, and very weakly, $\varphi \sim \log\log r$ in two space dimensions. In higher space dimensions, the phase converges. In particular, the target sources or target sinks created by inhomogeneities in three space dimensions will have only a finite number of rings, moving outward or inward, depending on the sign of $\varepsilon M\Omega''(0)$.

The proof of the theorem will occupy section 2. Coherent structures solve a degenerate elliptic PDE in x, t, with periodic boundary conditions in t and somewhat intricate boundary conditions at $|x| = \infty$. The main difficulty stems from the boundary conditions at infinity, which require the solution to be pointwise in x close to a homogeneous oscillation $u_*(\cdot+\phi)$ for some shift $\phi(x)$, where the function ϕ *and* its asymptotics as $|x| \to \infty$ are to be determined as part of the analysis.

We illustrate the main result with numerical simulations. Figure 3 shows snapshots of the dynamics of an oscillatory system in the presence of a localized inhomogeneity. The color coding reflects values of the v-component in Barkley's FitzHugh–Nagumo model

$$u_t = \Delta u + \frac{1}{\mu}u(1-u)\left(u - \frac{v+b}{a}\right), \quad v_t = \Delta v + u - v + \frac{\varepsilon}{1+|x/3|^2} \tag{1.11}$$

with parameters $a = 0.3$, $b = -0.45$, $\mu = 0.095$ on $\Omega = \{|x_j| \leq 90\}$. The convergence to the final state is much faster in the case of a source.

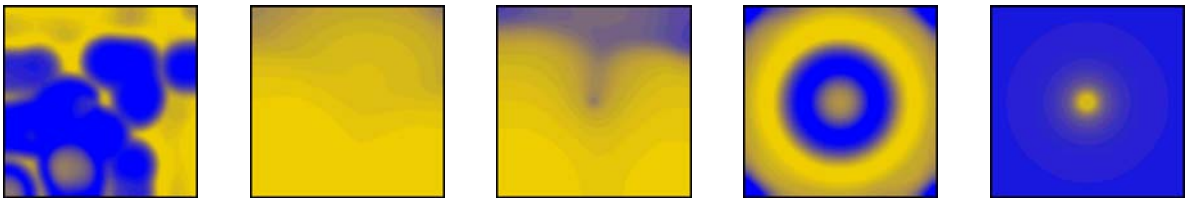


Figure 3. *Plotted are time snapshots of the inhibitor concentrations $v(x)$ of (1.11). From left to right, a spatially irregular initial condition (first) evolves into a homogeneous oscillation after time $T = 100$ (second) in a homogeneous medium. The next two pictures show how the same initial condition has evolved into a contact defect, $\varepsilon = -0.2$, and a source, $\varepsilon = 0.2$. The last picture shows the radially symmetric contact defect, which forms from a homogeneous initial condition, $\varepsilon = -0.2$. Simulations based on* EZ-SPIRAL *[2]. Clicking on the above images displays the accompanying movies (66695_01.mpg [10.3MB], 66695_02.mpg [10.4MB], 66695_03.mpg [12.9MB], and 66695_04.mpg [1.5MB]) of the time evolution.*

1.5. Outline. The remainder of this article is organized as follows. We introduce spatial dynamics and prove Theorem 1.1 in section 2. We rely on an ill-posed dynamical systems formulation, inspired by [8, 20]. The main difficulty is the analysis of an ODE in the far field, which possesses a highly degenerate equilibrium. We unfold the degeneracy using geometric blow-up methods; see [18, 11] for a recent account. We then discuss coherent structures which are not necessarily radially symmetric. We therefore derive a viscous eikonal equation that approximates the reaction-diffusion dynamics and discuss shapes of coherent structures in this approximation; see section 3. This discussion largely relies on the Hopf–Cole transformation, which links coherent structures in the eikonal equation to eigenfunctions of Schrödinger operators. We show that coherent structures still are approximately radial in the far field. We conclude section 3.3 with a discussion of moving inhomogeneities. Main results here again rely on the Hopf–Cole transformation. Coherent structures show more complicated nonradial patterns such as sonic cones and diffusive profiles. We conclude with a discussion in section 4, including a summary of results, discussion of topological and quantitative classifications of defects, and a note on asymptotic stability.

The results on reaction-diffusion systems, Theorem 1.1, are new. Some of the results on modulation equations in section 3 appear to be folklore. We refer to [22] and the references therein for related results and an overview of the experiments. We include this discussion of inhomogeneities in the eikonal approximation since it allows for a comparison with our main technical result, Theorem 1.1, and for extensions to nonsymmetric settings.

2. Radial dynamics. In this section, we prove Theorem 1.1. Coherent structures solve

$$-\omega u_\tau = Du_{rr} + \frac{n-1}{r}u_r + f(u) + \varepsilon g(r,u), \qquad \tau \in \mathbb{R} \bmod 2\pi\mathbb{Z},\ r \geq 0,$$

which we can (formally) rewrite as an abstract first-order differential equation in the radius r,

$$\begin{aligned} u_r &= v, \\ v_r &= -\frac{n-1}{r}v - D^{-1}\left(\omega\partial_\tau u + f(u) + \varepsilon g(r,u)\right). \end{aligned} \tag{2.1}$$

We consider (2.1) on the Hilbert space $X = H^{1/2}_{\mathrm{per}}(0,2\pi) \times L^2_{\mathrm{per}}(0,2\pi)$. The unbounded principal part of the right-hand side of (2.1) defines a closed operator with domain of definition

$X^1 = H^1_{\mathrm{per}}(0,2\pi) \times H^{1/2}_{\mathrm{per}}(0,2\pi)$. We say that $(u,v)(r,\tau)$ is a solution to (2.1) on $J \subset \mathbb{R}_+$ if $(u,v)(r,\cdot)$ is in $C^0(J,X) \cap C^1(\mathrm{Int}(J),X) \cap C^0(\mathrm{Int}(J),X^1)$. Note that the domain reflects the expected parabolic regularity with two derivatives in x *or* one derivative in t; see [8, 20] and [15, Lemma 3.1] for a simple computation justifying this choice of norms.

Note that (2.1) is invariant under the time shift $T_\phi : (u,v)(r,\cdot) \mapsto (u,v)(r,\cdot+\phi)$, which therefore maps solutions to solutions.

The differential equation (2.1) is ill-posed as a dynamical system, as its principal part amounts to solving the heat equation as an initial-value problem "sideways" in r.

Our strategy for the proof now consists of several steps. We construct manifolds of bounded solutions close to $(u_*(\cdot+\varphi),0)$ for $r \in [0,R]$, W^{cu}_- (Step 1), and for $r \in [R,\infty]$, W^{cs}_+ (Step 2). We then show that these two manifolds intersect transversely along $(u_*(\cdot+\varphi),0)$ (Step 3). In Step 4, we construct a two-dimensional manifold W^{c}_+ that contains this intersection, and we compute the vector field on this manifold in Step 5. Steps 6 and 7 are concerned with an analysis of this vector field, in particular tracking the points on W^{c}_+ that yield contact defects, sources, or sinks in $r \geq R$. The analysis involves a geometric blow-up construction, Step 6, and a Dulac map analysis, Step 7. In Step 8, we locate those points in W^{c}_+ which yield bounded solutions on $[0,R]$ to leading order in ε. In the final Step 9, we match these bounded solutions on $[0,R]$ with conditions for sources, sinks, and contact defects inside of W^{c}_+.

The general strategy is reminiscent of [20], where radially symmetric and time-periodic patterns have been studied close to a time- and space-independent equilibrium of the reaction-diffusion system. We will encounter some additional difficulties in the far field, which are similar to the difficulties arising in the study of one-dimensional contact defects; see [17, 18].

Step 1: Construction of W^{cu}_-.

Proposition 2.1. *For all $R > 0$ there exists a smooth manifold*

$$W^{\mathrm{cu}}_- = \bigcup_{0<r\leq R} W^{\mathrm{cu}}_-(r) \times \{r\} \subset X \times \mathbb{R}_+$$

such that $W^{\mathrm{cu}}_-(r)$ contains precisely all functions $(u(r),v(r))$ which are boundary values to bounded solutions of (2.1) on $(0,r]$, close to $(u_*(\cdot+\phi),0)$ in X for some $\phi \in \mathbb{R}$. The manifold is invariant under the action of the temporal shifts T_φ on X, and depends smoothly on ω and ε. Moreover, the map from solutions to boundary values is a smooth diffeomorphism, equivariant with respect to T_φ.

Proof. The proof is the same as in [20, Proposition 4.7]. ∎

Step 2: Construction of W^{cs}_+. We now turn to a description of the bounded solutions at $r = \infty$. We therefore introduce $\alpha = 1/r$ and rewrite (2.1) as

$$\begin{aligned} u_r &= v, \\ v_r &= -(n-1)\alpha v - D^{-1}\left(\omega\partial_\tau u + f(u) + \varepsilon g\left(\frac{1}{\alpha}, u\right)\right), \\ \alpha_r &= -\alpha^2. \end{aligned} \tag{2.2}$$

Note that the decay assumption on g implies that $g = \mathrm{O}(\alpha^{2+\beta}) \in C^2$ for $\alpha \geq 0$. In $\alpha = 0$, there is a circle of equilibria $u = u_*(\cdot+\varphi)$, $v = 0$. We are interested in solutions that stay in

a vicinity of this circle for all α small. We therefore consider the linearization in one of these equilibria:

$$\begin{aligned} u_r &= v, \\ v_r &= -D^{-1}\left(\omega_* \partial_\tau u + f'(u_*)\right), \\ \alpha_r &= 0. \end{aligned} \tag{2.3}$$

This system can be written in the short form $u'' = -D^{-1}L_0 u$, so that the spectrum of the right-hand-side operator in (2.3) is given by the square root of the spectrum of $-D^{-1}L_0$ (and the zero eigenvalue from the equation for α). In particular, there is no purely imaginary spectrum outside of zero since $D^{-1}L_0 u = \gamma^2 u$ for some $\gamma \neq 0$ violates our stability assumption (ii). All other eigenvalues come in pairs $(\nu, -\nu)$ by reversibility. The zero eigenvalue is geometrically simple, with kernel spanned by $(u_*', 0)$, again by the stability assumption. Then $(0, u_*')$ provides an obvious generalized eigenfunction. The quadratic expansion of the dispersion relation guarantees that the algebraic multiplicity is not larger than two; see [20].

We therefore have a $2+1$-dimensional center eigenspace E^{c} consisting of the generalized kernel, and infinite-dimensional stable and unstable eigenspaces $E^{\mathrm{s/u}}$.

Proposition 2.2 (W^{cs}_+). *There exists a smooth center-stable manifold to the circle of equilibria $(u_*(\cdot+\varphi), 0)$ at $\alpha = 0$, tangent to $E^{\mathrm{c}} \oplus E^{\mathrm{s}}$. Moreover, there is a local smooth semiflow Φ_r on W^{cs}_+, and all solutions that stay in a neighborhood of the circle of equilibria for all positive $r \geq R$ are trajectories to this semiflow on W^{cs}_+. Finally, the center-stable manifold is invariant, and the flow is equivariant with respect to the action of the symmetry group T_φ.*

Proof. The proof is analogous to [17, Theorem 3] and will be omitted here. ∎

Step 3: Intersecting W^{cu}_- and W^{cs}_+. By construction, the circle of equilibria at $(u_*(\cdot+\varphi), 0)$ belongs to the intersection of W^{cs}_+ and W^{cu}_-.

Proposition 2.3. *The manifolds W^{cs}_+ and W^{cu}_- intersect transversely along $(u_*(\cdot+\varphi), 0)$ in X at any fixed, finite $r \in (0, \infty)$. The intersection W^{c} is a circle, smoothly depending on ω, ε, and r.*

Proof. We show that tangent spaces intersect transversely along a one-dimensional subspace. Since the intersection necessarily contains the circle $(u_*(\tau+\varphi), 0)$ at $\varepsilon = 0$, $\omega = \omega_*$, this suffices to prove the theorem via Lyapunov–Schmidt reduction.

Transversality is encoded in the immersion map

$$\iota : T_* W^{\mathrm{cu}}_-(r) \times T_* W^{\mathrm{cs}}_+(r) \to X, \qquad (\mathbf{u}^{\mathrm{cu}}_-(r), \mathbf{u}^{\mathrm{cs}}_+(r)) \mapsto \mathbf{u}^{\mathrm{cu}}_-(r) - \mathbf{u}^{\mathrm{cs}}_+(r),$$

where $\mathbf{u}^{\mathrm{cu/cs}}_\pm = (u, v)$ solve the linearized equation

$$\begin{aligned} u_r &= v, \\ v_r &= -\frac{n-1}{r} v - D^{-1} L_0 u. \end{aligned} \tag{2.4}$$

Transverse intersection along a one-dimensional manifold is equivalent to ι being Fredholm of index one with minimal one-dimensional kernel. From [20, 16], we know that ι is Fredholm and the Fredholm index is given by the relative Morse index at $+\infty$. In order to compute the

relative Morse index, we compare the equation to the linearized equation

$$\begin{aligned} u_r &= v, \\ v_r &= -\frac{n-1}{r}v - D^{-1}L_0 u + \lambda u, \end{aligned}$$

with spectral parameter λ. For $\lambda > 0$, this equation possesses an exponential dichotomy; the immersion map ι is Fredholm of index 0. In the limit $\lambda = 0$, there is a double center eigenvalue, which we incorporated in the stable direction $T_* W_+^{\mathrm{cs}}$. For small positive λ, these two eigenvalues split in opposite directions. In other words, $T_* W_+^{\mathrm{cs}}$ is enlarged by one dimension compared to the hyperbolic Fredholm index 0 situation at $\lambda > 0$. A bordering lemma then shows that ι is Fredholm of index 1. It remains to show that the intersection is transverse, that is, that the kernel of ι is minimal one-dimensional.

We can block-diagonalize (2.4) by splitting off the one-dimensional kernel of $D^{-1}L_0$ with the associated spectral projection. In this center subspace, we find the unique bounded solution $u = u_*'(\tau)$, $v = 0$. Any linearly independent solution will exhibit a singularity at $r = 0$: indeed, solutions solve $u'' + \frac{n-1}{r}u' = 0$, with singularity r^{2-n}, $n \neq 2$, and $\log r$, $n = 2$, for any nonconstant solution. In the complement, any solution in the intersection would need to be exponentially localized at infinity. Any element in the kernel of ι in this hyperbolic component would therefore yield an exponentially localized solution to the system $D\Delta u + f'(u_*(\tau))u + \omega_* \partial_\tau u = 0$. In particular, the Fourier transform $\hat{u}(k, \tau)$ of this exponentially localized solution would be smooth in Fourier space and contribute to the kernel of L_k. Since all L_k with $k \neq 0$ are invertible by assumption, this implies $\hat{u} \equiv 0$ and concludes the proof. ∎

Step 4: Extending the intersection—W_+^{c}. The intersection W^{c} typically crosses the boundary of W_+^{cs} for $r \to \infty$; that is, solutions do not stay close to u_* for $r \to \infty$ for arbitrary parameters ε and ω. The construction of the center-stable manifold W_+^{cs} incorporates all solutions with mild growth. In order to single out the solutions that actually stay bounded as $r \to \infty$, we will analyze the flow in the center direction at $r = \infty$ more carefully. We therefore construct a center manifold $W_+^{\mathrm{c}} \subset W_+^{\mathrm{cs}}$, which is tangent to E^{c} at $r = \infty$ and contains W^{c}.

Proposition 2.4. *There exists a* $2+1$*-dimensional smooth center-manifold near the circle of equilibria* $(u(\cdot + \varphi), 0)$ *at* $\alpha = 0$*, which is tangent to* E^{c} *and contains all solutions that are bounded on* $(0, \infty)$*,*

$$W^{\mathrm{c}} \subset W_+^{\mathrm{c}} \subset W_+^{\mathrm{cs}},$$

for sufficiently large r*.*

Proof. The proof is analogous to [20, Theorems 3.7, 4.8]. ∎

Step 5: The vector field on W_+^{c}. Invariance of W_+^{c} together with the condition on tangency allows us to compute the Taylor expansion of the reduced vector field. We denote eigenvectors and adjoint eigenvectors in the generalized kernel by

$$e_0 = (u_*', 0), \quad e_1 = (0, u_*'), \quad e_0^* = (Du_{\mathrm{ad}}, 0), \quad e_1^* = (0, Du_{\mathrm{ad}}),$$

where $L_0^* u_{\mathrm{ad}} = 0$, and we normalize $(Du_{\mathrm{ad}}, u_*') = 1$, so that $(e_i, e_j^*) = \delta_{ij}$. We introduce coordinates on the center manifold by parameterizing the tangent bundle of the circle of

equilibria with θ, the symmetry action, and κ:

$$(u, v) = (u_*'(\cdot + \theta), 0) + \kappa e_1(\cdot + \theta) + \psi[\kappa](\cdot + \theta),$$

where ψ denotes the (symmetry-invariant) graph of the center manifold, $(e_j^*(\cdot), \psi(\cdot))) = 0$. The equation on the center manifold is independent of θ, by symmetry, and it therefore suffices to track the κ- and $\alpha = 1/r$-dependence of the vector field. A straightforward expansion shows that we necessarily recover the phase diffusion equation (1.6) at second order:

$$\begin{aligned} d(\kappa' + (n-1)\alpha\kappa) &= (\omega - \omega_*) + \frac{1}{2}\Omega''(0)\kappa^2 + \mathrm{o}(|\kappa|^2 + |\alpha|^2) + \mathrm{O}(|\omega - \omega_*|(|\kappa|^2| + |\alpha|^2)), \\ \alpha' &= -\alpha^2. \end{aligned} \tag{2.5}$$

The ε-dependent terms contribute only to the higher-order terms because of the rapid decay of g.

Step 6: Geometric blow-up. The vector field on W_+^{c} possesses a doubly degenerate equilibrium at $\alpha = \kappa = 0$: the linearization at this equilibrium vanishes, and leading-order terms are quadratic in α and κ. This degeneracy is unfolded by the parameter ω. Varying ω, the equilibrium undergoes a saddle-node bifurcation inside $\alpha = 0$. It turns out that we are interested in connections to the unstable equilibrium in this saddle-node bifurcation. In order to track its stable manifold, we invoke a desingularization method, the geometric blow-up [6, 11]. In fact, the system on the center manifold is similar to the system studied in [18], to which we will refer for more details on the construction. The similarity to the eigenvalue problem for the n-dimensional Laplacian in [18] comes as no surprise, as the formal eikonal long-wavelength approximation is conjugate to the linear heat equation via the Hopf–Cole transformation.

We briefly summarize the blow-up construction from [18]. To leading order, the system (2.5) is homogeneous, quadratic, with variables $(\kappa, \alpha, \delta) \in \mathbb{R}^3$, once we set $\omega - \omega_* = \pm d\delta^2$. We introduce polar coordinates $\mathbb{R}_+ \times S^2 \mapsto \mathbb{R}^3$, thus blowing up the origin into a 2-sphere, and then introduce local coordinates corresponding to stereographic projections. More explicitly, these coordinates are

$$\begin{aligned} \alpha_1 &= \alpha, \quad \kappa_1 = \frac{\kappa}{\alpha}, \quad \delta_1 = \frac{\delta}{\alpha}; \\ \alpha_2 &= \frac{\alpha}{\delta}, \quad \kappa_2 = \frac{\kappa}{\delta}, \quad \delta_2 = \delta. \end{aligned}$$

In the new coordinates, after rescaling time with the Euler multipliers α_1 and δ_2, respectively, the equations read

$$\begin{aligned} \kappa_1' &= -(n-2)\kappa_1 + \frac{\Omega''}{2d}\kappa_1^2 \pm \delta_1^2 + \mathrm{o}_{\alpha_1}(1), \\ \alpha_1' &= -\alpha_1, \\ \delta_1' &= \delta_1, \end{aligned}$$

and

$$\begin{aligned} \kappa_2' &= -(n-1)\alpha_2\kappa_2 + \frac{\Omega''}{2d}\kappa_2^2 \pm 1 + \mathrm{o}_{\delta_2}(1), \\ \alpha_2' &= -\alpha_2^2, \\ \delta_2' &= 0. \end{aligned}$$

In the following, we will assume $\Omega'' > 0$ and choose the "-"-sign in the two equations. The other case is completely analogous. We depict the phase portrait in this rescaled time in Figure 4. Note that, with the convention $\Omega'' > 0$, the equilibria with $\kappa > 0$ correspond to asymptotically positive outward pointing group velocity, and connections to those equilibria are the sources we are seeking.

Step 7: The Dulac map. The stable manifold of the family of saddles in the 2-chart can be continued into the 1-chart until it enters the section $\alpha = 1/R$. Its location in this section can be computed as follows. The stable manifold enters a neighborhood of the singular equilibrium $\kappa_1 = \delta_1 = \alpha_1 = 0$ along the weak unstable (or center) direction from the negative κ direction. It will leave the neighborhood of this equilibrium along the stable α-direction. We start by computing the transition map for the approximation

$$\begin{aligned}
&\kappa_1' = \frac{\Omega''}{2d}\kappa_1^2, \quad \alpha_1' = -\alpha, \quad \delta_1' = \delta_1, \ n = 2, \quad \text{or}\\
&\kappa_1' = -(n-2)\kappa_1, \quad \alpha_1' = -\alpha, \quad \delta_1' = \delta_1, \ n < 2.
\end{aligned}$$

In the section $\kappa_1 = m$, $m > 0$ small, the distance of the stable manifold from the singular sphere can be expanded as $\alpha_{\text{in}} = c_\alpha \delta + \mathrm{O}(\delta^2)$. We want to compute the location of this manifold after passage near the singular equilibrium, when it hits the section $\alpha_1 = m$. Since the flow in the direction of α_1 is linear, we find the time of flight as $T = \log \delta + \mathrm{O}(1)$, with error terms smooth in δ. In order to compute the κ_1-coordinate after time T, we need to compute the flow Φ_T in the κ-direction, with initial condition $\kappa_1 = m$. We find

$$\begin{aligned}
\Phi_T(m) &= \mathrm{e}^{-(n-2)T} m \qquad \text{for } n < 2,\\
\Phi_T(m) &= \frac{1}{\frac{1}{m} - \frac{\Omega''}{2d}T} \qquad \text{for } n = 2,
\end{aligned}$$

which gives

$$\begin{aligned}
\kappa_1 = \Phi_T(m) &= \delta^{2-n} m + \mathrm{O}(\delta^{2(2-n)}) \qquad \text{for } n < 2,\\
\kappa_1 = \Phi_T(m) &= -\frac{2d}{\Omega'' \log \delta} + \mathrm{O}\left(\frac{1}{(\log \delta)^2}\right) \qquad \text{for } n = 2.
\end{aligned} \tag{2.6}$$

In order to estimate the influence of the error terms, we set $\kappa_1 = \kappa_1^* + \hat{\kappa}_1$, where κ_1^* is the above approximation, with the exact initial conditions $\hat{\kappa}_1(t = T) = 0$. A straightforward fixed point argument then shows that

$$\hat{\kappa}_1(t) = \mathrm{O}(t^{-2}) \quad \text{for } n = 2, \qquad \hat{\kappa}_1(t) = \mathrm{O}(\mathrm{e}^{-(2-n+\beta)|t|}) \quad \text{for } n < 2.$$

This shows that the expansions (2.6) are valid for the full system incorporating the error terms, as well.

Step 8: The Melnikov integral. In order to conclude the construction, we have to match the stable manifold of the asymptotic wave train with the shooting manifold W_-^{cu}. At $\varepsilon = 0$, W_-^{cu} intersects W_+^{cs} transversely along $W^{\text{c}} \subset W_+^{\text{c}}$. It is therefore sufficient to compute the location of $W_-^{\text{cu}} \cap W_+^{\text{c}}$ and then propagate the solution to $\alpha = 0$. To leading order in δ, the location of

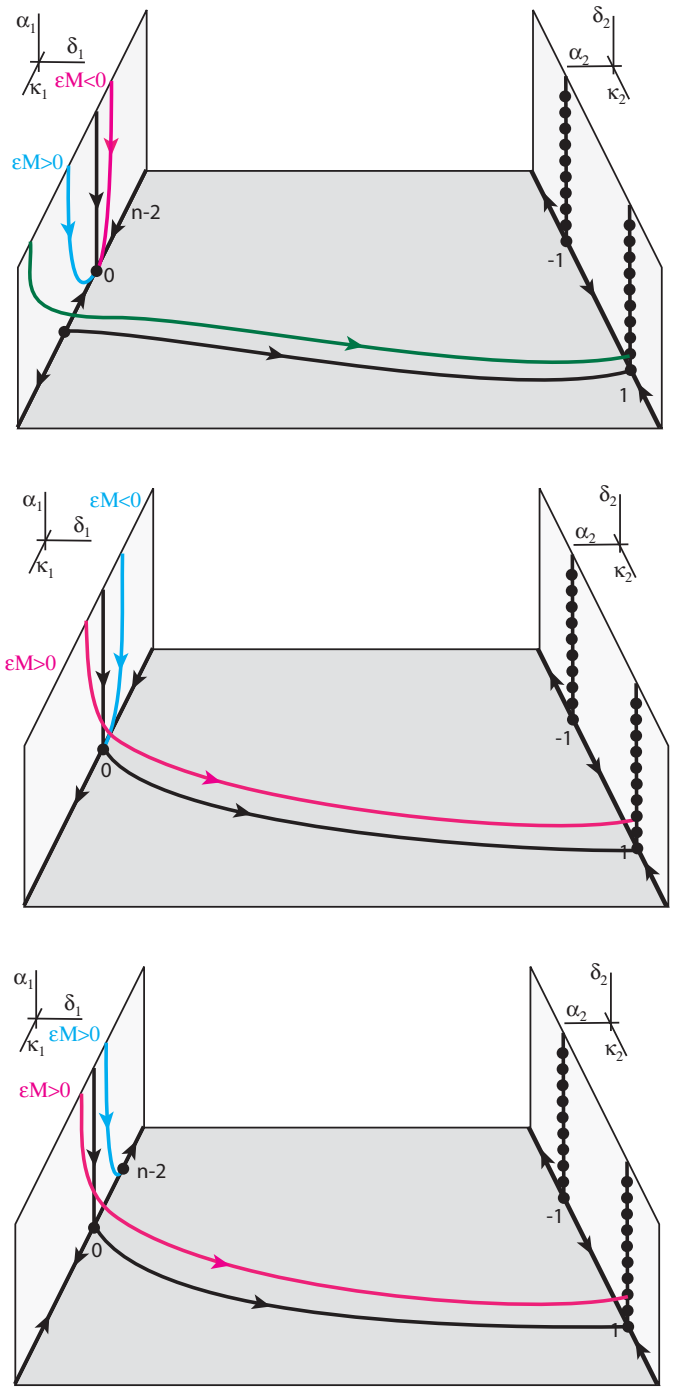


Figure 4. *The connecting orbits between $r = 0$ and the 1-chart and from the 1-chart to the 2-chart, for $n > 2$ (top), $n = 2$ (middle), and $n < 2$ (bottom). The entering trajectory in $\delta_1 = 0$, $\alpha_1 > 0$ is obtained from matching with the region $r \leq R$. We set $\Omega''/2d = 1$ in all cases.*

W^{c} remains unchanged, since δ enters the equation only at quadratic order. To leading order in ε, the location can be computed by solving the variational problem following the standard procedure in heteroclinic bifurcations as follows. We denote by $\Phi(\rho,\sigma)$ the evolution operator to the linearized equation

$$\begin{aligned} u_\rho &= \mathrm{e}^\rho v, \\ v_\rho &= -(n-1)v - \mathrm{e}^\rho D^{-1} L_0 u, \end{aligned} \tag{2.7}$$

which can be readily constructed using the exponential dichotomies at $r=\infty$ and $r\sim 0$. Similarly, we can define the adjoint operator $\Phi^*(\rho,\sigma)$ as the solution to

$$\begin{aligned} u_\rho &= \mathrm{e}^\rho L_0^* D^{-1} v, \\ v_\rho &= -\mathrm{e}^\rho u + (n-1)v. \end{aligned} \tag{2.8}$$

We are interested in the location $\partial_\varepsilon \kappa_-(R)$, the derivative of the manifold W^{c} in the direction of e_1, which is found by taking the scalar product of the perturbation, integrated along the linearized flow, with e_1^*,

$$\begin{aligned} \partial_\varepsilon \kappa_-(R) &= \left\langle e_1^*, \int_{-\infty}^{\mathrm{e}^R} \Phi(\mathrm{e}^R,\sigma)\mathrm{e}^\sigma (0, -D^{-1}\partial_\varepsilon g(\mathrm{e}^\sigma, u_*;0))\mathrm{d}\sigma \right\rangle \\ &= -\int_{-\infty}^{\mathrm{e}^R} \langle \Phi^*(\mathrm{e}^R,\sigma)e_1^*, (0, D^{-1}\partial_\varepsilon g(\mathrm{e}^\sigma,u_*;0))\mathrm{e}^\sigma\rangle \mathrm{d}\sigma. \end{aligned}$$

An explicit computation shows that

$$\Phi^*(\mathrm{e}^R,\sigma)e_1^* = (0, \mathrm{e}^{(n-1)\sigma} D u_{\mathrm{ad}}),$$

which then gives

$$\begin{aligned} \partial_\varepsilon \kappa_-(R) &= -\int_{-\infty}^{\mathrm{e}^R} \langle D u_{\mathrm{ad}}, D^{-1}\partial_\varepsilon g\rangle \mathrm{e}^{n\sigma}\mathrm{d}\sigma \\ &= -\int_0^R \langle u_{\mathrm{ad}}, \partial_\varepsilon g\rangle r^{n-1}\mathrm{d}r \\ &= -(\mathrm{Vol}\,(S^{n-1}))^{-1} \int_{\mathbb{R}^n}\int_0^{2\pi} (u_{\mathrm{ad}}(\tau), \partial_\varepsilon g(|x|, u_*(\tau);0))\mathrm{d}\tau \mathrm{d}x + \mathrm{o}_{1/R}(1). \end{aligned}$$

Step 9: Matching core and far field. We first consider the case of space dimension $n>2$. At $\delta=0$, the heteroclinic given by $u_*(\tau)$ connects to a sink in the singular blow-up chart. The connection is therefore robust under ε-perturbations, yielding a robust family of contact defects, asymptotic wavenumber zero. Unfolding in δ, a simple transitivity lemma shows that orbits pass near the equilibrium $\kappa_1=0$ in the singular chart and then connect to the sink in the 2-chart, corresponding to sinks with asymptotically negative group velocity in the reaction-diffusion system. We will derive expansions for the wavenumber below, which will then prove Theorem 1.1 in the case $n>2$.

In the case $n \leq 2$, the heteroclinic connects to a saddle and the above Melnikov analysis shows that turning on ε actually breaks the connection. For $\partial_\varepsilon \kappa_-(R) > 0$, we find a connection to the sink $\kappa_1 = 0$ in the 1-chart, thus yielding contact defects. Again, nonzero values of δ yield connections to the sink in the 2-chart, thereby accompanying sinks in the full reaction-diffusion system. In the case $\partial_\varepsilon \kappa_-(R) < 0$, we find an intersection with the stable manifold of the saddle in the 2-chart, whose location at $\alpha = 1/R$ is $\kappa = M\varepsilon + \mathrm{O}(\varepsilon^2)$. We match this expansion with the expansion (2.6) from the singular chart and solve for δ^{2-n} or $\mathrm{e}^{-1/\delta}$, respectively, with the implicit function theorem. We find the expansion

$$\begin{aligned} &\delta = c_* |M\varepsilon|^{1/(2-n)} (1 + \mathrm{o}_\varepsilon(1)) \quad \text{for } n < 2, \\ &\delta = c_* \mathrm{e}^{-\frac{2d}{\Omega'' M\varepsilon}} (1 + \mathrm{o}_\varepsilon(1)) \quad \text{for } n = 2, \end{aligned} \tag{2.9}$$

for some constants c_* with $c_* \Omega'' > 0$. Next, we replace δ by k, using the relations $\omega - \omega_* = \delta^2$ and $\omega - \omega_* = \Omega'' k^2/2 + \mathrm{O}(k^4)$, so that

$$\begin{aligned} &k = \hat{c} (M\varepsilon)^{1/(2-n)} (1 + \mathrm{o}_\varepsilon(1)) \quad \text{for } n < 2, \\ &k = \hat{c} \mathrm{e}^{-\frac{2d}{\Omega'' M\varepsilon}} (1 + \mathrm{o}_\varepsilon(1)) \quad \text{for } n = 2. \end{aligned} \tag{2.10}$$

This proves the existence of sources in the case $n \leq 2$ and the expansion for the asymptotic wavenumber. Again, there are also accompanying sinks with asymptotic wavenumbers $|k| > |k_{\mathrm{source}}(\varepsilon)|$. This proves existence and asymptotics for sources in Theorem 1.1 in the case $n \leq 2$.

The existence and asymptotics for contact defects, $\kappa = 0$, are simpler. The connection is robust in the 1-chart, with asymptotics

$$\kappa_1 \to \frac{2d(2-n)}{\Omega''}, \quad n < 2, \quad \kappa_1 \sim \frac{1}{\log r + (\varepsilon M)^{-1}}, \quad n = 2, \quad \text{and} \quad \kappa_1 \sim r^{2-n}, \quad n > 2,$$

which leads to wavenumber asymptotics

$$k \sim \frac{2d(2-n)}{\Omega'' r}, \quad n < 2, \quad k \sim \frac{\varepsilon M}{\varepsilon M r \log r + r}, \quad n = 2, \quad \text{and} \quad k \sim M\varepsilon r^{1-n}, \quad n > 2.$$

Remark 2.5. The analysis presented here simplifies considerably when studying systems with a gauge symmetry, such as $\lambda - \omega$ systems or Ginzburg–Landau equations. The gauge symmetry in these systems is represented as an action T of the circle group $\phi \in S^1$, so that u is a solution if and only if $T_\phi u$ is. The simplest periodic solutions then are equilibria with respect to this action, $u(t,x) = T_{kx-\omega(k)t} u_*(k)$. Coherent structures can be found as solutions of the form $u(t,x) = T_{-\omega t} u_*(x)$, so that u_* satisfies a (time-independent) elliptic equation with a free parameter ω. In the radially symmetric case, this reduces the problem to an ordinary differential equation; see [10] for an analysis of such a problem. For nonradially symmetric inhomogeneities, the resulting time-independent PDE problem is similar to the radially symmetric case, without gauge symmetry, that we analyzed here.

3. The eikonal approximation. In this section, we derive the viscous eikonal equation formally from the reaction-diffusion system and analyze profiles of coherent structures in this approximation. We find results similar to the results in Theorem 1.1, and we are able to discuss nonradially symmetric inhomogeneities, and the effect of moving inhomogeneities. The results differ from the results in the previous section since they assume a particular spatial scaling of the inhomogeneity for small amplitude ε. This special scaling is consistent with the modulation ansatz, albeit not typical in a given system. We also emphasize that this formal approach yields flat $\mathrm{e}^{-1/\varepsilon}$-expansions, while the approximation itself is correct only to order ε^2. It is therefore not clear at all why these formal results give asymptotics similar to the ones that we rigorously obtained in the previous section.

3.1. Stationary nonsymmetric inhomogeneities. We consider

$$u_t = D\triangle u + f(u) + \varepsilon^2 g(\varepsilon x, u), \tag{3.1}$$

with ε small. We follow the derivation in [5, section 4.3], where the computations are shown for a wave train with nonzero spatial wave number. We substitute the ansatz (1.5),

$$u(t,x) = u_*(\Phi(T,X) - \omega t; \varepsilon\nabla_X\Phi(T,X)) + \varepsilon^2 u_1(-\omega t, T, X), \quad X = \varepsilon x,\ T = \varepsilon^2 t, \tag{3.2}$$

into (1.7) and expand in powers of ε. At order ε^2, we find after a short computation

$$\Phi_T u_*' - \Delta_X\Phi D u_*' - |\nabla_X\Phi|^2 D u_*'' - g(X,u_*) = L_0 u_1(x). \tag{3.3}$$

Solvability requires that the left-hand side of (3.3) belong to the range of L_0. We denote by u_{ad} the kernel of the L^2-adjoint of L_0 with normalization $(u_{\mathrm{ad}}, u_*') = 1$. After some calculations, this solvability condition turns out to be equivalent to

$$\Phi_T = d\Delta_X\Phi - \frac{1}{2}\Omega''|\nabla_X\Phi|^2 + \bar{G}(X), \tag{3.4}$$

where

$$\bar{G}(X) = \int_0^{2\pi} (g(X, u_*(\zeta)), u_{\mathrm{ad}}(\zeta))\,\mathrm{d}\zeta;$$

see also [5] for a similar expansion. The Hopf–Cole transformation

$$A = \mathrm{e}^{a\Phi}, \qquad a = -\frac{2d}{\Omega''},$$

linearizes the eikonal equation so that we find

$$A_T = d\Delta A + \bar{G}(X)A. \tag{3.5}$$

Coherent structures are solutions with $\Phi_T = -\omega$, which is equivalent to $A_T = -\omega A$. They therefore correspond to eigenfunctions of the Schrödinger eigenvalue problem

$$-\omega A = d\Delta A + \bar{G}(X)A. \tag{3.6}$$

For $n = 1, 2$, expansions on the leading eigenvalue have been derived in many contexts, going back to Landau [12] and Simon [21]. For instance, for $d = 1$, [21, Theorem 3.4] states that

$$\omega \sim \mathrm{e}^{-\frac{4\pi}{\int G}} \tag{3.7}$$

for small $\int G$, which agrees with our expansion in Theorem 1.1. To see this, replace $\omega \sim \Omega'' k^2$ in (3.7), solve for k, and recall that $M = \int G/\mathrm{Vol}(S^1) = \int G/2\pi$.

From a phenomenological viewpoint, we are most interested in the asymptotics of the wavenumber in the far field, that is, in $\nabla\Phi(X)$ as $X \to \infty$. In the Schrödinger formulation, this amounts to computing decay and growth properties of the eigenfunction in the far field. For instance, assume that an eigenfunction A possesses exponential asymptotics $\mathrm{e}^{-\gamma|x|}$ for some $\gamma \in \mathbb{R}$. We can then infer phase asymptotics for $\Phi = \frac{1}{a}\log A$,

$$\Phi \sim \frac{1}{2d}\Omega''\gamma|x|,$$

and hence $\operatorname{sign} k = \operatorname{sign} \Omega''\gamma$. In particular, exponential decay generates outward group velocities, and exponential growth inward group velocities. The exponential localization of eigenfunctions outside of the essential spectrum yields sources in the eikonal equation.

For $n > 2$ the nonexistence of eigenvalues (point spectrum) for small potential follows readily from Hardy's inequality [13], which states that for any f with f' integrable,

$$\int_0^\infty \left|\frac{f(x)}{x}\right|^p \mathrm{d}x < \left(\frac{p}{p-1}\right)^p \int_0^\infty |f'(x)|^p \mathrm{d}x.$$

Indeed, for $n = 3$, say, and $\bar{G}$ radial, the symmetric form associated with the elliptic operator is positive whenever $\bar{G} < \frac{d}{4r^2}$,

$$\int_{\mathbb{R}^3} (d|\nabla A|^2 - \bar{G}(x)A^2)|x|^2 \mathrm{d}x \geq 4\pi d \int_0^\infty \left(r^2 A_r^2 - \frac{1}{4}A^2\right) \mathrm{d}r,$$

and by Hardy's inequality

$$\int_0^\infty r^2 A_r^2 \mathrm{d}r = \int_0^\infty \left((rA)_r^2 - A^2 - (A^2)'r\right) \mathrm{d}r = \int_0^\infty (rA)_r^2 \mathrm{d}r \geq \int_0^\infty \frac{1}{4}A^2 \mathrm{d}r. \tag{3.8}$$

Note, however, that in the scaling used to derive the eikonal approximation, $\bar{G}$ need not be small. In particular, potentials which are small in amplitude but long-range in the sense imposed by the scaling can create sources.

The following result appears to be standard, but we were unable to locate a good reference and therefore include a proof in the appendix.

Proposition 3.1. *Consider the eigenvalue problem* (3.6) *in* $\mathbb{R}^2$ *with* $d = 1$ *and smooth potential* $\bar{G}$ *with*

$$|\bar{G}(r,\varphi))| + |\partial_r \bar{G}(r,\varphi)| + |\partial_\varphi \bar{G}(r,\varphi))| = \mathrm{O}(r^{-1-\beta})$$

in polar coordinates, as $r = |x| \to \infty$*, for some* $\beta > 0$*. Let* $u \in C^2$ *be a positive eigenfunction to an eigenvalue* $\omega = -\gamma^2 < 0$*.*

We then have the following renormalized asymptotics with $\tilde{A}(r,\varphi) = \sqrt{r}\mathrm{e}^{\gamma r}A(r,\varphi)$*:*

$$|\tilde{A}(r,\cdot) - \tilde{A}_\infty(\cdot)|_{H^1} \to 0$$

for some $\tilde{A}_\infty(\cdot) \in H^1_{\mathrm{per}}(0,2\pi)$*,* $\tilde{A}_\infty \neq 0$*.*

There is a tremendous amount of literature on decay properties of solutions to elliptic equations; see, for instance, the early references [1] for upper bounds and [4, Theorem 3.2] for lower bounds. The decay properties for $\bar{G}$ needed here are weaker than the decay properties in the analysis of small potentials. For instance, long-range potentials, $\bar{G} \sim |x|^{-1-\beta}$, $\beta < 1$, create infinitely many bound states, while the short-range potentials from Theorem 1.1 create a unique bound state at small amplitude. The proposition states that the long range of the potential does not influence the decay rate of the eigenfunction at leading order.

We expect that contact defects are highly nonunique in the nonradially symmetric case: we expect an infinite-dimensional manifold, corresponding to different asymptotic azimuthal profiles. Asymptotics for this situation do not appear to be well understood.

From Proposition 3.1, we can conclude that level lines of the phase are corrections to increasingly large circles. Rings sent out by the sources do not converge to circles, and the deviation will remain O(1). Indeed, in the far field, level sets of the phase $\Phi = \log A = C$ solve

$$\log \tilde{A}_\infty(\varphi) + \gamma r + \frac{1}{2}\log r = C,$$

so that with the formal inverse $h(\gamma r + \frac{1}{2}\log r) = r$, we have the asymptotic parametrization $r = h(C - \log \tilde{A}_\infty(\varphi))$ of level sets. Since $h' \sim 1/\gamma$ for large r, r will not be constant on level sets if $\tilde{A}_\infty$ is not constant.

Note, however, that the wavenumber does converge to $kx/|x|$, since azimuthal gradients decay as $1/|x|$.

3.2. Transmission, sinks, and contacts. Deviating from our primary focus on higher-dimensional coherent structures, we recall results on transmission defects in one-dimensional media from [17], preparing for the analysis of moving inhomogeneities in the next section. Our interest is in the response of the inhomogeneity to incoming wave trains. A fairly complete answer can be provided in one space dimension.

We start with the viscous eikonal equation, $d = 1$, $\Omega'' = 1$, and a localized inhomogeneity,

$$-\omega = \Phi_{xx} - \frac{1}{2}\Phi_x^2 + \varepsilon G(x), \tag{3.9}$$

which we write as a nonautonomous first-order ODE,

$$\Phi_x = k, \qquad k_x = -\omega + \frac{1}{2}k^2 - \varepsilon G(x).$$

Note that the equation for k is independent of Φ and can therefore be solved independently,

$$k_x = -\omega + \frac{1}{2}k^2 - \varepsilon G(x). \tag{3.10}$$

At $x = \pm\infty$, the inhomogeneity disappears, and we find

$$k_x = -\omega + \frac{1}{2}k^2.$$

This equation possesses equilibria $k = \pm\sqrt{2\omega}$ for $\omega > 0$ corresponding to the planar wave trains, equilibria $k = 0$ for $\omega = 0$ corresponding to the homogeneous oscillations, and no bounded solutions for $\omega < 0$. As a consequence, bounded solutions to (3.10) exist only for $\omega \geq 0$. As $x \to +\infty$, there exists a unique solution which converges to $k = \sqrt{2\omega}$. All other bounded solutions converge to $k = -\sqrt{2\omega}$.

Let $W^{\mathrm{s}}_+ = (-\infty, k^*_+)$ be the set of initial values to (3.10) at $x = 0$ which converge to $k = -\sqrt{2\omega}$ for $x \to \infty$, the slice of the stable manifold at $x = 0$. The solution with $k(x = 0) = k^*_+$ then converges to $k = \sqrt{2\omega}$. Analogously, we define $W^{\mathrm{u}}_- = (k^*_-, +\infty)$ as the set of initial values that converge to $k = \sqrt{2\omega}$ for $x \to -\infty$, so that $k(x = 0) = k^*_-$ is the unique initial value to the solution which converges to $k = -\sqrt{2\omega}$ for $x = -\infty$. We set $\gamma = \sqrt{2\omega} \geq 0$. It is not difficult to see that

$$k^*_+ = \gamma - \varepsilon \int_0^\infty G + \mathrm{O}(\varepsilon^2 + \gamma^2), \qquad k^*_- = -\gamma + \varepsilon \int_{-\infty}^0 G + \mathrm{O}(\varepsilon^2 + \gamma^2). \tag{3.11}$$

We find bounded solutions whenever $k^*_+ - k^*_- \geq 0$, that is, when

$$2\gamma \geq \varepsilon \int_{-\infty}^\infty G + \mathrm{O}(\varepsilon^2). \tag{3.12}$$

If $\varepsilon \int G < 0$, bounded solutions exist for all values of γ. If $\varepsilon \int G > 0$, bounded solutions exist only for $\gamma \geq \varepsilon \int G/2$. The phase portrait in extended phase space is as depicted in Figure 5. The construction of the bifurcation diagram can be adapted to the case where $G(x) = \mathrm{o}(1/x)$, when the decay of G exceeds the decay of k, using a variation-of-constant formula. From the heteroclinic bifurcation picture in Figure 5, one can readily infer the bifurcation diagram in Figure 2.

We close with a short phenomenological interpretation of these findings. For $\varepsilon < 0$, incoming waves with wavenumber k_∞ are transmitted across the inhomogeneity with a phase jump $-\varepsilon \int G/k$. For $\varepsilon > 0$, the incoming waves are transmitted as long as their wavenumber is large enough, so that waves emitted by the source are pushed towards the inhomogeneity. When the wavenumber of the incoming waves is less than or equal to the emitted wavenumber, an interface between the incoming waves and the waves emitted by the source is pushed away from the inhomogeneity, so that the final state is the pure source. The latter scenario is the building block of the situation in two space dimensions, which we will address next in section 3.3.

An interesting question arises when trying to understand this bifurcation from a path-following point of view. Sources, as codimension-one heteroclinic orbits, are robust and can be followed in parameters such as ε. While global path-following results for heteroclinic orbits are not available, one can still try to continue the source ad hoc through $\varepsilon = 0$. The natural continuation becomes apparent in the dual picture of eigenvalues of Schrödinger operators: the eigenvalue corresponding to the source disappears in the essential spectrum. It can, however, be continued as a zero of the analytic continuation of the *pointwise* resolvent, or the

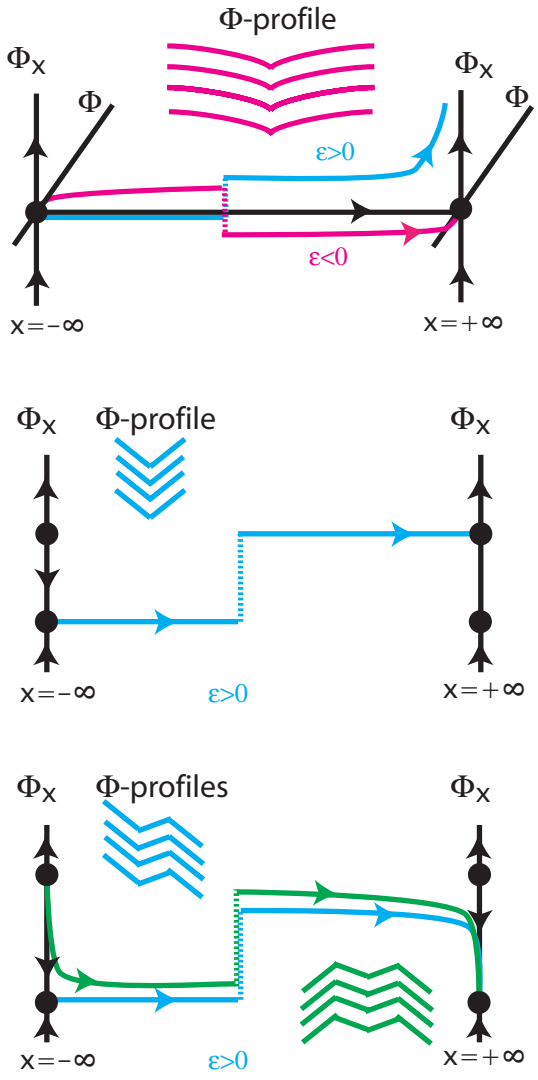


Figure 5. *Schematic picture of the existence of sources and sinks in the eikonal approximation for small potentials. Phase portraits are in the phase space of wavenumber (vertical) and physical space (horizontal and compactified). Phase is irrelevant and suppressed. In the first picture, we find contact defects; in the second, sources; and in the last picture, sinks and transmission defects, constructed from sources and sinks.*

scattering function for $\varepsilon < 0$. It then corresponds to a *resonance pole*, exhibiting pure exponential growth. In the ODE bifurcation picture for the eikonal equation, these resonance poles correspond to a heteroclinic connection between strong unstable and strong stable manifolds. The corresponding sinks separate two different kinds of sinks. For the first, regular type, with large enough incoming wavenumber, the wavenumber decreases monotonically along the sink. Sending in smaller wavenumbers, the profile changes to a nonmonotone profile, where wavenumbers (and therefore group velocities) increase in magnitude while approaching the sink, before they decrease close to the sink and change sign. The maximal wavenumber attained

by those sinks is the one selected by the resonance pole. In this sense, even though the source disappeared, the ghost of the source still selects a wavenumber in an intermediate range!

3.3. Moving sources—Conical and diffusive shock profiles. We now consider the case of a moving inhomogeneity or, equivalently, the system

$$u_t = D\triangle u + c\varepsilon u_{x_1} + f(u) + \varepsilon^2 g(\varepsilon x, u), \tag{3.13}$$

where $x = (x_1, y) \in \mathbb{R} \times \mathbb{R}^{n-1}$. With (1.5), we find the eikonal equation in a comoving frame

$$\Phi_T = d\Delta_X \Phi + c\Phi_{X_1} - \frac{1}{2}\Omega''|\nabla_X \Phi|^2 + \bar{G}(X), \tag{3.14}$$

with $X = (X_1, Y) = \varepsilon(x_1, y)$ and $T = \varepsilon^2 t$. The transformations

$$\Psi = \Phi + bx_1, \qquad b = \frac{c}{\Omega''}, \qquad \text{and} \qquad A = \mathrm{e}^{a\Psi}, \qquad a = -\frac{2d}{\Omega''}, \tag{3.15}$$

give

$$A_T = d\Delta A - \frac{c^2}{2\Omega''}A + \bar{G}(X)A. \tag{3.16}$$

Coherent structures again correspond to eigenfunctions of the self-adjoint Schrödinger eigenvalue problem

$$-\omega A = d\Delta A - \frac{c^2}{2\Omega''}A + \bar{G}(X)A. \tag{3.17}$$

If we adjust the frequency for the Doppler shift $\hat{\omega} := \omega - \frac{c^2}{2\Omega''}$, we recover the eigenvalue problem from the previous section. In particular, Proposition 3.1 gives wavenumber asymptotics

$$\nabla\Phi = \gamma\frac{x}{|x|} - be_1 + \mathrm{O}\left(\frac{1}{|x|}\right),$$

where $\gamma = \sqrt{-\hat{\omega}}$, b was defined in (3.15), and $e_1 \in \mathbb{R}^n$ denotes the unit vector in the direction of x_1. Level lines are radial where the phase gradient is perpendicular to x, that is, when $\gamma - b(e_1, x/|x|) = 0$. This gives the typical sonic cone with opening angle $\vartheta = \arccos(\gamma/b)$; see also the numerical simulations in Figure 6 below.

A more typical scenario would be a source moving with speed $c = \mathrm{O}(1)$,

$$u_t = D\triangle u + cu_{x_1} + f(u) + \varepsilon^2 g(\varepsilon^2 x_1, \varepsilon y, u). \tag{3.18}$$

We insert the ansatz

$$u(t,x) = u_*(\Phi(T,X) - \omega t; \varepsilon\nabla_X\Phi(T,X)) + \varepsilon^2 u_1(-\omega t, T, X), \quad X = (\varepsilon^2 x_1, \varepsilon y),\ T = \varepsilon^2 t, \tag{3.19}$$

and find

$$\Phi_T = d\Delta_Y \Phi + c\Phi_{X_1} - \frac{1}{2}\Omega''|\nabla_Y \Phi|^2 + \bar{G}(X_1, Y). \tag{3.20}$$

Transforming

$$A = \mathrm{e}^{a\Phi}, \qquad a = -\frac{2d}{\Omega''},$$

linearizes the equation, which results in

$$A_T = d\Delta_Y A + cA_{X_1} + \bar{G}(X_1, Y)A.$$

Coherent structures are solutions to the *heat equation*

$$-cA_{X_1} = d\Delta_Y A + \omega A + \bar{G}(X_1, Y)A.$$

Removing the exponential growth $\mathrm{e}^{-\omega X_1/c}$, induced again by a Doppler shift, and rescaling Y and X_1, we find

$$A_{X_1} = \Delta_Y A + \bar{G}(X_1, Y)A.$$

We focus on solutions with decay of the superimposed wavenumber, that is, decay of $\nabla(\log A)$, as $X_1 \to -\infty$. This is equivalent to the shooting condition in the analysis of the preceding section, where we imposed a wavenumber k at $-\infty$ and deduced wavenumber and phase jump at $+\infty$. Such solutions can be written in the form $A = 1 + B$ for a localized B, which solves the fixed point equation

$$B(X_1, \cdot) = \int_{-\infty}^{X_1} \big(T(X_1 - x_1)\bar{G}(x_1, \cdot)(1 + B(x_1, \cdot))\big)\,\mathrm{d}x_1.$$

Here, $T(\xi)$ is the heat semigroup, given as a convolution operator

$$(T(\xi)f)(y) = \frac{1}{\sqrt{4\pi\xi}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{(y-y')^2}{4\xi}} f(y')\mathrm{d}y'.$$

It is easy to see that localization of $\bar{G}$, for instance, $\bar{G} \in L^1$, implies that there exists a unique solution B such that B decays as $X_1 \to -\infty$. For $X_1 \to +\infty$, this unique solution will approach a self-similar solution to the heat equation with asymptotics

$$B(X_1, Y) = B_* \frac{1}{\sqrt{4\pi X_1}} \mathrm{e}^{-\frac{Y^2}{4X_1}} (1 + \mathrm{o}_{1/X_1}(1))$$

in $L^1 \cap L^\infty$. In particular, level lines are close to parabolas, $Y^2 - \frac{1}{2}X_1 \log X_1 = \textit{const.}$ as $X_1 \to \infty$. We note that these results should be valid only in an intermediate regime until the next-order correction in the derivation of the modulation equation becomes relevant. We therefore expect a crossover at a finite, but large, distance from a parabola to the conical sector described in section 3.1.

The simulations in Figure 6 confirm these predictions in the setting of a reaction-diffusion system with a homogeneous oscillation. In particular, we observe the conical and parabolic profile for weak and strong drift, respectively.

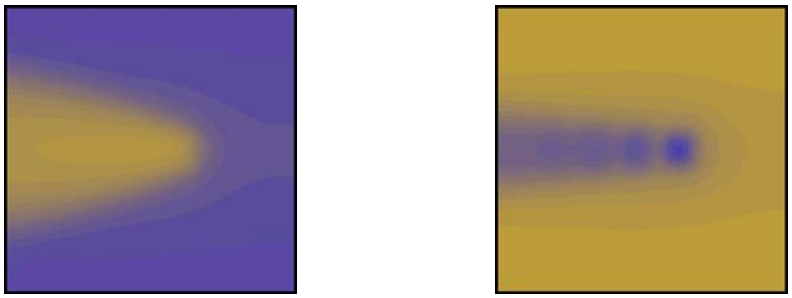


Figure 6. *Simulations of* (1.11), $\varepsilon = 0.2$, *with parameters from the introduction; the horizontal drift is* $c = 1.34$ *(left) and* $c = 2.68$ *(right). Plotted is the asymptotic state when starting with a spatially homogeneous initial condition. Clicking on the above images displays the accompanying movies* (66695_05.mpg [1.7MB] *and* 66695_06.mpg [1.5MB]) *of the time evolution.*

4. Discussion.

4.1. Summary. We proposed to study inhomogeneities in oscillatory media as an example for the creation and annihilation of coherent structures. Motivated by the one-dimensional case, we emphasized group velocities in the far field as the primary characteristic. In the examples we found we were able to compute these group velocities. In the simplest radial case, group velocities are radial, and coherent structures are either sources, sinks, or contact structures. In two (or fewer) space dimensions, small inhomogeneities can create sources. Contact defects are always *weak sinks*, and group velocities at finite but large distance point inward. In more than two space dimensions, small inhomogeneities create *weak sources or sinks*, with group velocities converging to zero in the far field. In fact, the phase jump between the center and infinity is finite, so that typically only a finite number of rings are observed in physical space. Our study of moving inhomogeneities revealed more subtle effects. Group velocities may point outward only in a sector and inward along the complement. In analogy to the one-dimensional situation, we would refer to these structures as transmission defects, since waves both enter and leave a fixed neighborhood of the defect.

4.2. Fluxes. Instead of retaining only the sign of $(c_g, x/|x|)$, that is, inward versus outward transport, we may look at the group velocity as a map from S^{n-1} to $\mathbb{R}^n$, on large centered spheres. In most of our cases, this map converges as the size of the sphere tends to infinity. There are then various ways to extract quantitative information from this map in order to characterize the coherent structures. For instance, we may define the *net flux* J associated with the phase of a coherent structure Φ as

$$J(\Phi) = \lim_{R\to\infty} (R^{n-1}\mathrm{Vol}\,(S^{n-1}))^{-1} \int_{|x|=R} j(x)\mathrm{d}\sigma \quad \text{with } j(x) = \left(c_g(x), \frac{x}{|x|}\right),$$

where σ is the $(n-1)$-dimensional Lebesgue surface element on $|x| = R \subset \mathbb{R}^n$. In addition to the distinction between sources, $J > 0$, contact defects, $J = 0$, and sinks, $J < 0$, J also retains the strength of the source, and thus $J \sim \varepsilon$ in one space dimension and $J \sim \mathrm{e}^{-1/\varepsilon}$ in two space dimensions, where we have neglected normalizing constants. This can be readily seen from our main results on the expansion for the wavenumber k in the far field. Indeed, $c_g \sim k$

since $\omega \sim k^2$, which immediately gives the expansion using Theorem 1.1 in the radial case and Proposition 3.1 in the case of the eikonal approximation; see also the remark at the end of section 3.1. As a limit, J does not retain the correction terms which distinguish between weak sources and sinks. For the transmission defects that we discussed in the case of slowly moving sources, the mean drift cancels in the integral so that we are lead to interpreting the case $\varepsilon M > 0$ as a source rather than a transmission defect: transmission in the horizontal direction is superimposed by a source term, visible in particular in the vertical direction. For inhomogeneities moving with finite speed, it is not difficult to see that the limit actually vanishes: group velocities generated by the inhomogeneity decay almost everywhere.

4.3. Degrees. A more phenomenological classification would look at constant level lines of Φ, or u, directly. We showed that level lines are expanding circles, with bounded and converging correction terms in case of anisotropic inhomogeneities. In case of moving inhomogeneities, we found cones and parabolas in the far field. Level lines need not form closed curves, as the case of spiral waves illustrates. In our case, however, it is not difficult to see that spirals cannot form. Indeed, one can define the topological degree of the defect as the degree i_{phase} of the phase as a map from a large circle into a circle [25]. Since the phase is everywhere defined in our case of weak inhomogeneities, this map extends to a map from the disc into the circle so that the degree vanishes. Note that fluxes do not distinguish between target patterns and spiral waves, while the topological degree does not distinguish between sources and sinks. Another degree i_{grad} would be the Brouwer degree of c_{g}. If $j > 0$ pointwise (source), then $i_{\text{grad}} = 1$, and if $j < 0$ pointwise, then $i_{\text{grad}} = -1$, regardless of i_{phase}. For a plane wave, $i_{\text{grad}} = 0$. Note, however, that i_{grad} need not be defined for all coherent structures, since group velocities may vanish even in the far field.

4.4. Stability. Using the methods employed to prove existence, one can also track eigenvalues of the linearization. We conjecture that all coherent structures discussed here are stable. In the eikonal approximation, this can be seen after a Cole–Hopf transformation: solutions to the heat equation with source term approach the eigenfunction with the largest eigenvalue, exponentially when there is a spectral gap. For the reaction-diffusion systems, we expect that the methods developed in section 2 should give spectral stability. We are not aware of nonlinear stability proofs for sources in reaction-diffusion system; see, however, [7] for a related result for a transmission defect.

5. Appendix. We prove Proposition 3.1. Our proof follows the ideas outlined in section 2, setting up a dynamical systems framework where asymptotics are a consequence of expansions on stable and unstable manifolds. We rewrite (3.6) as

$$A_{rr} + \frac{1}{r}A_r + \frac{1}{r^2}A_{\varphi\varphi} + G(r,\varphi)A - \gamma^2 A = 0, \quad \omega = -\gamma^2$$

in the form

$$\begin{aligned} A_r &= \mathcal{M}B, \\ B_r &= \mathcal{M}A - \frac{\gamma^2}{r}\mathcal{M}^{-2}B - \mathcal{M}^{-1}GA, \end{aligned} \tag{5.1}$$

where $\mathcal{M}$ is the unbounded self-adjoint operator

$$\mathcal{M} = \sqrt{-\frac{1}{r^2}\partial_{\varphi\varphi} + \gamma^2} \geq \gamma,$$

and $A, B \in L^2(S^1)$. We set $w_\pm = A \pm B$. Since $\mathcal{M}$ is self-adjoint and positive, and the correction terms are bounded as operators, with norm $\mathrm{O}(r^{-1})$, by the assumptions on $\bar{G}$, this equation possesses an exponential dichotomy [14]: there is $\Psi(r) : L^2 \to L^2$, $\Psi = \mathrm{O}(r^{-1})$, so that any solution that is bounded as $r \to \infty$ satisfies $w_+ = \Psi w_-$. We end up with an evolution equation for w_- which describes the asymptotics of any bounded solution:

$$w_-' = \mathcal{L}_-(r)w_-, \quad \mathcal{L}_- = -\mathcal{M} + \mathcal{N}, \text{ with } \mathcal{N} \text{ bounded}, \qquad \mathcal{N} = -\frac{\gamma^2}{2r}\mathcal{M}^{-2} + \mathrm{O}(r^{-(1+\beta)}).$$

We start analyzing the asymptotics neglecting the bounded $\mathrm{O}(r^{-(1+\beta)})$-terms contained in $\mathcal{N}$. The truncated equation is diagonal in Fourier modes $w_- = \sum_{k\in\mathbb{Z}} w_k \mathrm{e}^{\mathrm{i}k\varphi}$,

$$(w_-^k)' = \mathcal{L}_-^k w_-^k, \qquad \mathcal{L}_-^k = -\gamma - \frac{1}{2r} + \mathrm{O}(r^{-(1+\beta)}).$$

We next set $w_k := \sqrt{r}\mathrm{e}^{\gamma r} w_-^k$ and find

$$w_k' = \left(\gamma - \sqrt{\frac{k^2}{r^2} + \gamma^2}\right) w_k + \frac{k^2}{2r(k^2 + \gamma^2 r^2)} w_k + \mathrm{O}(r^{-(1+\beta)} w),$$

or more briefly,

$$w_k' = \frac{1}{r^2}\varphi\left(\frac{1}{r}; k\right) w_k + \mathrm{O}\left(r^{-(1+\beta)} w\right).$$

Error terms are coupling all w_k but are uniformly bounded as operators on ℓ^2. Since the right-hand side is $\mathrm{O}(r^{-(1+\beta)})$, we can introduce the compactified time variable $\tau = -1/(\beta r^\beta)$ and find

$$\dot{w}_k = \tau^{1-\beta}\varphi(\tau; k) w_k + \mathrm{O}(w).$$

More precisely, we have

$$\varphi(\tau; k) = \tau^{1-\beta}(\gamma - \sqrt{\gamma^2 + \tau^2 k^2}) + \tau \frac{k}{2(k^2\tau^2 + \gamma^2)} + \mathrm{O}\left(\tau^{-(1+\beta)} G\left(-\frac{1}{\tau}\right)\right).$$

Here we assume $\beta \leq 1$ without loss of generality. The flow map from a fixed time $\tau = -1/r_0$ to $\tau = 0$ is readily verified to be bounded (and actually differentiable in r_0), using the explicit representation in Fourier modes and a variation-of-constant formula. This provides us with the desired limiting profile w, and, substituting back, with asymptotics for w_-, w_+, u, and v. The gradient estimates readily follow from the same argument, carried out in the space ℓ^1_2, with $\sum_k |w_k|^2 |k|^2 < \infty$.

Acknowledgments. A. Scheel acknowledges discussions with D. Aronson and O. Zeitouni on properties of Schrödinger operators.

REFERENCES

[1] S. Agmon, *Lectures on Exponential Decay of Solutions of Second-Order Elliptic Equations: Bounds on Eigenfunctions of N-Body Schrödinger Operators*, Math. Notes 29, Princeton University Press, Princeton, NJ, 1982.

[2] D. Barkley, *EZ-SPIRAL, A Code for Simulating Spiral Waves*, available online from http://www.maths.warwick.ac.uk/~barkley/Software/ez_software.html.

[3] B. Blasius and L. Stone, *Chaos and phase synchronization in ecological systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 10 (2000), pp. 2361–2380.

[4] R. Carmona and B. Simon, *Pointwise bounds on eigenfunctions and wave packets in N-body quantum systems.* V. *Lower bounds and path integrals*, Comm. Math. Phys., 80 (1981), pp. 59–98.

[5] A. Doelman, B. Sandstede, A. Scheel, and G. Schneider, *The dynamics of modulated wave trains*, Mem. Amer. Math. Soc., to appear.

[6] F. Dumortier, *Techniques in the theory of local bifurcations: Blow-up, normal forms, nilpotent bifurcations, singular perturbations*, in Bifurcations and Periodic Orbits of Vector Fields, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. 408, D. Schlomiuk, ed., Kluwer, Dordrecht, The Netherlands, 1993, pp. 19–73.

[7] T. Gallay, G. Schneider, and H. Uecker, *Transport of information near essentially unstable localized structures*, Discrete Contin. Dyn. Syst. Ser. B, 4 (2004), pp. 349–390.

[8] G. Iooss and A. Mielke, *Bifurcating time-periodic solutions of Navier-Stokes equations in infinite cylinders*, J. Nonlinear Sci., 1 (1991), pp. 107–146.

[9] R. Kapral and K. Showalter, *Chemical Waves and Patterns*, Kluwer Academic, Dordrecht, The Netherlands, 1995.

[10] N. Kopell and L. N. Howard, *Target patterns and horseshoes from a perturbed central-force problem: Some temporally periodic solutions to reaction-diffusion equations*, Stud. Appl. Math., 64 (1981), pp. 1–56.

[11] M. Krupa and P. Szmolyan, *Extending slow manifolds near transcritical and pitchfork singularities*, Nonlinearity, 14 (2001), pp. 1473–1491.

[12] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, Course of Theoretical Physics, Vol. 3, Pergamon Press, Elmsford, NY, 1958.

[13] B. G. Pachpatte, *Mathematical Inequalities*, North–Holland Math. Library 67, Elsevier, Amsterdam, 2005.

[14] D. Peterhof, B. Sandstede, and A. Scheel, *Exponential dichotomies for solitary wave solutions of semilinear elliptic equations on infinite cylinders*, J. Differential Equations, 140 (1997), pp. 266–308.

[15] B. Sandstede and A. Scheel, *Essential instability of pulses and bifurcations to modulated travelling waves*, Proc. Roy. Soc. Edinburgh Sect. A, 129 (1999), pp. 1263–1290.

[16] B. Sandstede and A. Scheel, *On the structure of spectra of modulated travelling waves*, Math. Nachr., 232 (2001), pp. 39–93.

[17] B. Sandstede and A. Scheel, *Defects in oscillatory media: Toward a classification*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 1–68.

[18] B. Sandstede and A. Scheel, *Evans function and blow-up methods in critical eigenvalue problems*, Discrete Contin. Dyn. Syst., 10 (2004), pp. 941–964.

[19] B. Sandstede, A. Scheel, and C. Wulff, *Bifurcations and dynamics of spiral waves*, J. Nonlinear Sci., 9 (1999), pp. 439–478.

[20] A. Scheel, *Radially Symmetric Patterns of Reaction-Diffusion Systems*, Mem. Amer. Math. Soc. 165, AMS, Providence, RI, 2003.

[21] B. Simon, *The bound state of weakly coupled Schrödinger operators in one and two dimensions*, Ann. Phys., 97 (1976), pp. 279–288.

[22] M. Stich and A. S. Mikhailov, *Target patterns in two-dimensional heterogeneous oscillatory reaction-diffusion systems*, Phys. D, 215 (2006), pp. 38–45.

[23] F. Varela, J. P. Lachaux, E. Rodriguez, and J. Martinerie, *The brainweb: Phase synchronization and large-scale integration*, Nat. Rev. Neuroscience, 2 (2001), pp. 229–239.

[24] A. T. Winfree, *Electrical turbulence in three-dimensional heart muscle*, Science, 266 (1994), pp. 1003–1006.

[25] A. T. Winfree, *The Geometry of Biological Time*, Interdiscip. Appl. Math. 12, Springer-Verlag, New York, 2001.

SIAM J. APPLIED DYNAMICAL SYSTEMS
Vol. 6, No. 1, pp. 263–292

Patterns and Features of Families of Traveling Waves in Large-Scale Neuronal Networks*

William C. Troy† and Vladimir Shusterman‡

Abstract. We study traveling wave solutions of a system of integro-differential equations which describe the activity of large-scale networks of excitatory neurons on spatially extended domains. The independent variables are the activity level u of a population of excitatory neurons, which have long range connections, and a recovery variable v. We have found a critical value of the parameter β ($\beta_* > 0$) that appears in the equation for v, at which the eigenvalues of the linearization of the system around the rest state $(u, v) = (0, 0)$ change from real to complex. In contrast to previous studies which analyzed properties of traveling waves when the eigenvalues are real, we examine the range $\beta > \beta_*$, where the eigenvalues are complex. In this case we show that there is a range of parameters over which families of wave fronts and 1-pulse and more general N-pulse waves can coexist as stable solutions. In two space dimensions our numerical experiments show how single-ring and multiring waves form in response to a Gaussian-shaped stimulus. With a spatially invariant coupling function, outwardly propagating waves can be periodically produced when a ring-shaped wave receives an appropriately timed perturbation. When the coupling is inhomogeneous the periodic production of waves eventually breaks down, and a stable one-armed rotating spiral wave emerges and fills the entire domain. It is noteworthy that all of these phenomena can be initiated at any point in the medium, that they are not driven by an underlying time dependent periodic pacemaker, and that they do not depend on the persistent or periodic presence of an external input.

Key words. waves, integro-differential equation, nonlocal, excitatory

AMS subject classifications. 34B15, 34C23, 34C11

DOI. 10.1137/06066638X

1. Introduction.

Functional behavior of the central nervous system includes such diverse phenomena as information processing from different receptor zones, sleep, and the control of vital autonomic functions [16, 28, 29, 38, 52]. These processes require coordination between ensembles of cells organized into large-scale spatially extended neuronal networks. The physical laws that govern the behavior of large-scale networks are different from those for a system consisting of small numbers of cells [14, 26, 33, 53]. Considerable attention has been given to the study of traveling waves of activity in spatially extended neuronal networks. This includes both experimental [4, 5, 6, 10, 18, 23, 32, 35, 36, 43, 45, 47, 51, 55] and theoretical [1, 2, 12, 13, 17, 19, 20, 22, 24, 25, 31, 37, 38, 40, 54, 56] studies.

In this paper we investigate properties of traveling wave solutions of the excitable spatially

*Received by the editors July 30, 2006; accepted for publication (in revised form) by B. Sandstede December 20, 2006; published electronically March 30, 2007. Both authors contributed equally to this work.
http://www.siam.org/journals/siads/6-1/66638.html

†Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260 (troy@math.pitt.edu). This author was supported by NSF grant DMS0412370.

‡Cardiovascular Institute, University of Pittsburgh, Pittsburgh, PA 15260 (shustermanv@upmc.edu). This author was supported by NIH SCOR grant P50 HL52338.

extended neuronal field model

$$\begin{aligned} \frac{\partial u(x,t)}{\partial t} &= -u(x,t) - v(x,t) + \int_{-\infty}^{\infty} w(x-x')f(u(x',t)-\theta)dx', \\ \frac{\partial v(x,t)}{\partial t} &= \epsilon(\beta u(x,t) - v(x,t)). \end{aligned} \tag{1.1}$$

Systems of this form were introduced by Pinto and Ermentrout [40] to model the spread of excitation waves in slices of brain cortex in which synaptic inhibition is pharmacologically blocked [8, 10, 30, 55]. The variable u denotes the activity level of the population of excitatory neurons with long-range connections. The equation for v represents a negative feedback recovery mechanism in which "the negative feedback could represent spike frequency adaptation, synaptic depression or some other process that limits excitation of the network" [40]. The coupling function w is positive, continuous, and integrable. The firing rate function f is nonnegative and sigmoidal-shaped. The parameters ϵ and β are positive and control the rate of change of v; θ is a positive constant which denotes the threshold level for u. In order to allow for comparison of our results with those of previous studies, we follow [2, 19, 40, 46] and assume for simplicity that the coupling and firing rate functions are

$$w(x) = \frac{1}{2}e^{-|x|} \quad \text{and} \quad f(u-\theta) = H(u-\theta) \quad \forall x, u \in \mathrm{R}, \tag{1.2}$$

where H is the Heaviside function defined by

$$H(u-\theta) = \begin{cases} 1 & \forall u \geq \theta, \\ 0 & \forall u < \theta. \end{cases} \tag{1.3}$$

Pinto and Ermentrout investigated the existence of 1-pulse traveling waves in parameter regimes such that the linearization of (1.1) around the rest state $(u, v) = (0, 0)$ has real eigenvalues. Recently, Richardson, Schiff, and Gluckman [46] made use of these results in their study of the effects of electric fields on 1-pulse traveling waves in mammalian cortex. The existence and stability of solutions for more general couplings has also been studied [12, 42, 48].

In this paper we extend the results described above and analyze traveling waves in parameter regimes where the linearization of the system around the rest state has *complex* eigenvalues. In particular, when $\epsilon > 0$ the eigenvalues of the linearization change from real to complex as β passes through the critical value $\beta^* = \frac{(\epsilon-1)^2}{4\epsilon}$ from below. We focus our attention on the range $0 < \epsilon < 1$ and $\beta > \beta^*$, and find that the dynamics of (1.1) are richer than in the real eigenvalue case. Furthermore, these dynamics closely resemble electrophysiological phenomena observed in clinical and experimental studies [38]. Our specific aims and results are summarized below.

I. In one space dimension (sections 2–5) we study families of wave fronts and 1-pulse and more general N-pulse traveling waves. There is a range of parameters where these different types of waves can coexist as stable solutions. Our analysis explains why multipulse waves are expected to exist only in the complex eigenvalue regime. Because the eigenvalues are now complex, technical difficulties arise which make existence proofs more challenging than in the

real eigenvalue case. These difficulties lead to several open problems which will be stated as we proceed.

II. We extend our investigation to two space dimensions in section 6. In analogy with the one dimensional case, we study both single- and multipulse traveling waves. Classical in vivo experiments showed that such waves exist in feline cortex [4, 5, 6], and recently they have been discovered also in the intact brain of freely moving mice [18]. They have also been found in both tangential and coronal brain slice experiments [30, 55]. Our numerical experiments will show how single- and multiring waves form and propagate outward from a point of initial stimulus. When an appropriately timed stimulus breaks one of these rings, a new solution emerges which periodically emits ring-shaped waves. When the coupling is inhomogeneous we will see how the periodic production of outwardly propagating waves eventually breaks down and the solution evolves into a stable one-armed spiral wave which fills the entire domain.

III. Conclusions and directions for future research are given in section 7. The appendix (section 8) contains an analysis of asymptotic behavior of solutions in the real eigenvalue regime.

2. Traveling waves. In this section we begin our study of traveling wave solutions of (1.1). These have the form $(u, v) = (U(z), V(z))$, where $z = x + ct$, and satisfy

$$\begin{aligned} cU'(z) &= -U - V + \int_{-\infty}^{\infty} w(z - z')H(U(z') - \theta)dz', \\ cV'(z) &= \epsilon(\beta U - V), \end{aligned} \tag{2.1}$$

where $w(z - z') = \frac{1}{2}e^{-|z-z'|}$. It is easily verified that (2.1) is equivalent to

$$c^2U'' + c(1 + \epsilon)U' + \epsilon(\beta + 1)U = c\frac{d}{dz}\int_{-\infty}^{\infty} w(z - z')H(U - \theta)dz' + \epsilon\int_{-\infty}^{\infty} w(z - z')H(U - \theta)dz'. \tag{2.2}$$

Linearizing (2.2) around the rest state $U = 0$, we obtain

$$c^2U'' + c(1 + \epsilon)U' + \epsilon(\beta + 1)U = 0. \tag{2.3}$$

Following [42], we assume without loss of generality that $c \geq 0$. When $c > 0$ the eigenvalues associated with (2.3) are

$$\mu^{\pm} = \frac{\lambda^{\pm}}{c} = \frac{-(\epsilon + 1) \pm \mathrm{i}\sqrt{4\beta\epsilon - (\epsilon - 1)^2}}{2c}. \tag{2.4}$$

We restrict our attention to the regime $0 < \epsilon < 1$. From this and (2.4) it follows that

$$\mu^{\pm} \text{ are real} \iff 0 < \beta \leq \beta_* = \frac{(\epsilon - 1)^2}{4\epsilon}. \tag{2.5}$$

For real eigenvalues, properties of traveling waves were examined in [2, 19, 20, 40, 42, 46]. Here we extend these results and study traveling waves when $\mu^{\pm}$ are complex. When $c > 0$ we will make use of the quantities

$$\alpha = \mathrm{Re}(\mu^{\pm}) = \frac{-(\epsilon + 1)}{2c} < 0 \quad \text{and} \quad \gamma = \mathrm{Im}(\mu^{\pm}) = \frac{\sqrt{4\epsilon\beta - (\epsilon - 1)^2}}{2c} > 0. \tag{2.6}$$

When $\mu^{\pm}$ are complex, underlying oscillatory terms lead to technical difficulties which make rigorous existence proofs especially challenging. These difficulties suggest several open problems which will be discussed as we proceed.

Our investigation indicates that stable traveling waves exist when

$$0 < \epsilon < 1, \quad 0 < \theta < \min\left(\frac{2\epsilon}{(\epsilon+1)^2}, \frac{1}{4(\epsilon+1)}\right), \quad \beta > \beta_* = \frac{(\epsilon-1)^2}{4\epsilon}. \tag{2.7}$$

The first two sets of inequalities in (2.7) are mild restrictions which allow technical arguments to be completed, and the third inequality means that $\mu^{\pm}$ are complex. Throughout the paper we perform numerical experiments for parameters which satisfy (2.7). An important implication of (2.7) is that

$$\beta_* = \frac{(\epsilon-1)^2}{4\epsilon} < \frac{1}{2\theta} - 1. \tag{2.8}$$

For example, in section 3 we will see that a branch of wave fronts comes into existence at the critical value $\beta = \frac{1}{2\theta} - 1$. In our analysis of 1-pulse traveling waves in section 4 it is shown that infinitely many wave speeds are possible at this same value of β. More general N-pulses are described in section 5.

3. Wave fronts. In this section we analyze properties of wave front solutions when the eigenvalues $\mu^{\pm}$ are complex. These solutions satisfy

$$c^2U'' + c(1+\epsilon)U' + \epsilon(\beta+1)U = c\frac{d}{dz}\int_0^\infty w(z-z')dz' + \epsilon\int_0^\infty w(z-z')dz', \tag{3.1}$$

where $c > 0$, and

$$\begin{cases} U(z) < \theta \quad \forall z < 0, \\ (U(z), U'(z)) \to (0,0) \quad \text{as } z \to -\infty, \\ U(0) = \theta, \quad U(z) > \theta \quad \forall z > 0, \\ U'(z) \to 0 \quad \text{as } z \to \infty. \end{cases} \tag{3.2}$$

Thus far we do not have a complete proof of existence of wave fronts when $\mu^{\pm}$ are complex. However, a combination of analysis and numerical experiments suggests that two branches of solutions exist. To understand how these results are obtained we devote the remainder of this section to the following:

A. Analysis of solutions on $(-\infty, 0]$.
B. Analysis of solutions on $(0, \infty)$.
C. Numerical evidence for the existence of two branches of solutions.
D. Open problems.

A. Analysis of solutions on $(-\infty, 0]$. To begin we recall that $w(x) = \frac{1}{2}e^{-|x|}$. Then (3.1) becomes

$$c^2U'' + c(1+\epsilon)U' + \epsilon(\beta+1)U = g(z), \tag{3.3}$$

where

$$g(z) = \begin{cases} .5(c+\epsilon)e^z & \forall z \le 0, \\ .5(c-\epsilon)e^{-z} + \epsilon & \text{if } z > 0. \end{cases} \tag{3.4}$$

On $(-\infty, 0]$ the general solution of (3.3) is

$$U_0(z) = b_1 e^{\alpha z}\cos(\gamma z) + b_2 e^{\alpha z}\sin(\gamma z) + P_0(z), \tag{3.5}$$

where b_1 and b_2 are constants, and $P_0(z)$ is the particular solution

$$P_0(z) = \frac{.5(\epsilon + c)}{c^2 + (1+\epsilon)c + \epsilon(\beta+1)} e^z \quad \forall z \le 0. \tag{3.6}$$

The oscillatory terms $b_1 e^{\alpha z}\cos(\gamma z)$ and $b_2 e^{\alpha z}\sin(\gamma z)$ in (3.5) are due to $\mu^{\pm}$ being complex. Recall from (2.6) that $\alpha = \mathrm{Re}(\mu^{\pm}) < 0$. Thus, to satisfy the condition $U_0(-\infty) = U_0'(-\infty) = 0$, we conclude that $b_1 = b_2 = 0$, and (3.5)–(3.6) reduce to

$$U_0(z) = \frac{.5(\epsilon + c)}{c^2 + (1+\epsilon)c + \epsilon(\beta+1)} e^z \quad \forall z \le 0. \tag{3.7}$$

Substituting the condition $U_0(0) = \theta$ into (3.7) gives the algebraic equation

$$\frac{.5(\epsilon + c)}{c^2 + (1+\epsilon)c + \epsilon(\beta+1)} = \theta. \tag{3.8}$$

It follows from (3.7)–(3.8) that

$$U_0(z) = \theta e^z \quad \forall z \le 0. \tag{3.9}$$

Note that $(U_0(-\infty), U_0'(-\infty)) = (0,0)$, as required by (3.2). Furthermore, $U_0(z)$ has no oscillatory component and is strictly increasing on $(-\infty, 0]$ (Figure 1, left panel).

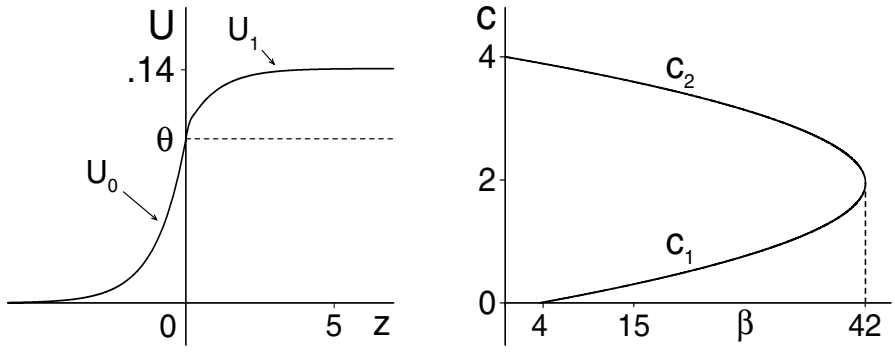


Figure 1. *Left panel: U_0 and U_1 when $\epsilon = \theta = .1$, $\beta = 4$, and $c = c_1$. Right panel: Wave speeds c_1 and c_2 versus β when $(\epsilon, \theta) = (.1, .1)$ (see (3.12)). Note that the positive branch of c_1 begins at the critical value $\beta = \frac{1}{2\theta} - 1 = 4$. The two branches meet at $\beta \approx 42$ where the functions c_1 and c_2 become complex.*

B. Analysis of solutions on $(0,\infty)$. When $z>0$ the first step in analyzing solutions is to solve (3.8) for wave speed c. This gives the two values (Figure 1, right panel)

$$c_1=\frac{.5-\theta(\epsilon+1)-\sqrt{(.5-\theta(1+\epsilon))^2-4\epsilon\theta\left(\theta(\beta+1)-.5\right)}}{2\theta}, \tag{3.10}$$

$$c_2=\frac{.5-\theta(\epsilon+1)+\sqrt{(.5-\theta(1+\epsilon))^2-4\epsilon\theta\left(\theta(\beta+1)-.5\right)}}{2\theta}. \tag{3.11}$$

Recall from (2.7) that $0<\epsilon<1$ and $0<\theta<\frac{1}{4(\epsilon+1)}$. This and (3.10)–(3.11) imply that

$$\begin{cases} c_1<0 \text{ and } c_2>0 \text{ if } \beta_*<\beta<\frac{1}{2\theta}-1,\\ c_1=0 \text{ and } c_2>0 \text{ if } \beta=\frac{1}{2\theta}-1,\\ 0<c_1\le c_2 \text{ if } \frac{1}{2\theta}-1<\beta\le\frac{1}{2\theta}-1+\frac{1}{4\epsilon\theta^2}\left(.5-\theta(1+\epsilon)\right)^2,\\ c_1 \text{ and } c_2 \text{ are complex if } \beta>\frac{1}{2\theta}-1+\frac{1}{4\epsilon\theta^2}\left(.5-\theta(1+\epsilon)\right)^2. \end{cases} \tag{3.12}$$

When $z>0$ the solution of (3.3) is

$$U_1=k_1e^{\alpha z}\cos(\gamma z)+k_2e^{\alpha z}\sin(\gamma z)+P_1(z), \tag{3.13}$$

where $\alpha<0$ and $\gamma>0$ are defined in (2.6), and $P_1(z)$ is the particular solution

$$P_1(z)=\frac{.5(c-\epsilon)}{c^2-(1+\epsilon)c+\epsilon(\beta+1)}e^{-z}+\frac{1}{\beta+1}. \tag{3.14}$$

To preserve continuity at $z=0$ we require that $(U_1(0),U_1'(0))=(U_0(0),U_0'(0))$. This and (3.13)–(3.14) show that k_1 and k_2 are uniquely defined by

$$k_1=\theta-P_1(0)\quad\text{and}\quad k_2=\frac{1}{\gamma}\left(\theta(1-\alpha)-P_1{}'(0)+\alpha P_1(0)\right). \tag{3.15}$$

To complete the proof that a solution satisfies all of the conditions in (3.2) for a traveling wave front we need to show that $U_1(z)$ (Figure 1, left panel) satisfies

$$U_1(z)>\theta\ \ \forall z>0\quad\text{and}\quad\lim_{z\to\infty}U_1'(z)=0. \tag{3.16}$$

C. Numerical evidence for the existence of two branches of solutions. Our numerical experiments suggest that conditions (3.16) are satisfied along two branches of solutions (see Figure 2). The lower branch Γ_0^- corresponds to $c=c_1$, and the upper branch Γ_0^+ corresponds to $c=c_2$. Below we describe existence and stability properties of solutions along each branch.

The lower branch. When $c=c_1$ it follows from (3.13), (3.14), and (3.15) that

$$U_1(z)=k_1e^{\alpha z}\cos(\gamma z)+k_2e^{\alpha z}\sin(\gamma z)+\frac{.5(c_1-\epsilon)}{(c_1)^2-(1+\epsilon)c_1+\epsilon(\beta+1)}e^{-z}+\frac{1}{\beta+1}\qquad\forall z>0. \tag{3.17}$$

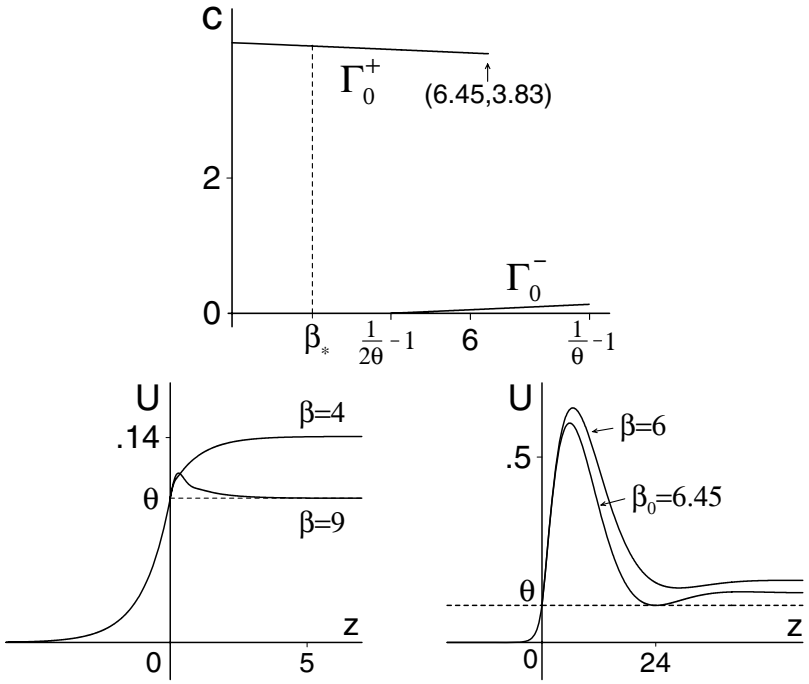


Figure 2. *Upper panel:* Γ_0^- *and* Γ_0^+ *when* $(\epsilon, \theta) = (.1, .1)$. *The horizontal axis gives* β *values, and the vertical axis represents the wave speed* c. *The eigenvalues* $\mu^{\pm}$ *become complex at* $\beta_* = 2.025$. *Lower left: solutions on* Γ_0^- *at its endpoints* $\beta = \frac{1}{2\theta} - 1 = 4$ *and* $\beta = \frac{1}{\theta} - 1 = 9$. *The solution at* $\beta = 4$ *is the stationary solution defined in* (3.21). *Lower right: solutions on* Γ_0^+ *at* $\beta = 6$ *and* $\beta = 6.45$ *The solution at* $\beta = 6$ *is a wave front solution since it remains above* θ *on* $(0, \infty)$. *The solution at* $\beta = 6.45$ *is not a wave front since it is tangent to* $U = \theta$ *at* $z \approx 24$. *When* $\beta > 6.45$ *the functions dip below* θ *and therefore cannot be wave fronts. Thus,* Γ_0^+ *ends at* $\beta = 6.45$. *Clicking on the top image displays the accompanying movie* (66638_01.mpg [1.7MB]) *showing wave formation when* $\beta = 6.45$.

If $U_1(z) > \theta \; \forall z > 0$, then (3.17) implies that $U_1'(z) \to 0$ as $z \to \infty$, and therefore (3.16) holds. Thus, it is sufficient to show that

$$U_1(z) > \theta \quad \forall z > 0. \tag{3.18}$$

However, it is difficult to prove (3.18) since the oscillatory component $k_1 e^{\alpha z} \cos(\gamma z) + k_2 e^{\alpha z} \sin(\gamma z)$ of (3.17) can cause $U_1(z)$ to dip below the threshold level θ at some point in $(0, \infty)$. Our numerical experiments indicate that there is an entire branch Γ_0^- (Figure 2) of traveling wave front solutions satisfying

$$U(z) = \begin{cases} \theta e^z < \theta & \forall z < 0, \\ U_1(z) > \theta & \forall z > 0. \end{cases} \tag{3.19}$$

Along Γ_0^- it follows from (3.10) and (3.12) that

$$c_1 \to 0^+ \quad \text{as } \beta \to \left(\frac{1}{2\theta} - 1\right)^+. \tag{3.20}$$

Thus, the lower end of Γ_0^- begins at $\beta = \frac{1}{2\theta} - 1$, where $c_1 = 0$ (Figure 2, upper panel). At this point the independent variable $z = x + ct$ reduces to $z = x$, and we obtain the stationary solution

$$U(x) = \begin{cases} \theta e^x & \forall x \le 0, \\ \frac{\theta}{2}\left(2 - e^{-x}\right) & \forall x > 0. \end{cases} \tag{3.21}$$

It follows from (3.21) that U is increasing on the entire interval $(-\infty, \infty)$, and that all of the requirements in (3.2) are satisfied. To determine the upper end of Γ_0^- we substitute the condition $U(\infty) \ge \theta$ into (3.19) and obtain $U(\infty) = \frac{1}{\beta+1} \ge \theta$. This implies that the branch Γ_0^- cannot extend past $\beta = \frac{1}{\theta} - 1$. Figure 2 (lower left) shows solutions at the two endpoints of Γ_0^- when $(\epsilon, \theta) = (.1, .1)$. At the lower end where $\beta = \frac{1}{2\theta} - 1 = 4$ the solution is the monotonic stationary front $U(x)$ given in (3.21). At the upper end where $\beta = \frac{1}{\theta} - 1 = 9$ the solution is no longer monotone, but it does remain above the threshold level θ $\forall z > 0$. For each $\beta \in [4, 9]$ our computations indicate that $U(z) > \theta$ $\forall z > 0$. Thus, we conjecture that the interval of existence of Γ_0^- is the entire interval $[\frac{1}{2\theta} - 1, \frac{1}{\theta} - 1] = [4, 9]$. Our numerical study also indicates that solutions on Γ_0^- are unstable.

Remarks. To obtain Γ_0^- we kept (ϵ, θ) fixed and let β increase from the critical value $\beta_0 = \frac{1}{2\theta} - 1$, where $c_1 = 0$ and the solution is the stationary front defined in (3.21). The eigenvalues $\mu^{\pm}$ are complex at $\beta = \beta_0$ since (ϵ, θ) satisfy (2.7). It is interesting to contrast this bifurcation result with that in [2], where a similar stationary solution is found when (ϵ, θ) are chosen so that $\mu^{\pm}$ are real. In that study (θ, β) are kept fixed, and counter propagating fronts bifurcate from the stationary solution as ϵ varies. It would be interesting to determine whether a similar phenomenon occurs when $\mu^{\pm}$ are complex.

The upper branch. We let Γ_0^+ denote the branch of wave fronts corresponding to $c = c_2$ (Figure 2, upper panel). Our study of the example $(\epsilon, \beta) = (.1, .1)$ indicates that Γ_0^+ extends below $\beta = \beta_*$ down to $\beta = 0$. When $0 < \beta \le \beta_*$ the eigenvalues $\mu^{\pm}$ are real, and solutions along Γ_0^+ are monotone for large z. When $\beta > \beta_*$ and the eigenvalues are complex, solutions are oscillatory when $z > 0$ and have the form

$$U(z) = \begin{cases} \theta e^z & \forall z \le 0, \\ k_1 e^{\alpha z}\cos(\gamma z) + k_2 e^{\alpha z}\sin(\gamma z) + \frac{.5(c_2 - \epsilon)}{(c_2)^2 - (1+\epsilon)c_2 + \epsilon(\beta+1)} e^{-z} + \frac{1}{\beta+1} & \forall z > 0, \end{cases} \tag{3.22}$$

where α, γ, k_1, and k_2 are evaluated at $c = c_2$. As we noted above, to complete the proof that a solution is a wave front we need to prove that $U(z) > 0$ $\forall z > 0$. Again, this is difficult to prove when $\beta > \beta_*$ since the oscillatory component of (3.22) can cause the function U to dip below the threshold level θ at some positive z. In Figure 2 (lower right panel) the oscillatory characteristics of U are clearly visible. The solution at $\beta = 6$ qualifies as a wave front since it remains above θ on the entire interval $(0, \infty)$. However, when $\beta = 6.45$ the solution is tangent to $U = \theta$ at $z \approx 21$; hence it does not satisfy the definition of a wave front. Likewise, when $\beta > 6.45$ the function U cannot be a wave front since it dips below θ at a finite positive value of z. Thus, we conjecture that the interval of existence of Γ_0^+ is $(0, 6.45)$. Along Γ_0^+ our numerical study also indicates that solutions are stable.

D. Open problems. It remains an open problem to prove the existence of solutions along the two branches Γ_0^- and Γ_0^+ described above. In particular, whether $\mu^\pm$ are real or complex, new methods are needed in order to overcome the technical difficulties in showing that solutions satisfy $U(z) > \theta$ on the entire interval $(0, \infty)$. Although generalizations and insightful approximations have previously been given [2, 40], this property has not yet been verified for any set of parameters, even in the real eigenvalue regime. In the appendix we consider parameter values where $\mu^\pm$ are real and develop a comparison method to address similar issues that arise in the analysis of 1-pulse solutions. It is hoped that an extension of these techniques will be of help in completing the proof of existence of wave fronts for wider parameter regimes, and also for more general coupling and firing rate functions. The proof of stability of solutions along Γ_0^+ and the instability of solutions along Γ_0^- also remain an open problem. It is possible that the analysis of stability properties might be accomplished by extensions of Evans function methods developed in [12, 42, 48].

4. 1-pulse traveling waves. In this section we analyze properties of 1-pulse traveling waves when the eigenvalues $\mu^\pm$ are complex. These solutions satisfy

$$c^2U'' + c(1+\epsilon)U' + \epsilon(\beta+1)U = c\frac{d}{dz}\int_0^a w(z-z')dz' + \epsilon\int_0^a w(z-z')dz', \tag{4.1}$$

where $c > 0$, and

$$\begin{cases} U(0) = U(a) = \theta \quad \text{for some } a = a(c) > 0, \\ U(z) \neq \theta \quad \text{if } z \neq 0 \text{ or } z \neq a, \\ (U(z), U'(z)) \to (0,0) \quad \text{as } |z| \to \infty. \end{cases} \tag{4.2}$$

Following [42], we assume without loss of generality that $c > 0$. Because $w(x) = \frac{1}{2}e^{-|x|}$, equation (4.1) can be written in the equivalent form

$$c^2U'' + c(1+\epsilon)U' + \epsilon(\beta+1)U = f(z), \tag{4.3}$$

where

$$f(z) = \begin{cases} .5(c+\epsilon)e^z(1-e^{-a}) & \forall z \leq 0, \\ .5c\left(e^{-z} - e^{(z-a)}\right) + .5\epsilon\left(2 - e^{-z} - e^{(z-a)}\right) & \text{if } 0 < z < a, \\ .5(\epsilon - c)(e^a - 1)e^{-z} & \forall z \geq a. \end{cases} \tag{4.4}$$

We make use of both (4.1) and (4.3) to analyze 1-pulse waves. The remainder of this section consists of the following:

A. the construction of a family of stationary 1-pulse waves,
B. the statement and proof of Theorem 4.1 concerning the coexistence of infinitely many 1-pulse waves for positive wave speeds,
C. a description of open problems based on the results of numerical experiments.

A. Stationary solutions. We set $c = 0$ in (4.1) and look for time independent solutions

$$U(x) = \frac{1}{\beta + 1} \int_0^a w(x - x')dx', \tag{4.5}$$

which satisfy (4.2). Setting $w(x - x') = \frac{1}{2}e^{-|x-x'|}$, we define a branch of stationary pulse-shaped solutions (Figure 3) by

$$U(x) = \begin{cases} \frac{.5(1-e^{-a})}{\beta+1} e^x & \forall x \le 0, \\ \frac{.5}{\beta+1}\left(2 - e^{-x} - e^{x-a}\right) & \text{if } 0 < z < a, \\ \frac{.5(e^a-1)}{\beta+1} e^{-x} & \forall x \ge a, \end{cases} \tag{4.6}$$

where

$$a = \ln\left(\frac{1}{1 - 2\theta(\beta + 1)}\right) \quad \forall \beta \in \left(0, \frac{1}{2\theta} - 1\right). \tag{4.7}$$

A similar family of stationary solutions was found by Pinto, Jackson, and Wayne [42]. Figure 3 illustrates specific solutions for the representative parameter set $(\epsilon, \theta) = (.1, .1)$ and for β values where $\mu^{\pm}$ are complex. Our numerical experiments indicate that the solutions defined by (4.5)–(4.6) are unstable. Again, an extension of the Evans function methods in [12] might help prove this conjecture.

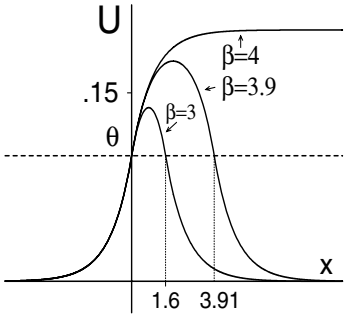


Figure 3. *Stationary 1-pulse wave solutions defined by (4.5)–(4.6). Lower solution: $(\beta, a) = (3, 1.6)$. Middle solution $(\beta, a) = (3.9, 3.91)$. Upper solution: as $a \to \infty$, the stationary solutions evolve into a limiting monotonic solution at $\beta = \frac{1}{2\theta} - 14$. This limiting solution is the stationary wave front defined in (3.21).*

B. Positive wave speeds. When $c > 0$ our goal is to show that infinitely many 1-pulse traveling waves can coexist when the eigenvalues are complex. We prove the following result.

Theorem 4.1. *Let (ϵ, θ) satisfy (2.7), and let $\beta = \frac{1}{2\theta} - 1$. There are infinitely many values $c \in (c_1, c_2)$ and $a(c) > 0$, and corresponding solutions U of (4.3)–(4.4), such that*

$$U(0) = U(a(c)) = \theta. \tag{4.8}$$

Furthermore, $U(z)$ satisfies

$$U(z) = \theta e^z \quad \forall z \le 0. \tag{4.9}$$

Remarks. To prove that the solutions described in Theorem 4.1 are truly 1-pulse traveling waves we must also prove that $z = 0$ and $z = a(c)$ are the *only* solutions of $U(z) = \theta$, and that $(U(z), U'(z)) \to (0, 0)$ as $z \to \infty$. Because the eigenvalues are now complex, technical difficulties arise which make the verification of these properties a challenging problem. Further details of this and related open problems are discussed following the proof of the theorem.

Proof of Theorem 4.1. The first step is to consider the interval $(-\infty, 0)$, where (4.3)–(4.4) reduces to

$$c^2U'' + c(1+\epsilon)U' + \epsilon(\beta+1)U = .5(c+\epsilon)(1-e^{-a})e^z \quad \forall z \le 0. \tag{4.10}$$

On $(-\infty, 0)$ the general solution of (4.10) is

$$U_2 = h_1 e^{\alpha z}\cos(\gamma z) + h_2 e^{\alpha z}\sin(\gamma z) + P_2(z), \tag{4.11}$$

where $P_2(z)$ is the particular solution

$$P_2 = \frac{.5(\epsilon+c)(1-e^{-a})}{c^2+(1+\epsilon)c+\epsilon(\beta+1)}e^z \quad \forall z \le 0. \tag{4.12}$$

We need to show that there are values $c > 0$ and $a > 0$ such that

$$U_2(z) < \theta \ \ \forall z \in (-\infty, 0), \quad U_2(-\infty) = U_2'(-\infty) = 0, \quad \text{and} \quad U_2(0) = \theta. \tag{4.13}$$

Recall from (2.6) that $\alpha = \mathrm{Re}(\mu^{\pm}) < 0$. Thus, to satisfy the condition $U_2(-\infty) = U_2'(-\infty) = 0$ we conclude that $h_1 = h_2 = 0$, and therefore

$$U_2(z) = \frac{.5(\epsilon+c)(1-e^{-a})}{c^2+(1+\epsilon)c+\epsilon(\beta+1)}e^z \qquad \forall z \le 0. \tag{4.14}$$

Substituting the requirement $U_2(0) = \theta$ into (4.14) leads to

$$\frac{.5(\epsilon+c)(1-e^{-a})}{c^2+(1+\epsilon)c+\epsilon(\beta+1)} = \theta. \tag{4.15}$$

Combining (4.14) and (4.15) gives $U_2(z) = \theta e^z \ \forall z \le 0$ (Figure 4).

The interval $(0, a)$. On the interval $(0, a)$ the problem (4.3)–(4.4) reduces to

$$c^2U'' + c(1+\epsilon)U' + \epsilon(\beta+1)U = .5c\left(e^{-z} - e^{(z-a)}\right) + .5\epsilon\left(2 - e^{-z} - e^{(z-a)}\right). \tag{4.16}$$

The general solution of (4.16) is

$$U_3(z) = m_1 e^{\alpha z}\cos(\gamma z) + m_2 e^{\alpha z}\sin(\gamma z) + P_3(z), \tag{4.17}$$

where α and γ are defined in (2.6) and P_3 is the particular solution

$$P_3(z) = \frac{.5(c-\epsilon)}{c^2-(1+\epsilon)c+\epsilon(\beta+1)}e^{-z} - \frac{.5(\epsilon+c)}{c^2+(1+\epsilon)c+\epsilon(\beta+1)}e^{z-a} + \frac{1}{\beta+1}. \tag{4.18}$$

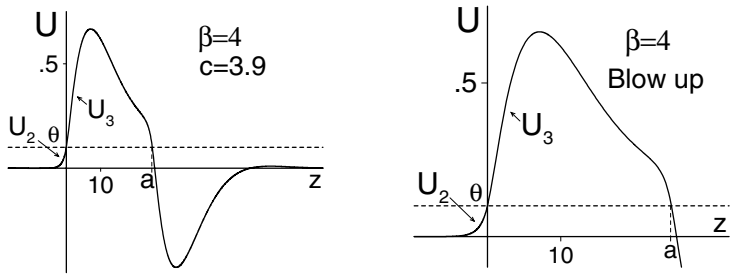


Figure 4. *Left panel:* 1*-pulse traveling wave when* $\epsilon = \theta = .1$, $\beta = \frac{1}{2\theta} - 1 = 4$, $c \approx c_2 = 3.9$, *and* $a \approx 25$. *The component* U_2 *of the solution is defined by* $U_2(z) = \theta e^z \ \forall z \le 0$. *The component* $U_3(z)$ *is defined on* $[0, \infty)$ *and has the form given by* (4.17)–(4.18). *Right panel: blow up of the solution.*

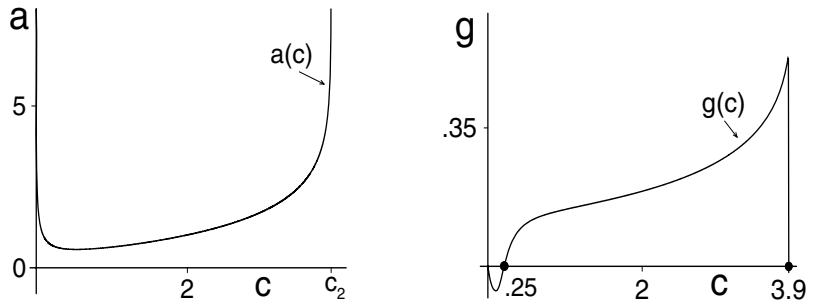


Figure 5. *Left panel:* $a(c)$ *defined in* (4.20) *when* $(\epsilon, \theta) = .(1, .1)$ *and* $\beta = \frac{1}{2\theta} - 1 = 4$. *Right panel:* $g(c)$ *defined in* (4.20).

We need to show that there are values $c > 0$ and $a = a(c) > 0$ which satisfy the algebraic condition (4.15) such that

$$U_3(0) = U_2(0) = U_3'(0) = U_2'(0) = U_3(a) = \theta. \tag{4.19}$$

We begin by solving (4.15) for e^{-a}. This gives

$$e^{-a(c)} = \frac{.5(\epsilon + c) - \theta\left(c^2 + (1+\epsilon)c + \epsilon(\beta+1)\right)}{.5(\epsilon + c)}. \tag{4.20}$$

Since the right-hand side of (4.20) is zero at $c = c_1$ and $c = c_2$, it follows (see Figure 5) that

$$a(c) > 0 \ \ \forall c \in (c_1, c_2) \quad \text{and} \quad \lim_{c \to c_2^-} a(c) = \lim_{c \to c_1^+} a(c) = \infty. \tag{4.21}$$

To preserve continuity at $z = 0$ we require that $(U_3(0), U_3'(0)) = (U_2(0), U_2'(0)) = (\theta, \theta)$. Combining this with (4.17) and (4.18) shows that the coefficients m_1 and m_2 in (4.17) are uniquely defined by

$$m_1 = \theta - P_3(0) \quad \text{and} \quad m_2 = \frac{1}{\gamma}\left(\theta(1-\alpha) - P_3'(0) + \alpha P_3(0)\right). \tag{4.22}$$

We need to show that there are infinitely many speeds $c \in (c_1, c_2)$, and corresponding values $a = a(c) > 0$ such that

$$U_3(a(c)) = \theta. \tag{4.23}$$

Substituting (4.17) into (4.23) gives

$$m_1 e^{\alpha a(c)} \cos(\gamma a(c)) + m_2 e^{\alpha a(c)} \sin(\gamma a(c)) + P_3(a(c)) = \theta. \tag{4.24}$$

Thus, to prove that 1-pulse traveling waves exist, it is necessary to show that there are values $c > 0$ such that the function

$$g(c) = m_1 e^{\alpha a(c)} \cos(\gamma a(c)) + m_2 e^{\alpha a(c)} \sin(\gamma a(c)) + P_3(a(c)) - \theta \tag{4.25}$$

satisfies $g(c) = 0$. The remainder of the proof is devoted to showing that $g(c)$ has infinitely many zeros on (c_1, c_2). The first step is to write (4.20) as

$$\frac{.5(\epsilon + c)}{c^2 + (1+\epsilon)c + \epsilon(\beta+1)} = \frac{.5(\epsilon + c)e^{-a}}{c^2 + (1+\epsilon)c + \epsilon(\beta+1)} + \theta. \tag{4.26}$$

Substituting (4.26) into (4.18) and using the hypothesis that $\beta = \frac{1}{2\theta} - 1$ gives

$$P_3(a(c)) - \theta = e^{-a(c)} \left(\frac{.5(c - \epsilon)}{c^2 - (1+\epsilon)c + \epsilon(\beta+1)} - \frac{.5(\epsilon + c)}{c^2 + (1+\epsilon)c + \epsilon(\beta+1)} \right). \tag{4.27}$$

Next, substitute (4.27) into (4.25) and obtain

$$g(c) = e^{\alpha a(c)} g_1(c), \tag{4.28}$$

where

$$\begin{aligned} g_1 &= m_1 \cos(\gamma a(c)) + m_2 \sin(\gamma a(c)) \\ &\quad + e^{-a(c)(1+\alpha)} \left(\frac{c^2 - \epsilon^2(\beta+1)}{(c^2 - (1+\epsilon)c + \epsilon(\beta+1))(c^2 + (1+\epsilon)c + \epsilon(\beta+1))} \right). \end{aligned} \tag{4.29}$$

Because $e^{\alpha a(c)} > 0$, it suffices to show that $g_1(c)$ changes sign infinitely often on (c_1, c_2). To analyze $g_1(c)$ we need to determine the limiting behavior of the terms $1 + \alpha$, $\gamma a(c)$, $c^2 - (1+\epsilon)c + \epsilon(\beta+1)$, and m_1 as $c \to c_2^-$. First, from (2.6) it follows that

$$\lim_{c \to c_2^-} (1 + \alpha) = 1 - \frac{(1+\epsilon)}{2c_2} = \frac{2c_2 - 1 - \epsilon}{2c_2}. \tag{4.30}$$

It follows from (3.11) and (2.7) that $c_2 = \frac{.5 - \theta(1+\epsilon)}{\theta} > 0$ when $\beta = \frac{1}{2\theta} - 1$. We substitute this into (4.30) and find, after an algebraic manipulation, that

$$\lim_{c \to c_2^-} (1 + \alpha) = \frac{1 - 3\theta(1+\epsilon)}{1 - 2\theta(1+\epsilon)} > 0. \tag{4.31}$$

Remarks. The restrictions in (2.7) imply that $0 < \theta < \frac{1}{3(\epsilon+1)} < \frac{1}{2(\epsilon+1)}$, and therefore the right-hand side of (4.31) is positive. Thus, if we show that $\gamma a(c) \to \infty$ as $c \to c_2^-$, then the term $e^{-a(c)(1+\alpha)}$ on the right-hand side of (4.29) decays exponentially fast as $c \to c_2^-$. If we also show that $c^2 - (1+\epsilon)c + \epsilon(\beta+1)$ and m_1 are bounded away from zero as $c \to c_2^-$, then the first two terms on the right-hand side of (4.29) will dominate when $c \approx c_2$, and we will be able to prove that g_1 has infinitely many zeros.

First, we analyze $\gamma a(c)$. Recall that $\beta = \frac{1}{2\theta} - 1$. This and (2.6)–(2.7) imply that

$$\gamma = \frac{(\epsilon+1)}{2c\sqrt{\theta}} \sqrt{\frac{2\epsilon}{(\epsilon+1)^2} - \theta} > 0. \tag{4.32}$$

From this and (4.21) we conclude that

$$\gamma a(c) > 0 \quad \forall c \in (c_1, c_2) \quad \text{and} \quad \lim_{c \to c_2^-} \gamma a(c) = \infty. \tag{4.33}$$

Next, we prove that $c^2 - (1+\epsilon)c + \epsilon(\beta+1)$ is bounded away from zero as $c \to c_2^-$. It follows from (3.10) that $c_2 = \frac{.5-\theta(1+\epsilon)}{\theta} > 0$ when $\beta = \frac{1}{2\theta} - 1$. This implies that

$$(c_2)^2 - (1+\epsilon)c_2 + \epsilon(\beta+1) = c_2 \left(\frac{1}{2\theta} - 2(1+\epsilon) \right) + \frac{\epsilon}{2\theta}. \tag{4.34}$$

It follows from (2.7) that $\frac{1}{2\theta} - 2(1+\epsilon) > 0$. This and (4.34) imply that $(c_2)^2 - (1+\epsilon)c_2 + \epsilon(\beta+1)$ is bounded away from zero as $c \to c_2^-$.

Finally, we show that m_1 is bounded away from 0 as $c \to c_2^-$. Since $a(c) \to \infty$ as $c \to c_2^-$, it follows from the definition of m_1 given in (4.22) and (4.18) that

$$\lim_{c \to c_2^-} m_1 = \theta - \frac{.5(c_2 - \epsilon)}{(c_2)^2 - (1+\epsilon)c_2 + \epsilon(\beta+1)} - \frac{1}{\beta+1}. \tag{4.35}$$

We substitute (4.34) into (4.35), combine the first two terms, and obtain

$$\lim_{c \to c_2^-} m_1 = \frac{\theta(\epsilon - 2\theta(1+\epsilon)c_2)}{(.5 - 2\theta(1+\epsilon))c_2 + .5\epsilon} - \frac{1}{\beta+1}. \tag{4.36}$$

Setting $c_2 = \frac{.5-\theta(1+\epsilon)}{\theta} > 0$ in the numerator of the first term in (4.36) gives

$$\lim_{c \to c_2^-} m_1 = \frac{\theta(-1 + 2\theta(1+\epsilon)^2)}{(.5 - 2\theta(1+\epsilon))c_2 + .5\epsilon} - \frac{1}{\beta+1}. \tag{4.37}$$

Recall from (2.7) that $0 < \epsilon < 1$ and $0 < \theta < \frac{1}{4(\epsilon+1)}$. This implies that $-1 + 2\theta(1+\epsilon)^2 < .5(\epsilon - 1) < 0$ and $.5 - 2\theta(1+\epsilon) > 0$. Thus, the first term in (4.37) is negative, and therefore

$$\lim_{c \to c_2^-} m_1 < -\frac{1}{\beta+1} = -2\theta. \tag{4.38}$$

We now use the estimates in (4.31), (4.33), (4.34), and (4.38) to determine the behavior of $g_1(c)$. It follows from (4.33) that there is an increasing sequence $\{c_n\}$ such that

$$c_n \to c_2^- \quad \text{as } n \to \infty \quad \text{and} \quad \gamma a(c_n) = 2n\pi \quad \text{for large } n. \tag{4.39}$$

Let $c = c_n$ in (4.29). Then $\sin(\gamma a(c_n)) = 0$, $\cos(\gamma a(c_n)) = 1$, and (4.29) reduces to

$$g_1(c_n) = m_1 + e^{-a(c_n)(1+\alpha)} \left(\frac{c_n^2 - \epsilon^2(\beta+1)}{(c_n^2 - (1+\epsilon)c_n + \epsilon(\beta+1))(c_n^2 + (1+\epsilon)c_n + \epsilon(\beta+1))} \right). \tag{4.40}$$

Combining the estimates in (4.31), (4.33), (4.34), (4.38), and (4.39), we conclude that

$$g_1(c_n) < -\frac{1}{2(\beta+1)} \quad \text{for } n \gg 1. \tag{4.41}$$

Likewise, there is an increasing sequence $\{c^n\}$ such that

$$c^n \to c_2^- \quad \text{as } n \to \infty \quad \text{and} \quad \gamma a(c^n) = (2n+1)\pi \quad \text{for } n \gg 1. \tag{4.42}$$

Let $c = c^n$ in (4.29). From (4.31), (4.33), (4.34), (4.38), and (4.42) it follows that

$$g_1(c^n) > \frac{1}{2(\beta+1)} \quad \text{for large } n \gg 1. \tag{4.43}$$

It now follows from (4.41) and (4.43) and continuity that $g_1(c)$, and therefore $g(c)$, have infinitely many zeros on (c_1, c_2). Thus, we we have proved that there are infinitely many $c \in (c_1, c_2)$ and $a(c) > 0$, and corresponding solutions U of (2.2), such that

$$U(z) = \begin{cases} \theta e^z & \forall z \le 0, \\ U_3(z) & 0 < z < a(c), \end{cases} \tag{4.44}$$

where $U_3(z)$ is defined in (4.17)–(4.18) and satisfies

$$U_3(0) = U_3'(0) = U_3(a(c)) = \theta. \tag{4.45}$$

This completes the proof of Theorem 4.1.

C. Open problems. As mentioned earlier, to show that solutions are truly 1-pulse traveling waves we must also prove that $z = 0$ and $z = a(c)$ are the *only* solutions of $U(z) = \theta$, and that $(U(z), U'(z)) \to (0,0)$ as $|z| \to \infty$. It follows from (4.44) that

$$U(z) < 0 \ \ \forall z < 0 \quad \text{and} \quad (U(z), U'(z)) \to (0,0) \ \ \text{as } z \to -\infty. \tag{4.46}$$

Thus, the function $U(z) - \theta$ has no zeros on $(-\infty, 0)$. Our numerical experiments indicate that solutions satisfy $U(z) > \theta$ on the entire interval $(0, a(c))$. However, because the eigenvalues $\mu^{\pm}$ are complex, solutions have oscillatory tails which can cause $U(z)$ to exceed the threshold level θ at some point in (a, ∞). This implies that the function $U(z) - \theta$ has at least three zeros, and therefore $U(z)$ cannot be a 1-pulse wave. Figure 6 illustrates the different types of behavior for the representative parameter set $(\epsilon, \theta) = (.1, .1)$. In the right panel we let

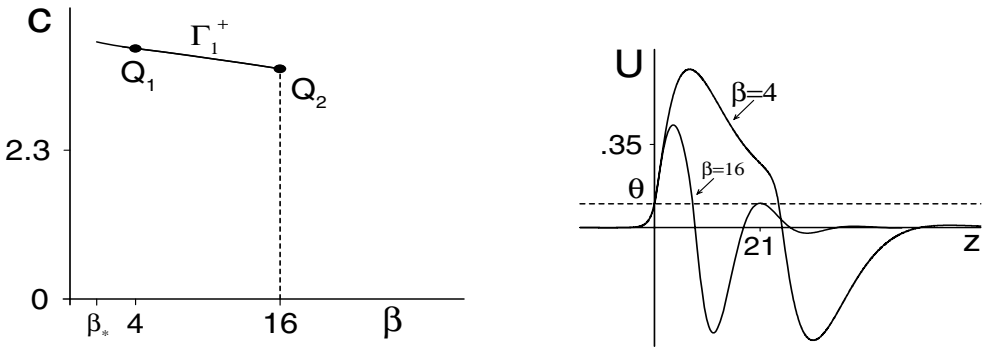


Figure 6. *Left panel: the bifurcation curve Γ_1^+ for a branch of 1-pulse traveling waves when $(\epsilon, \theta) = (.1, .1)$. The vertical axis gives wave speed c, and the horizontal axis denotes values of β. The branch Γ_1^+ extends from $\beta = \beta_* = 2.025$ to $\beta = 16$. Right panel: solutions on Γ_1^+ at Q_1, where $\beta = \frac{1}{2\theta} - 1 = 4$ and $c \approx 3.9$, and at Q_2, where $(\beta, c) = (16, 3.56)$. The solution at Q_2 is not a 1-pulse wave since it is tangent to $U = \theta$ at $z \approx 21$. When $\beta > 16$ solutions rise above θ at the second maximum, and therefore cannot be 1-pulse waves. Thus, we conjecture that Γ_1^+ ends at $\beta = 16$.*

$\beta = \frac{1}{2\theta} - 1 = 4$ and graph the solution corresponding to the second positive zero of $g(c)$. The width of this solution is $a \approx 25$; hence $e^{-a} = e^{-25} \approx 10^{-11}$. Using the approximation $e^{-a} = 0$ in (4.15) and solving for c, we find that $c = 3.9 \approx c_2$, where c_2 is the "fast" wave speed defined in (3.11). In addition, the graph of this solution indicates that all of the conditions (4.2) for a 1-pulse wave are satisfied. The bifurcation curve Γ_1^+ (left panel) denotes the family of 1-pulse traveling waves which form a continuation of the solution computed at $\beta = 4$. This branch extends from $\beta = \beta_* = 2.025$, where $\mu^\pm$ become complex, to $\beta = 16$. The solution at $\beta = 16$ is graphed in the right panel. This solution is tangent to $U = \theta$ at $z = a \approx 20.8$; hence it does not satisfy conditions (4.2). When $\beta > 16$ solutions also intersect $U = \theta$ more than twice, and therefore these functions cannot be 1-pulse traveling waves either. Thus, the branch Γ_1^+ ends at $\beta = 16$. Our results also suggest that $c \approx c_2(\beta)$ along Γ_1^+, and that solutions on Γ_1^+ are stable. A rigorous analysis of the properties described above is beyond the scope of this investigation, and therefore their proofs remain as open problems for future study.

Finally, we point out that numerical and analytical evidence suggests that 1-pulse traveling waves can also exist for parameter values where $\mu^\pm$ are real [2, 40, 46]. In this setting, although insightful approximations have previously been given, the proof that $U(z) = \theta$ exactly twice on $(-\infty, \infty)$ has not yet been completed. As a first step towards resolving this issue we again let $w(x) = \frac{1}{2}e^{-|x|}$, and in the appendix we develop a comparison method which shows that when $\mu^\pm$ are real, $U(z) = \theta$ has no solution on (a, ∞). It remains an open problem to extend our technique and prove that $U(z) = \theta$ has no solution on $(0, a)$. It also remains an open problem to extend our methods to the complex eigenvalue regime, and also to systems with more general coupling and firing rate functions.

5. N-pulse waves. Finite wave trains have been observed in feline cortex [4, 5, 6], in the brain of freely moving mice [18], and also in tangential and coronal brain slice experiments [30, 55]. Thus, in this section our goal is to understand how N-pulse waves arise in the one-dimensional model. The insights we obtain provide a basis for the study of two-dimensional wave formation in section 6.

In the appendix we explain why N-pulse waves are not expected when $\mu^{\pm}$ are real. Thus, we focus our attention on the regime $\beta > \beta_* = \frac{(\epsilon-1)^2}{4\epsilon}$, where $\mu^{\pm}$ are complex. To understand how such waves can form when $\beta > \beta_*$ we need to analyze the linearization of (2.2) around the rest state $U = 0$:

$$c^2 H'' + c(1+\epsilon)H' + \epsilon(\beta+1)H = 0. \tag{5.1}$$

When $\mu^{\pm}$ are complex the general solution of (5.1) is

$$H(z) = b_1 e^{\alpha z}\sin(\gamma z) + b_2 e^{\alpha z}\cos(\gamma z), \tag{5.2}$$

where

$$\alpha = \mathrm{Re}(\mu^{\pm}) = \frac{-(\epsilon+1)}{2c} < 0 \quad \text{and} \quad \gamma = \mathrm{Im}(\mu^{\pm}) = \frac{\sqrt{4\epsilon\beta - (\epsilon-1)^2}}{2c} > 0. \tag{5.3}$$

It follows from (5.2)–(5.3) that the frequency of oscillation of $H(z)$ increases as β increases from β_*. In turn, this causes solutions of (1.1) to become more oscillatory as β increases, making it increasingly likely that an initial perturbation will evolve into an N-pulse traveling wave. The remainder of this section is devoted to presenting numerical evidence for the existence of N-pulse waves. To gain insight we consider the representative parameter set $(\epsilon, \theta) = (.1, .1)$ and study the dynamics of wave formation when $\beta > \beta_* = 2.025$. At specific β values (see Figure 7) we solve the initial value problem

$$\begin{aligned} u_t(x,t) &= -u - v + \frac{1}{2}\int_{-100}^{100} e^{-|x-x'|} H(u(x',t)-\theta)dx', \\ v_t(x,t) &= \epsilon(\beta u - v), \\ u(x,0) &= .6e^{-x^2} \quad \text{and} \quad v(x,0) = 0 \quad \forall x \in [-100,100], \end{aligned} \tag{5.4}$$

where the limits $(-\infty, \infty)$ in the integral term in (1.1) have been replaced with $[-100, 100]$. Other choices for initial conditions give results similar to those described below. A second approach which leads to N-pulse wave formation is to initially keep u and v at their resting levels, i.e., $u(x,0) = v(x,0) = 0$, and perturb the system with an external stimulus applied to the right-hand side of the equation for u. To solve (5.4) we approximate the integral term with a Riemann sum and use an Euler time step $\Delta t = .1$. When $\beta > \beta_*$ the solution of (5.4) evolves into a traveling wave, and the number of peaks in the waves increases as β increases. There is a transition from a wave front (Figure 7, top panel) to a 1-pulse wave as β passes through $\beta = 6.45$ from below. This is consistent with the observation in section 3 that $\beta = 6.45$ is the critical value where the upper branch Γ_0^- of wave fronts ceases to exist (Figure 2). As β increases from $\beta = 6.45$, the 1-pulse wave continues to exist (Figure 7, lower left panel) until a second critical value, $\beta \approx 10.5$, is reached, where the transition to a 2-pulse wave occurs. It follows from (2.2) that 2-pulse waves (lower panel, middle) satisfy

$$\begin{aligned} c^2 U'' + c(1+\epsilon)U' + \epsilon(\beta+1)U = {} & c\frac{d}{dz}\int_0^a w(z-z')dz' + \epsilon\int_0^a w(z-z')dz' \\ & + c\frac{d}{dz}\int_b^d w(z-z')dz' + \epsilon\int_b^d w(z-z')dz', \end{aligned} \tag{5.5}$$

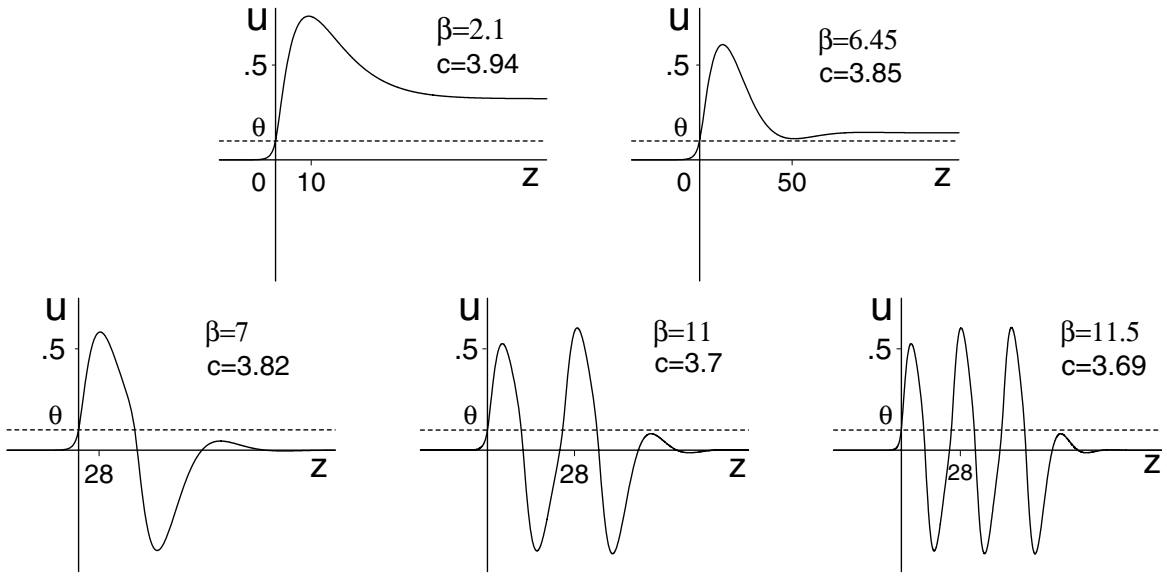


Figure 7. *Traveling wave solutions when* $(\epsilon, \theta) = (.1, .1)$ *and* $\beta > \beta_*$. *Upper row: stable wave fronts. As* β *passes through* $\beta = 6.45$ *from below there is a transition from a wave front to a* 1*-pulse traveling wave. Lower row:* 1*-pulse wave,* 2*-pulse wave, and* 3*-pulse wave. Clicking on the above images displays the accompanying movies* (66638_02.mpg [1.6MB], 66638_01.mpg [1.7MB], 66638_03.mpg [2MB], 66638_04.mpg [3MB], *and* 66638_05.mpg [5.3MB]).

where

$$\begin{cases} U(0) = U(a) = U(b) = U(d) = \theta \quad \text{for some } d > b > a > 0, \\ U(z) \neq \theta \quad \text{if } z \notin \{0, a, b, d\}, \\ (U(z), U'(z)) \to (0, 0) \quad \text{as } |z| \to \infty. \end{cases} \tag{5.6}$$

As β increases further, the number of peaks in the wave also increases. For example, a 3-pulse wave exists at $\beta = 11.5$ (lower panel, right). For general $N > 2$ a simple extension of (5.5)–(5.6) gives the criteria satisfied by N-pulse traveling waves. In particular, it is necessary to prove that the solution intersects $U = \theta$ exactly 2N times. However, as with 1-pulse waves, the nonlocal terms in the equation lead to technical difficulties in proving this property, and therefore existence proofs remain a challenging open problem. In contrast, nonlocal terms do not appear in reaction-diffusion equations where the existence of N-pulse waves can be proved using topological shooting methods [27]. Finally, we note that our numerical experiments suggest that the traveling waves in Figure 7 are all stable as solutions of (1.1). It would be of interest to develop Evans function methods to prove this conjecture.

6. Two dimensions. In this section we extend the one-dimensional model (1.1) to the two-dimensional system

$$\begin{aligned} u_t &= -u - v + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x, y, x', y') H(u(x', y', t) - \theta) dx' dy' + \zeta(x, y, t), \\ v_t &= \epsilon(\beta u - v). \end{aligned} \tag{6.1}$$

Here we have replaced the coupling function $w(x-x') = \frac{1}{2}e^{-|x-x'|}$ in (1.1) with the more general two-dimensional coupling

$$w(x,y,x',y') = 2.1e^{-\sqrt{(x-x')^2+(y-y')^2}}g(x',y'). \tag{6.2}$$

When $g(x',y') \equiv 1$ the function w represents symmetric connections, whereas $g(x',y') \neq 1$ corresponds to asymmetric connections. The function $\zeta(x,y,t)$ denotes external input. As a first step towards understanding the dynamics of wave formation we use a Fourier transform method developed in [34] and approximate (6.1) by the PDE system

$$\begin{aligned} &\left(\nabla^4 + 2M\nabla^2 + B + M^2\right)(u_t + u + v - \zeta(x,y,t)) = Ag(x,y)H(u-\theta), \\ &v_t(x,y,t) = \epsilon(\beta u - v), \end{aligned} \tag{6.3}$$

where $(A,B,M) = (7,.52,-2.5)$. We solve (6.3) with an Euler time step of length $\Delta t = .35$, and use finite differences to approximate the spatial derivatives on disk-shaped spatial domains $\Omega = \{(x,y) \mid x^2+y^2 < \mathrm{R}^2\}$, with Neumann conditions on the boundary. Traveling waves can be initiated with initial conditions of the form

$$(u(x,y,0),v(x,y,0)) = \left(\lambda e^{-\kappa\sqrt{x^2+y^2}},0\right) \quad \text{and} \quad \zeta(x,y,0) = 0, \quad 0 \le x^2+y^2 \le \mathrm{R}. \tag{6.4}$$

Alternatively, we could initially keep u and v in the rest state, i.e., $(u(x,y,0),v(x,y,0)) \equiv (0,0)$, and initiate waves with an external stimulus of the form $\zeta(x,y,0) = \lambda e^{-\kappa\sqrt{x^2+y^2}}$.

As in section 5, we restrict our attention to the representative parameter set $(\epsilon,\theta) = (.1,.1)$ and investigate the dynamics of wave formation when $\beta > \beta_*$. Below we give a brief description of the results of our numerical experiments. These include the following:

(I) the formation and propagation of ring-shaped waves;
(II) the periodic formation of traveling waves;
(III) the formation of spiral waves.

(I) *Ring-shaped waves.* Here we let $g(x',y') \equiv 1$, so that the function w defined in (6.2) describes symmetric connections. We assume that $\zeta(x,y,t) \equiv 0$ (i.e., no external input), and that the initial conditions are given by

$$(u(x,y,0),v(x,y,0)) = \left(5e^{-.05\sqrt{x^2+y^2}},0\right), \qquad 0 \le x^2+y^2 \le 65. \tag{6.5}$$

When $\beta > \beta_*$, axially symmetric ring-shaped traveling waves form and propagate outwards from $(x,y) = (0,0)$. Figure 8 illustrates single-ring and double-ring waves when $\beta = 3$ and $\beta = 6$, respectively. As β increases further, the number of rings in the solution also increases. Furthermore, when $\beta = 6$ our numerical experiments show that an initial condition of smaller magnitude evolves into a single-ring wave. This leads us to conjecture that multiring waves can coexist as stable solutions.

(II) *The periodic formation of traveling waves.* To understand how waves can periodically form we fix β at the representative value $\beta = 3$ and let the initial conditions be given by (6.5). As in (I), the solution evolves into a ring-shaped wave which propagates outwards from $(x,y) = (0,0)$ (Figure 8 (left)). At $t = 10$ we apply an external stimulus, which causes a small break in the ring. Figure 9 illustrates how the solution evolves when $t > 10$, and

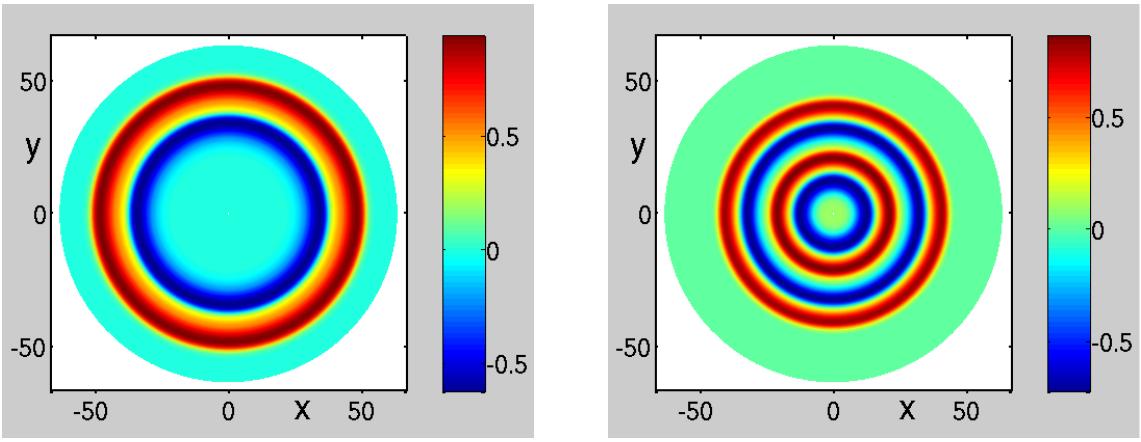


Figure 8. *Single-ring and double-ring wave solutions when $\beta = 3$ (left) and $\beta = 6$ (right). The initial conditions are given by (6.5). Clicking on the above images displays the accompanying movies (66638_06.mpg [320KB] and 66638_07.mpg [680KB]).*

outwardly propagating waves are periodically produced (panels (e) and (f)). Since the system is translationally invariant this phenomenon can be initiated at any point in two-dimensional space. It depends only on the presence of connections and occurs without any persistent external input present. To our knowledge this mechanism for the periodic production of traveling waves has not previously been reported for this type of model.

(III) *The formation of spiral waves.* For brevity we describe only our numerical experiments when the coupling function is asymmetric. For this we set $\beta = 3$ and let $g(x, y)$ be a small perturbation from $g = 1$ which, in polar form, is given by

(6.6)
$$g = \frac{1}{2}\left(1 + \exp\left(.01r\sin\left(.2\left[\theta + \frac{3\pi}{4}\right]\right)\right)\right), \qquad r = \sqrt{x^2 + y^2} \quad \text{and} \quad \theta = \tan^{-1}\left(\frac{y}{x}\right).$$

We describe two ways in which spiral waves can form. First, one can proceed as in (II) and set $\zeta(x, y, 0) \equiv 0$, assuming that the initial conditions are given by (6.5). A ring-shaped wave forms and propagates outwards from $(x, y) = (0, 0)$, and at $t = 10$ we apply a perturbation which breaks the leading edge of a ring. As in (II), the two free ends of the break form into rotors. However, if the perturbation is of sufficiently large magnitude, then the rotors cannot meet. Instead, the solution immediately evolves into a spiral wave which fills the entire two-dimensional domain. A second method (Figure 10) is to initially keep u and v at their resting values, i.e., $(u(x, y, 0), v(x, y, 0)) \equiv (0, 0)$, and apply a sequence of appropriately timed external stimuli which are focused at an arbitrarily chosen point, for example, $(x, y) = (0, 0)$. We consider the representative external stimulus defined by

(6.7)
$$\zeta(x, y, t) = \begin{cases} 5e^{-.05\sqrt{x^2+y^2}} & \text{if } \ t \in \{0, t_{12}, t_{24}, t_{36}, \ldots, t_{96}\}, \\ 0 & \text{otherwise.} \end{cases}$$

Here $t_N = .35N$ denotes the Nth time step in the numerical solution scheme. During the time interval $[0, t_{12}] = [0, 4.2]$ a ring-shaped wave forms and begins to propagate outwards

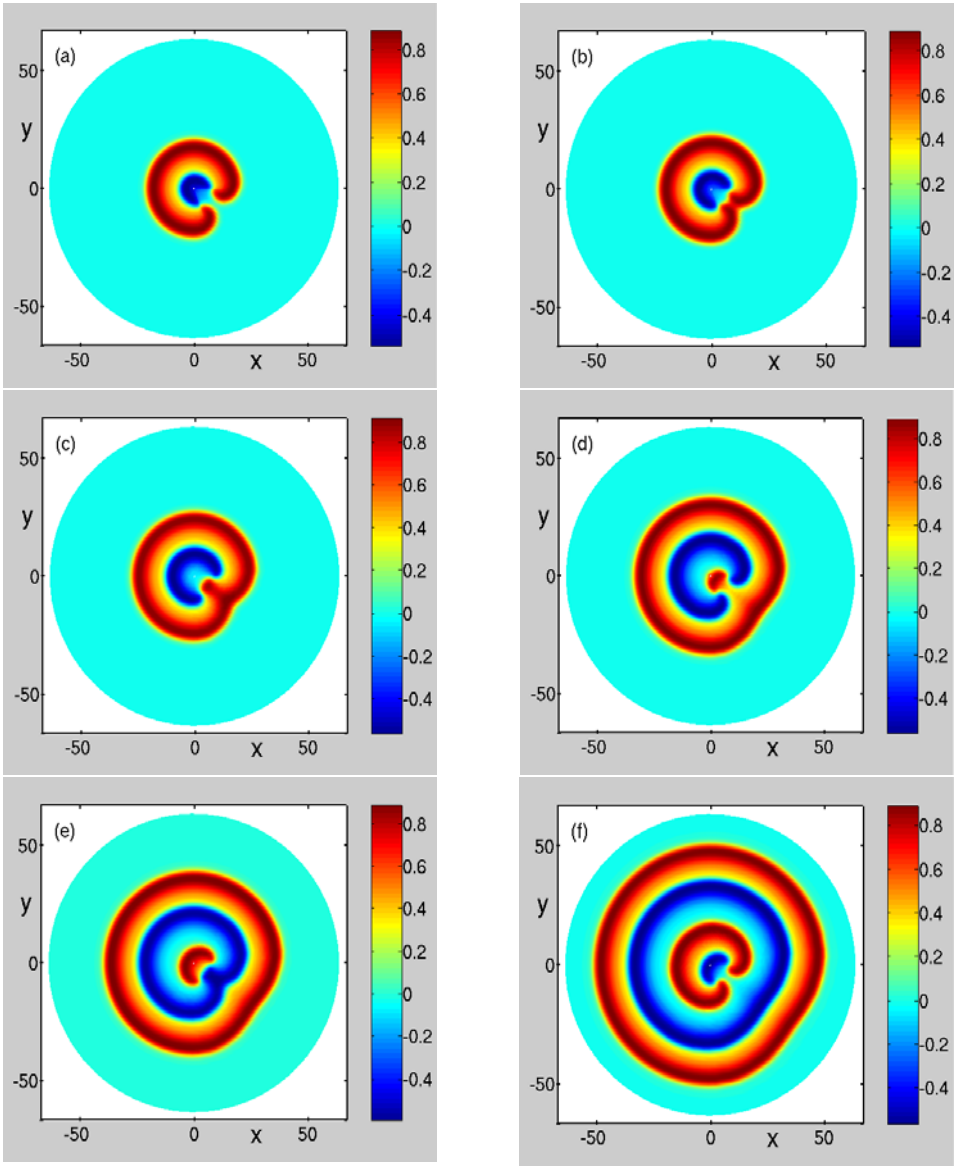


Figure 9. *Outwardly propagating waves are periodically produced when a ring-shaped wave receives an appropriately timed external stimulus that breaks the ring. Here $\beta = 3$, and the initial conditions are given by (6.5). At first, the two free ends form rotors (panel* (a)*). Their leading edges propagate outwards until the rotors meet (panel* (b)*), and then a transition takes place in which two new rotors are formed, and a newly formed traveling wave breaks away and propagates outwards (panels* (c) *and* (d)*). This process is continuously repeated (panels* (e) *and* (f)*). Clicking on the above images displays the accompanying movie* (66638_08.mpg [2.8MB]).

from $(0, 0)$. The second external stimulus is applied at $t = t_{12} = 4.2$, while the region behind the leading edge of the wave is in the refractory state. A second wave forms and propagates outwards, and we apply the third timed stimulus at $t = t_{24} = 8.4$, just before the recovery period for the formation of the second wave is complete. Because the coupling is slightly

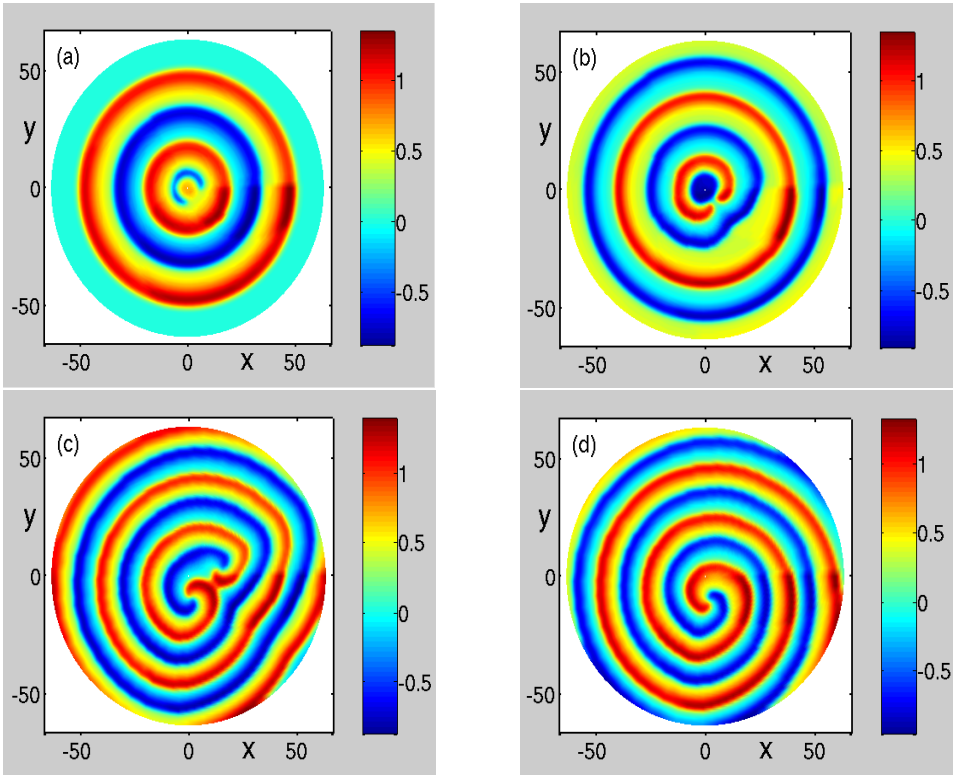


Figure 10. *The formation of a spiral wave in response to the sequence of stimuli defined in* (6.7)*. At* $t = 96$ *the outer edge of a wave breaks at a point where the activity has minimal amplitude, and the two free ends form into rotors (panel* (a)*). As* t *increases further, the leading edges of the rotors propagate outwards until they meet, and then a transition occurs in which two new rotors are formed, and a newly formed wave breaks away and propagates outwards (panel* (b)*). The asymmetry in the coupling causes one of the rotors to rotate more slowly than the other (panel* (c)*). Eventually this prevents their outer edges from meeting, causing the formation of outgoing waves to cease, and the solution evolves into a spiral wave which fills the entire domain (panel* (d)*). Clicking on the above images displays the accompanying movie* (66638_09.mpg [6.1MB]).

inhomogeneous, there is a unique point on the leading edge of each wave where the amplitude of the activity is a minimum. Repeated stimulation of the system during the recovery periods causes the leading edge of the waves to weaken at the point of minimal amplitude, and by $t = t_{96} = 33.6$ a small break occurs at this point. Figure 10 illustrates the evolution of the solution when $t > t_{96}$. The asymmetry in the coupling causes one of the rotors to rotate slightly faster than the other, and eventually this prevents the two rotors from intersecting, and the solution evolves into a spiral wave which fills the entire two-dimensional region (Figure 10(d)). Chu, Milton, and Cowan [9] have numerically demonstrated that a sequence of timed stimuli produces a spiral wave in a discrete integrate-and-fire network.

7. Conclusions.

In this paper we analyzed the dynamic behavior of a system of integro-differential equations that models the activity of excitatory neurons on large-scale spatially extended domains. The independent variables represent the activity level of a population of excitatory neurons with long-range connections (u) and recovery (v). We considered positive

connection functions and a Heaviside firing rate. Much of our focus has been on understanding the mechanisms responsible for the formation of traveling wave solutions. Such waves have been observed in feline cortex [4, 5, 6], in the brain of freely moving mice [18], and also in tangential and coronal brain slice experiments [30, 55].

In one space dimension we assumed symmetric connections and examined the existence of traveling wave solutions in different parameter regimes such that the linearization of the system around the rest state $(u, v) = (0, 0)$ has real or complex eigenvalues. When the eigenvalues are real both wave fronts and 1-pulse traveling waves can exist. In the appendix we explain why multipulse traveling waves are not expected in the real eigenvalue case. In contrast, when the eigenvalues are complex the range of behavior is much richer. In particular, our analysis provides strong evidence for the coexistence of at least two families of wave fronts and, in addition, infinitely many families of 1-pulse waves. To our knowledge this is the first description of infinitely many families of 1-pulse traveling waves for this class of nonlocal model. Furthermore, our numerical experiments suggest that wave fronts, 1-pulse waves, and multipulse waves can coexist as stable solutions. In all cases formidable technical difficulties preclude the completion of the the final step of existence proofs. However, for a particular coupling we developed a technique that overcomes some of the difficulties which arise in proving the existence of 1-pulse traveling waves in the real eigenvalue regime (see the appendix). It remains an open problem to extend our methods so that existence proofs can be completed for a wider range of parameters, and also for more general coupling and firing rate functions.

The one-dimensional results have facilitated our study of wave formation in two space dimensions. Here we considered a PDE approximation to the nonlocal model which includes both symmetric and asymmetric couplings. For a representative symmetric coupling we showed numerically how single-ring and double-ring-shaped traveling waves can form. In addition, we demonstrated a mechanism in which an appropriately timed stimulus causes a break in a ring, and subsequently the solution periodically produces outgoing traveling waves. When the coupling is asymmetric we investigated how a sequence of appropriately timed stimuli focused at one point can cause ring-shaped waves to form, and how one of the rings eventually breaks at a point of minimal activity on the leading edge of the wave. Subsequently, the stimuli cease, and the asymmetry in the coupling causes the solution to evolve into a spiral wave which fills the entire two-dimensional domain. Thus, we have demonstrated how appropriately timed stimuli can provide a natural pathway to the formation of spiral waves in a system which is initially at rest. Our results were obtained for the representative parameter set $(\epsilon, \theta, \beta) = (.1, .1, 3)$, where (i) the eigenvalues $\mu^{\pm}$ are complex and (ii) there is no spatially independent periodic solution which can act as a periodic pacemaker. In addition, all of these phenomena can be initiated at any point in the medium, and they do not depend on the presence of an external input. To our knowledge, this is the first demonstration of the periodic formation of traveling waves in the absence of an underlying spatially independent periodic solution for the class of integro-differential equations studied in this paper. The periodic formation of traveling waves is the result of the effects of large-scale connectivity and the complex eigenvalues. The analysis of the formation of spiral waves has practical significance since it has recently formed the basis for the prediction and discovery of rotating waves in the disinhibited cortex of a rat [30]. It remains an open problem to consider more general

parameter regimes, firing rates, and coupling functions, and to extend our results for the PDE system (6.3) to the full nonlocal model.

The numerical experiments described in this paper were performed when the firing rate is the Heaviside function. To test the robustness of our results we have considered more general sigmoidal-shaped firing rates of the form

$$f(u) = \frac{1}{1 + Ke^{-r(u-\theta)}}, \quad K > 0,\ r > 0, \tag{7.1}$$

and

$$f(u) = Ke^{-\frac{r}{(u-\theta)^2}} H(u-\theta), \quad K > 0,\ r > 0. \tag{7.2}$$

When the firing rate is given by (7.1) or (7.2) we find that our numerical results continue to hold when M is of moderate size and R is large (e.g., $K \approx 1$ and $R \geq 50$). We have also verified that our main results hold for more general coupling functions such as

$$w(x, y) = De^{-d(x^2+y^2)}, \quad D > 0,\ d > 0. \tag{7.3}$$

It remains an open problem to determine the maximal range of parameters, and generality of firing rate and coupling functions, over which the numerical results are valid.

It is of interest to contrast our results with [19, 20], where it is shown numerically that a breather exists when an appropriate external input is present. For some parameters the breather periodically emits traveling waves. When the external input is missing it is pointed out in [2] that "the homogeneous network does not support the existence of a breather that can act as a source of these waves." By contrast, our results indicate that the periodic production of traveling waves can occur without any persistent external input.

Our theoretical results might have important implications for experimental and clinical neurophysiology. In particular, our finding that the dynamics of the system undergo a qualitative transition when the eigenvalues of the linearization of the system around the rest state become complex offers a plausible explanation of trailing-end instabilities observed in cortical experiments [43, 45]. Further explanation of experimentally observed variability in the trailing end of cortical waves might be provided by our finding of the coexistence of entire families of traveling wave solutions. The unpredictable behavior of the trailing end could be caused by the solutions switching from one member of the family to another. A possible biophysical mechanism of such switching may involve a variable neurohormonal concentration affecting neuronal recovery and strength of intercellular connections [30].

Our observation that a bifurcation of the system behavior occurs at the critical value $\beta = \beta_*$ also has important practical correlates. It predicts that by lowering β below β_*, or by pushing the system above β_*, one can qualitatively change the system behavior and obtain a broad range of dynamical phenomena. One experimental example of such macrobehavior is an evoked response, which might persist long after the stimulus [45]. Understanding the cellular mechanisms responsible for such important functional changes in neuronal networks requires further study.

8. Appendix. In sections 3 and 4 we showed how technical difficulties can arise in completing the proof of existence of wave fronts and 1-pulse traveling waves. Here we focus on 1-pulse waves and develop a comparison argument which allows us to resolve some of the technical issues when $\mu^{\pm}$ are real and $w(x) = \frac{1}{2}e^{-|x|}$.

Recall that 1-pulse traveling waves satisfy the following conditions:

$$\begin{cases} U(z) < \theta \ \ \forall z < 0 \ \text{ and } \ (U(z), U'(z)) \to (0,0) \text{ as } z \to -\infty, \\ U(0) = U(a) = \theta, \text{ and } z = 0,\ z = a \text{ are the only solutions of } U(z) = \theta, \\ (U(z), U'(z)) \to (0,0) \text{ as } z \to \infty. \end{cases} \tag{8.1}$$

In section 4 we showed how difficulties can arise in proving the second condition, that

$$z = 0 \text{ and } z = a \text{ are the only solutions of } U(z) = \theta. \tag{8.2}$$

It follows from (8.1) that $U(z) < \theta \ \forall z < 0$. Our goal here is to develop a technique which allows us to prove that $U(z) < \theta \ \forall z > a$ when $\mu^{\pm}$ are real. It remains an open problem to extend our methods and show that $U(z) > \theta$ on the entire interval $(0, a)$.

To understand how real eigenvalues arise, recall from section 2 that the linearization of (2.2) around the rest state $U = 0$ is the homogeneous equation

$$c^2 H'' + c(1+\epsilon)H' + \epsilon(\beta+1)H = 0. \tag{8.3}$$

When $c > 0$ the eigenvalues $\mu^{\pm}$ associated with (8.3) are

$$\mu^{\pm} = \frac{-(\epsilon+1) \pm \sqrt{(\epsilon-1)^2 - 4\beta\epsilon}}{2c}. \tag{8.4}$$

We assume that $\epsilon > 0$, $c > 0$, and $\beta > 0$. It follows from (8.4) that

$$\mu^{\pm} \text{ are real} \iff \beta \le \beta_* = \frac{(\epsilon-1)^2}{4\epsilon}. \tag{8.5}$$

Note that $\mu^{\pm} < 0$ when $0 < \beta \le \beta_*$. For reasons which will be made clear below, we restrict our study to the parameter regime $c \ge \epsilon$. This is reasonable since our numerical experiments show that stable waves have relatively large wave speeds. Resolution of the case $0 < c < \epsilon$ remains open.

Theorem 8.1. *Let $\epsilon > 0$ and $0 < \beta \le \beta_*$. Suppose that there are values $c \ge \epsilon$, $a > 0$, and a corresponding solution U of (2.2) such that*

$$\begin{cases} U(z) < \theta \ \ \forall z \le 0 \ \text{ and } \ (U(z), U'(z)) \to (0,0) \text{ as } z \to -\infty, \\ U(0) = U(a) = \theta \ \text{ and } \ U'(a) < 0. \end{cases} \tag{8.6}$$

Then

$$U(z) < \theta \ \ \forall z > a \ \text{ and } \ (U(z), U'(z)) \to (0,0) \text{ as } z \to \infty. \tag{8.7}$$

Remarks. (i) It follows from (8.7) that if $c \ge \epsilon$, then $U(z)$ cannot exceed $U = \theta$ anywhere on (a, ∞). Thus, since multipulse waves must exceed $U = \theta$ more than twice, they cannot exist

when $\mu^{\pm}$ are real and $w(x) = \frac{1}{2}e^{-|x|}$. (ii) If $\mu^{\pm}$ are complex, then solutions oscillate around $U = 0$, and the comparison method developed below does not apply. In fact, it is precisely this oscillatory property that underlies the existence of multipulse waves in the complex eigenvalue regime (Figure 7).

Proof of Theorem 8.6. From the condition $U(a) = \theta$ and $U'(a) < 0$ in (8.6) we conclude that $U(z) < \theta$ on a maximal interval (a, z_{max}). It follows from (4.3)–(4.4) that $U(z)$ satisfies

$$c^2U'' + c(1+\epsilon)U' + \epsilon(\beta+1)U = .5(\epsilon - c)(e^a - 1)e^{-z} \qquad \forall z \in (a, z_{max}). \tag{8.8}$$

On (a, z_{max}) the general solution of (8.8) is

$$U_5 = m_5 e^{\mu^+ z} + m_6 e^{\mu^- z} + \frac{.5(\epsilon - c)(e^a - 1)}{c^2 - c(1+\epsilon) + \epsilon(\beta+1)} e^{-z}. \tag{8.9}$$

The values m_5 and m_6 are uniquely determined by the continuity conditions

$$U_5(a) = U(a) = \theta \quad \text{and} \quad U_5'(a) = U'(a) < 0. \tag{8.10}$$

We need to prove that $z_{max} = \infty$ and that $U_5(z)$ satisfies (8.7). We consider two possibilities. The first is that $U_5'(z) < 0\ \forall z \in [a, z_{max})$. Because (8.8) is a linear nonhomogeneous equation which contains no discontinuous functions, we conclude that $z_{max} = \infty$. It follows from (8.9) that $(U_5(z), U_5'(z)) \to (0, 0)$ as $z \to \infty$, as required.

The second possibility is that there is a $\hat{z} > a$ such that

$$U_5'(z) < 0 \ \ \forall z \in [a, \hat{z}) \quad \text{and} \quad U_5'(\hat{z}) = 0. \tag{8.11}$$

Then $U_5''(\hat{z}) \geq 0$. From this, (8.8), and the restriction $c \geq \epsilon$, it follows that $U_5(\hat{z}) < 0$. Therefore, there is a value $z_1 \in (a, \hat{z})$ such that

$$U_5(z_1) = 0 \quad \text{and} \quad U_5'(z) < 0 \ \ \forall z \in [z_1, \hat{z}). \tag{8.12}$$

Suppose that $U_5(z) < 0\ \forall z > \hat{z}$. Then (8.9) holds $\forall z > \hat{z}$, from which we again conclude that $(U_5(z), U_5'(z)) \to (0, 0)$ as $z \to \infty$, as required. Suppose, however, that there is a first $z_2 > \hat{z}$ such that $U_5(z_2) = 0$. Then

$$U_5(z) < 0 \ \ \forall z \in [\hat{z}, z_2), \quad U_5(z_2) = 0, \quad \text{and} \quad U_5'(z_2) \geq 0. \tag{8.13}$$

To obtain a contradiction we let H denote the solution of (8.3) such that

$$H(z_1) = U_5(z_1) = 0 \quad \text{and} \quad H'(z_1) = U_5'(z_1) < 0. \tag{8.14}$$

There are two parameter regimes to consider:

(i) $\beta < \frac{(1-\epsilon)^2}{4\epsilon}$. Then $\mu^+ - \mu^- = \frac{\sqrt{(\epsilon-1)^2 - 4\beta\epsilon}}{c} > 0$ and

$$H(z) = \frac{U_5'(z_1)}{(\mu^- - \mu_+)} \left(e^{\mu^-(z - z_1)} - e^{\mu^+(z - z_1)} \right) < 0 \quad \forall z > z_1. \tag{8.15}$$

(ii) $\beta = \frac{(1-\epsilon)^2}{4\epsilon}$. Then $\mu^+ = \mu^- = -\frac{\epsilon+1}{2c} < 0$ and

$$H(z) = U_5'(z_1)(z - z_1)e^{\mu^+(z-z_1)} < 0 \quad \forall z > z_1. \tag{8.16}$$

We assume that (i) holds. The details for (ii) are the same and are omitted.

It follows from (8.3), (8.8), and (8.14) that

$$\left(\left[U_5'(z)H(z) - H'(z)U_5(z)\right] e^{\frac{(1+\epsilon)}{c}z}\right)' = H(z)\frac{.5(\epsilon - c)(e^a - 1)}{c^2}e^{-z}e^{\frac{(1+\epsilon)}{c}z}, \quad z_1 < z < z_2, \tag{8.17}$$

and

$$U_5'(z_1)H(z_1) - H'(z_1)U_5(z_1) = 0. \tag{8.18}$$

The first case is that $c > \epsilon$. Then (8.15) implies that the right-hand side of (8.17) is strictly positive. Integrating (8.17)–(8.18) from $z = z_1$ to $z = z_2$ gives

$$\left[U_5'(z_2)H(z_2) - H'(z_2)U_5(z_2)\right] e^{\frac{(1+\epsilon)}{c}z_2} > 0. \tag{8.19}$$

However, (8.13) and (8.15) imply that

$$U_5'(z_2)H(z_2) - H'(z_2)U_5(z_2) = U_5'(z_2)H(z_2) \leq 0, \tag{8.20}$$

contradicting (8.19).

The second case is that $c = \epsilon$. Then (8.17) reduces to

$$\left(\left[U_5'(z)H(z) - H'(z)U_5(z)\right] e^{\frac{(1+\epsilon)}{c}z}\right)' = 0 \quad \forall z \in [z_1, z_2]. \tag{8.21}$$

It follows from (8.18) and an integration of (8.21) from z_1 to z that

$$U_5'(z)H(z) - H'(z)U_5(z) = 0 \quad \forall z \in [z_1, z_2]. \tag{8.22}$$

Dividing (8.21) by $H^2(z)$ gives

$$\left(\frac{U_5(z)}{H(z)}\right)' = 0 \quad \forall z \in [z_1, z_2]. \tag{8.23}$$

Finally, we integrate (8.23) from z_1 to z and obtain

$$\frac{U_5(z)}{H(z)} = 1 \quad \forall z \in [z_1, z_2]. \tag{8.24}$$

At $z = z_2$ this gives $U_5(z_2) = H(z_2) < 0$, contradicting the definition of z_2. This completes the proof.

Acknowledgments. The authors thank the referees for providing valuable comments and insights. We also thank Jack Cowan for useful discussions, and especially for pointing out the early papers of Burns and his colleagues.

REFERENCES

[1] S. Amari, *Dynamics of pattern formation in lateral inhibition type neural fields*, Biolog. Cybernet., 27 (1977), pp. 77–87.
[2] P. C. Bressloff and S. E. Folias, *Front bifurcations in an excitatory neural network*, SIAM J. Appl. Math., 65 (2004), pp. 131–151.
[3] E. Brown, J. Gao, P. Holmes, R. Bogacz, M. Gilzenrat, and J. D. Cohen, *Simple neural networks that optimize decisions*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (2005), pp. 803–826.
[4] B. D. Burns, *Some properties of the cat's isolated cerebral cortex*, J. Physiol., 111 (1950), pp. 50–68.
[5] B. D. Burns, *Some properties of the isolated cerebral cortex in the unanaesthetized cat*, J. Physiol., 112 (1951), pp. 156–175.
[6] B. D. Burns and B. Grafstein, *The function and structure of some neurones in the cat's cerebral cortex*, J. Physiol., 118 (1952), pp. 412–433.
[7] G. Buzsaki and A. Draguhn, *Neuronal oscillations in cortical networks*, Science, 304 (2004), pp. 1926–1929.
[8] R. D. Chervin, P. A. Pierce, and B. W. Connors, *Periodicity and directionality in the propagation of epileptiform discharges across neocortex*, J. Neurophys., 60 (1988), pp. 1695–1713.
[9] P. H. Chu, J. Milton, and J. D. Cowan, *Connectivity and the dynamics of integrate and fire neural networks*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 4 (1994), pp. 237–243.
[10] B. W. Connors and Y. Amati, *Generation of epileptiform discharge by local circuits of neocortex*, in Epilepsy: Models, Mechanisms and Concepts, P. A. Schwartkroin, ed., Cambridge University Press, Cambridge, UK, 1993, pp. 388–423.
[11] S. Coombes, *Waves and bumps in neural field theories*, Biolog. Cybernet., 93 (2005), pp. 91–108.
[12] S. Coombes and M. R. Owen, *Evans functions for integral neural field equations with Heaviside firing rate function*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 574–600.
[13] S. Coombes and M. R. Owen, *Bumps, breathers, and waves in a neural network with spike frequency adaptation*, Phys. Rev. Lett., 94 (2005), paper 148102.
[14] J. D. Drover and B. Ermentrout, *Nonlinear coupling near a degenerate Hopf (Bautin) bifurcation*, SIAM J. Appl. Math., 63 (2003), pp. 1627–1647.
[15] J. S. Ebersol and J. Milton, *The electroencephalogram (EEG): A measure of neural synchrony*, in Epilepsy as a Dynamic Disease, Biol. Med. Phys. Ser., Springer, New York, 2003, Chapter 5.
[16] J. C. Eccles, *The Understanding of the Brain*, McGraw–Hill, New York, 1973.
[17] G. B. Ermentrout and J. B. McLeod, *Existence and uniqueness of traveling waves for a neural network*, Proc. Roy. Soc. Edinburgh Sect. A, 123 (1993), pp. 461–478.
[18] I. Ferezou, S. Bolea, and C. Petersen, *Visualizing the cortical representation of whisker touch: Voltage sensitive dye imaging in freely moving mice*, Neuron, 50 (2006), pp. 617–629.
[19] S. Folias and P. Bressloff, *Breathers in two dimensional excitable neural media*, Phys. Rev. Lett., 95 (2004), paper 208107.
[20] S. E. Folias and P. C. Bressloff, *Breathing pulses in an excitatory neural network*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 378–407.
[21] J. Glanz, *Mastering the nonlinear brain*, Science, 277 (1997), pp. 1758–1760.
[22] D. Golomb, *Models of neuronal transient synchrony during propagation of activity through neocortical circuitry*, J. Neurophys., 79 (1998), pp. 1–12.
[23] D. Golomb and Y. Amitai, *Propagating neuronal discharges in neocortical slices: Computational and experimental study*, J. Neurophys., 78 (1997), pp. 1199–1211.
[24] Y. Guo and C. C. Chow, *Existence and stability of standing pulses in neural networks:* I. *Existence*, SIAM J. Appl. Dyn. Syst., 4 (2005), pp. 217–248.
[25] Y. Guo and C. C. Chow, *Existence and stability of standing pulses in neural networks:* II. *Stability*, SIAM J. Appl. Dyn. Syst., 4 (2005), pp. 249–281.

[26] B. Gutkin, D. Pinto, and B. Ermentrout, *Mathematical neuroscience: From neurons to circuits to systems*, J. Physiol. Paris, 97 (2003), pp. 209–219.
[27] S. P. Hastings, *Single and multiple pulse waves for the FitzHugh–Nagumo equations*, SIAM J. Appl. Math., 42 (1982), pp. 247–260.
[28] J. A. Hobson, *Dreaming: An Introduction to the Science of Sleep*, Oxford University Press, London, 2004.
[29] J. A. Hobson and R. W. McCarley, *The brain as a dream state generator. An activation-synthesis of the dream process*, Amer. J. Psychiatry, 134 (1977), pp. 1335–1348.
[30] X. Huang, W. C. Troy, Q. Yang, H. Ma, C. Laing, S. Schiff, and J. Y. Wu, *Spiral waves in disinhibited mammalian cortex*, J. Neurosci., 24 (2004), pp. 9897–9902.
[31] M. A. P. Idiart and L. F. Abbott, *Propagation of excitation in neural network models*, Network, 4 (1993), pp. 285–294.
[32] D. Kleinfield, K. R. Delaney, M. S. Fee, J. A. Flores, D. W. Tank, and A. Galperin, *Dynamics of propagating waves in the olfactory network of a terrestrial mollusk: An electrical and optical study*, J. Neurophys., 72 (1994), pp. 1402–1419.
[33] N. Kopell and B. Ermentrout, *Chemical and electrical synapses perform complementary roles in the synchronization of interneuronal networks*, Proc. Natl. Acad. Sci. USA, 101 (2004), pp. 15482–15487.
[34] C. R. Laing and W. C. Troy, *PDE methods in nonlocal models*, SIAM J. Appl. Dyn. Syst., 2 (2003), pp. 487–516.
[35] Y. W. Lam, L. B. Cohen, M. Wachowiak, and M. R. Zochowski, *Odors elicit three different oscillations in the turtle olfactory bulb*, J. Neurosci., 20 (2000), pp. 749–762.
[36] R. Miles, R. D. Traub, and R. K. Wong, *Spread of synchronous firing in longitudinal slices from the CA3 region of the hippocampus*, J. Neurophys., 60 (1988), pp. 1481–1496.
[37] J. Milton, T. Mundel, U. an der Heiden, J. Sprire, and J. Cowan, *Traveling activity waves*, in Handbook of Brain Theory and Neural Networks, MIT Press, Cambridge, MA, 1994, pp. 994–996.
[38] J. Milton and P. Jung, eds., *Epilepsy as a Dynamic Disease*, Biol. Med. Phys. Ser., Springer, New York, 2003.
[39] C. Peskin, *Mathematical Aspects of Heart Physiology*, Courant Institute of Mathematical Sciences Publication, New York, 1975.
[40] D. J. Pinto and G. B. Ermentrout, *Spatially structured activity in synaptically coupled neuronal networks:* I. *Traveling fronts and pulses*, SIAM J. Appl. Math., 62 (2001), pp. 206–225.
[41] D. J. Pinto and G. B. Ermentrout, *Spatially structured activity in synaptically coupled neuronal networks:* II. *Lateral inhibition and standing pulses*, SIAM J. Appl. Math., 62 (2001), pp. 226–243.
[42] D. J. Pinto, R. K. Jackson, and C. E. Wayne, *Existence and stability of traveling pulses in a continuous neuronal network*, SIAM J. Appl. Dyn. Syst., 4 (2005), pp. 954–984.
[43] D. J. Pinto, S. A. Patrick, H. W. Huang, and B. Connors, *Initiation, propagation and termination of epileptiform activity in rodent neocortex in vitro involve distinct mechanisms*, J. Neurosci., 25 (2005), pp. 8131–8140.
[44] D. Pinto and W. C. Troy, *The Effects of Asymmetric Coupling on Traveling Waves in Neocortex*, in preparation.
[45] J. C. Prechtl, L. B. Cohen, B. Pasaram, P. P. Mitra, and D. Kleinfeld, *Visual stimuli induce waves of electrical activity in turtle cortex*, Proc. Natl. Acad. Sci. USA, 94 (1997), pp. 7621–7626.
[46] K. Richardson, S. J. Schiff, and B. J. Gluckman, *Control of traveling waves in mammalian cortex*, Phys. Rev. Lett., 94 (2005), paper 028103.
[47] A. Rosenbluth and W. B. Cannon, *Cortical responses to electrical stimulation*, A. J. Physiol., 135 (1942), pp. 690–741.
[48] B. Sandstede, *Evans functions and nonlinear stability of travelling waves in neuronal network models*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., to appear.
[49] B. Schechter, *How the brain gets its rhythm*, Science, 274 (1996), p. 339–340.
[50] J. Schofflen, R. Oostenveld, and P. Fries, *Neuronal coherence as a mechanism of effective corticospinal interaction*, Science, 308 (2003), pp. 111–113.
[51] I. A. Shevlev, E. N. Tsicalov, A. M. Gorbach, K. P. Budko, and G. A. Sharaev, *Temperature tomography of the brain cortex: Thermoencephalography*, J. Neurosci. Methods, 46 (1992), pp. 49–57.

[52] V. Shusterman, P. J. Jannetta, B. Aysin, A. Beigel, M. Glukhovskoy, and I. Usiene, *Direct mechanical stimulation of brainstem modulates cardiac rhythm and repolarization in humans*, J. Electrocardiol., 35 (2002), pp. 247–256.
[53] S. Strogatz, *SYNC*, Hyperion Books, New York, 2003.
[54] H. R. Wilson and J. D. Cowan, *A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue*, Kybernetik, 13 (1973), pp. 55–80.
[55] J. Y. Wu, L. Guan, and Y. Tsau, *Propagating activation during oscillations and evoked responses in neocortical slices*, J. Neurosci., 19 (1999), pp. 5005–5015.
[56] L. Zhang, *On stability of traveling wave solutions in synaptically coupled neuronal networks*, Differential Integral Equations, 16 (2003), pp. 513–536.

SIAM J. APPLIED DYNAMICAL SYSTEMS
Vol. 6, No. 2, pp. 293–318

Uniform Attractors for Nonautonomous Wave Equations with Nonlinear Damping*

Chunyou Sun†, Daomin Cao‡, and Jinqiao Duan§

Abstract. We consider dynamical behavior of nonautonomous wave-type evolutionary equations with nonlinear damping, critical nonlinearity, and time-dependent external forcing which is translation bounded but not translation compact (i.e., external forcing is not necessarily time-periodic, quasi-periodic, or almost periodic). A sufficient and necessary condition for the existence of uniform attractors is established using the concept of uniform asymptotic compactness. A new method for verifying uniform asymptotic compactness is devised. The required compactness for the existence of uniform attractors is then fulfilled by some new a priori estimates for concrete wave-type equations arising from applications. The structure of uniform attractors is obtained by constructing a skew product flow on the extended phase space.

Key words. nonautonomous systems, wave equation, nonlinear damping, critical exponent, uniform attractor, norm-to-weak continuous process

AMS subject classifications. 35L05, 35B40, 35B41

DOI. 10.1137/060663805

1. Introduction.

We consider the following nonautonomous wave equations with nonlinear damping on a bounded domain Ω in $\mathbb{R}^3$ with smooth boundary:

$$u_{tt} + h(u_t) - \Delta u + f(u) = g(x, t), \quad x \in \Omega, \tag{1.1}$$

under the boundary condition

$$u|_{\partial\Omega} = 0, \tag{1.2}$$

and initial conditions

$$u(x, 0) = u_0(x), \quad u_t(x, 0) = v_0(x). \tag{1.3}$$

Here h is the nonlinear damping function, f is the nonlinearity, and g is a given external time-dependent forcing.

Equation (1.1) arises as an evolutionary mathematical model in various systems—for example, (i) modeling a continuous Josephson junction with specific h, g, and f; (ii) modeling

*Received by the editors June 27, 2006; accepted for publication (in revised form) by G. Wayne December 20, 2006; published electronically April 17, 2007. This work was partly supported by the NSF grant 0620539, the NSFC (grant 10601021), the key project of NSFC (grant 10631030), the Knowledge Innovation Funds of CAS in China, and the Outstanding Overseas Chinese Scholars Fund of the Chinese Academy of Sciences.
http://www.siam.org/journals/siads/6-2/66380.html

†Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, 100080, China, and School of Mathematics and Statistics, Lanzhou University, Lanzhou, 730000, China (cysun@amss.ac.cn).
‡Institute of Applied Mathematics, Chinese Academy of Sciences, Beijing, 100080, China (dmcao@amt.ac.cn).
§Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616 (duan@iit.edu).

a hybrid system of nonlinear waves and nerve conduct; and (iii) modeling a phenomenon in quantum mechanics. A relevant physical issue is to investigate the asymptotic dynamical behavior of these mathematical models. See [31, 36, 39].

For the special, autonomous case of (1.1), i.e., when g does not depend on time t explicitly, the solution operator defines a flow or semigroup. The asymptotic behaviors of the solutions have been studied extensively by using the concept of global attractors; see, for example, [1, 2, 3, 11, 21, 25, 34, 39] for the linear damping case, and [13, 14, 15, 19, 20, 37] for the nonlinear damping case.

In the general case of the nonautonomous system (1.1), the solution operator does *not* define a flow or semigroup but a *process*; see sections 2 and 5 below. A proper extension of the notion of a global attractor for semigroups to the case of processes is the so-called uniform attractor (see, e.g., [23, 9, 11]). About the basic concepts of nonautonomous dynamical systems, uniform attractors, and processes, we refer the reader to [23, 9, 11] or section 2 below for more details. See also [4, 6, 7, 12, 18, 32] for other methods or frameworks considering the nonautonomous dynamics.

The basic assumptions about nonlinear damping h and nonlinearity f are

$$h \in C^1(\mathbb{R}), \quad h(0) = 0, \quad h \text{ is strictly increasing}, \tag{1.4}$$

$$\liminf_{|s|\to\infty} h'(s) > 0, \tag{1.5}$$

$$|h(s)| \le C_1(1 + |s|^p), \tag{1.6}$$

where $p \in [1, 5)$; and $f \in C^1(\mathbb{R})$ and satisfies

$$|f'(s)| \le C_2(1 + |s|^q), \tag{1.7}$$

$$\liminf_{|s|\to\infty} \frac{f(s)}{s} > -\lambda_1, \tag{1.8}$$

where $0 \leqslant q \leqslant 2$ and λ_1 is the first eigenvalue of $-\Delta$ in $H_0^1(\Omega)$. These assumptions are similar to those for an autonomous system and come from [13, 14, 19].

In this paper, we consider the nonautonomous system (1.1)–(1.3) via the uniform attractors of the corresponding family of processes $\{U_\sigma(t, \tau)\}$, $\sigma \in \Sigma$; Σ is the so-called symbol space and is explained later, especially by (i) the nonlinear damping (i.e., h is a nonlinear function), (ii) the nonlinearity $f(u)$ has critical exponent ($q = 2$), and (iii) the external forcing $g(x, t)$ is *not* translation compact in $L^2_{loc}(\mathbb{R}; L^2(\Omega))$.

In Chepyzhov and Vishik [11], for the linear damping case $h(v) = kv$ with a constant $k > 0$ and $q < 2$ (subcritical), for system (1.1)–(1.3), the authors obtained the existence of a bounded uniformly absorbing set if g is translation bounded and the existence of a uniform attractor when g is translation compact (e.g., when g is time-periodic, quasi-periodic, or almost periodic). Under the assumptions that g and $\partial_t g$ are both in the space of bounded continuous functions $C_b(\mathbb{R}, L^2(\Omega))$ and h satisfies the growth bounds $0 < \alpha \leqslant h'(s) \leqslant \beta < \infty$ for some constants α and β, Zhou and Wang [42] have proved the existence of kernel sections and obtained the estimation of the Hausdorff dimension of the kernel sections.

For the *existence* of uniform attractors, as in the autonomous case, some kind of compactness of the family of processes is a key ingredient. The corresponding compactness assumption in [23, 9, 11] is that the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, has a compact uniformly absorbing set. The number $q = 2$ is called the critical exponent, since the nonlinearity f is not compact in this case, which is an essential difficulty in studying the asymptotic behavior even for the autonomous case [1, 2, 3, 13, 14, 19, 34, 37].

About the case of $1 < p < 5$ for the nonlinear damping exponent p, as mentioned in Haraux [22], even for the bounded dissipation, it becomes much more difficult when g depends on t, and the characterizations of dynamics for this case are unknown to the authors. Moreover, the nonlinearity of h also brings some difficulties in proving the compactness. For example, for the autonomous linearly damped wave equations, Ball [3] proposed a very nice energy method to verify the necessary asymptotic compactness; see also [29, 32, 33, 39] for this method in the contexts of both the autonomous and nonautonomous cases. However, for our problem, due to the nonlinear damping, it seems difficult to directly apply the method of [3, 32].

The purpose of this paper is to obtain the existence and structure of the compact (in the norm topology of $H_0^1 \times L^2$), uniform attractor when the external forcing $g(x,t)$ is *not* translation compact in $L^2_{loc}(\mathbb{R}; L^2(\Omega))$. For the existence of uniform attractors, a main approach in [9, 11] is to construct skew product flow on the extended phase space $X \times \Sigma$. They require that the symbol space has some compactness so that the skew product flow has some corresponding compactness, i.e., the concept of translation compact functions (e.g., see [9, 10, 11]). Consequently, the compact uniform attractors are obtained for the systems with symbols of compact hulls, and the weakly compact uniform attractors for the systems with symbols of weakly compact hulls. However, there are some results which show that one can obtain the compact uniform attractors for the system with translation noncompact external forcing: by generalizing the methods in [30], the authors in [28] obtain the existence of a uniform attractor for the two-dimensional Navier–Stokes equation in a bounded domain with a kind of translation noncompact external forcing; in Zelik [40], by use of a bootstrap argument together with a sharp use of Gronwall-type lemmas, when $h(v) \equiv kv$ and $g, \partial_t g \in L^\infty(\mathbb{R}; L^2(\Omega))$, the author obtains some regularity estimates for the solutions of (1.1), which implies naturally the existence of an uniform attractor; see also the results in Chepyzhov and Vishik [8].

Furthermore, we consider the *structure* of the uniform attractor by investigating the kernel sections of a process (see [9, 11] for more details).

Here, for system (1.1)–(1.3), we further assume that

$$g(\cdot,t) \in L^\infty(\mathbb{R}; L^2(\Omega)) \tag{1.9}$$

and

$$\partial_t g \in L^r_b(\mathbb{R}; L^r(\Omega)) \text{ with } r > \frac{6}{5}, \tag{1.10}$$

where the space $L^r_b(\mathbb{R}; L^r(\Omega))$ of "translation bounded" functions will be defined in the beginning of the next section. Roughly speaking, the two conditions (1.9)–(1.10) mean that the external forcing g is bounded in time and its time-derivative $\partial_t g$ is translation bounded. It is clear that a function g satisfies (1.9) and (1.10) does not need to be translation compact in $L^2_{loc}(\mathbb{R}; L^2(\Omega))$. Moreover, we remark that the technical hypothesis (1.9) is mainly for the

existence of a bounded uniformly absorbing set, and (1.10) is for the uniform asymptotic compactness.

Note that for uniform dissipation, using a method similar to that in the proof of Theorem 5.3, we can replace (1.9) by $g \in W_b^{1,1}(\mathbb{R}; H^{-1})$, and, in order to obtain the asymptotic compactness, we need only the weaker assumption $g \in W_b^{1,1+\delta}(\mathbb{R}; H^{-1+\delta})$ with arbitrary $\delta > 0$. Please see more on this in section 5.

It is interesting to note that if (1.10) is replaced by the assumption that g is translation compact (e.g., g is a periodic, quasi-periodic, or almost periodic function in $L^2_{loc}(\mathbb{R}; L^2(\Omega))$), then our result on uniform attractors (see Theorem 5.12) in section 5 still holds, but the proof can be largely simplified; see (5.31)–(5.32), Remark 5.10, and Remark 5.13 below. At the same time, the method in [28] cannot be applied to our problem as (1.1) is a hyperbolic-type equation.

This paper is organized as follows. After introducing some basic materials in section 2, we first present a criterion for the existence of a compact uniform attractor in section 3, using the concept of uniform asymptotic compactness (different from the corresponding concept in [9, 11]) which is introduced by Moise, Rosa, and Wang in [32] for the family of semiprocesses. We apply this concept to the family of processes; see Definition 3.1 and Theorem 3.4. Then, we investigate the *structure* of the uniform attractor via kernel sections of a process. In fact, we present results on uniform attractors and their decompositions into kernel sections for norm-to-weak continuous processes (see Definition 3.5, Theorem 3.8, and Theorem 3.10). Note that the norm-to-weak continuity here is weaker than the usual *norm-to-norm* and *weak-to-weak* continuities.

In section 4, partially inspired by the results and ideas in [13, 14, 15, 24], we present a simple method for verifying the uniform asymptotic compactness for processes generated by wave-type evolutionary equations like (1.1); see Theorem 4.2.

In section 5, as applications to concrete wave-type evolutionary equations, we first prove the existence of compact uniform (w.r.t. $\sigma \in \Sigma$) attractors when the external forcing $g_0 = \sigma_0$ satisfies (1.9) and (1.10); see Theorem 5.12. Then we show that the uniform attractor w.r.t. the initial time τ of a process $\{U_{\sigma_0}(t,\tau)\}$ coincides with the uniform attractor w.r.t. the symbol $\sigma \in \Sigma'$ when the external forcing $g_0 = \sigma_0 \in W_b^{1,1+\delta}(\mathbb{R}; H^{-1+\delta}(\Omega))$ for arbitrary $\delta > 0$, and we further decompose this uniform attractor into kernel sections; see Theorem 5.15. Finally, we conclude the paper with some remarks in section 6.

2. Preliminaries. In this section, we recall some basic concepts about nonautonomous systems. We refer to [23, 9, 11] and the references therein for more details.

The space of translation bounded functions in $L^r_{loc}(\mathbb{R}; L^k(\Omega))$, with $r, k \geqslant 1$, is

$$L^r_b(\mathbb{R}; L^k(\Omega)) = \left\{ g \in L^r_{loc}(\mathbb{R}; L^k(\Omega)) : \sup_{t\in\mathbb{R}} \int_t^{t+1} \left(\int_\Omega |g(x,s)|^k dx \right)^{\frac{r}{k}} ds < \infty \right\}.$$

The space of translation compact functions in $L^2_{loc}(\mathbb{R}; L^2(\Omega))$ is

$$L^2_c(\mathbb{R}; L^2(\Omega)) = \Big\{ g \in L^2_{loc}(\mathbb{R}; L^2(\Omega)) : \text{For any interval } [t_1, t_2] \subset \mathbb{R},$$
$$\{g(x, h+s) : h \in \mathbb{R}\}|_{[t_1,t_2]} \text{ is precompact in } L^2(t_1, t_2; L^2(\Omega)) \Big\}.$$

Let X be a complete metric space and Σ be a parameter set.

The operators $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, are said to be a family of processes in X with symbol space Σ if for any $\sigma \in \Sigma$

$$U_\sigma(t,s) \circ U_\sigma(s,\tau) = U_\sigma(t,\tau) \quad \forall\ t \geqslant s \geqslant \tau,\ \tau \in \mathbb{R}, \tag{2.1}$$

$$U_\sigma(\tau,\tau) = \mathrm{Id}\ \ (\text{identity}) \quad \forall\ \tau \in \mathbb{R}. \tag{2.2}$$

Let $\{T(s)\}_{s\geqslant 0}$ be the translation semigroup on Σ; we say that a family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, satisfies the translation identity if

$$U_\sigma(t+s,\tau+s) = U_{T(s)\sigma}(t,\tau) \quad \forall\ \sigma \in \Sigma,\ t \geqslant \tau,\ \tau \in \mathbb{R},\ s \geqslant 0, \tag{2.3}$$

$$T(s)\Sigma = \Sigma \quad \forall\ s \geqslant 0. \tag{2.4}$$

By $\mathcal{B}(X)$ we denote the collection of the bounded sets of X, and $\mathbb{R}^\tau = \{t \in \mathbb{R},\ t \geqslant \tau\}$.

Definition 2.1 (see [11]). *A bounded set $B_0 \in \mathcal{B}(X)$ is said to be a bounded uniformly (w.r.t. $\sigma \in \Sigma$) absorbing set for $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, if for any $\tau \in \mathbb{R}$ and $B \in \mathcal{B}(X)$ there exists $T_0 = T_0(B,\tau)$ such that $\bigcup_{\sigma\in\Sigma} U_\sigma(t;\tau)B \subset B_0$ for all $t \geqslant T_0$.*

Definition 2.2 (see [11]). *A set $A \subset X$ is said to be uniformly (w.r.t. $\sigma \in \Sigma$) attracting for the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, if for any fixed $\tau \in \mathbb{R}$ and any $B \in \mathcal{B}(X)$*

$$\lim_{t\to+\infty} \left(\sup_{\sigma\in\Sigma} dist(U_\sigma(t;\tau)B;\ A) \right) = 0;$$

here $dist(\cdot,\cdot)$ is the usual Hausdorff semidistance in X between two sets.

In particular, a closed uniformly attracting set $\mathscr{A}_\Sigma$ is said to be the uniform (w.r.t. $\sigma \in \Sigma$) attractor of the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, if it is contained in any closed uniformly attracting set (minimality property).

Obviously, if the uniform (w.r.t. $\sigma \in \Sigma$) attractor exists, it is unique.

In order to obtain the structure as well as the existence of the uniform attractor, under the condition (2.3)–(2.4), the authors in [11] construct the skew product flow in $X \times \Sigma$,

$$S(t)(u,\sigma) = (U_\sigma(t,0)u,\ T(t)\sigma), \quad t \geqslant 0,\ (u,\sigma) \in X \times \Sigma, \tag{2.5}$$

and $\{S(t)\}_{t\geqslant 0}$ forms a semigroup on $X \times \Sigma$.

3. Abstract results.

3.1. Existence of the uniform attractor. In this subsection, we present a criterion for the existence of a compact uniform attractor using the concept of uniform (w.r.t. $\sigma \in \Sigma$) asymptotical compactness, which is different from the corresponding concept in [9, 11], and it is introduced in Moise, Rosa, and Wang [32] for a family of semiprocesses; see [26] for autonomous systems. Now, we use this concept for the family of processes.

Definition 3.1 (see [32]). *A family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, on a complete metric space X is said to be uniformly (w.r.t. $\sigma \in \Sigma$) asymptotically compact if and only if for any fixed $\tau \in \mathbb{R}$, a bounded sequence $\{u_n\}_{n=1}^\infty \subset X$, $\{\sigma_n\}_{n=1}^\infty \subset \Sigma$, and any $\{t_n\}_{n=1}^\infty \subset \mathbb{R}^\tau$ with $t_n \to \infty$ as $n \to \infty$, the sequence $\{U_{\sigma_n}(t_n,\tau)u_n\}_{n=1}^\infty$ is precompact in X.*

Similarly, define the uniform ω-limit set of $B \subset X$ at initial time τ by

$$\omega_{\tau,\Sigma}(B) = \bigcap_{t\geqslant\tau} \overline{\bigcup_{\sigma\in\Sigma} \bigcup_{s\geqslant t} U_\sigma(s,\tau)B}, \tag{3.1}$$

where $\overline{A}$ means the closure of A in X.

Then, we have the following characterizations for the uniform ω-limit set (see [11, 32]).

Lemma 3.2. *For any bounded set $B \subset \mathcal{B}(X)$, $u \in \omega_{\tau,\Sigma}(B)$ if and only if there exist $\{u_n\}_{n=1}^\infty \subset B$, $\{\sigma_n\}_{n=1}^\infty \subset \Sigma$, and $\{t_n\}_{n=1}^\infty \subset \mathbb{R}^\tau$ with $t_n \to \infty$ as $n \to \infty$ such that $U_{\sigma_n}(t_n,\tau)u_n \to u$.*

In the following, similar to [11, 28, 32], we give some characterizations for the uniform (w.r.t. $\sigma \in \Sigma$) asymptotically compact processes.

Lemma 3.3. *Let $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, be a family of uniform (w.r.t. $\sigma \in \Sigma$) asymptotically compact processes on a complete metric space X; then for any $\tau \in \mathbb{R}$ and any nonempty set $B \in \mathcal{B}(X)$, we have the following:*

(i) *$\omega_{\tau,\Sigma}(B)$ is nonempty and compact in X;*

(ii) *$\lim_{t\to+\infty} \sup_{\sigma\in\Sigma} dist\,(U_\sigma(t,\tau)B, \omega_{\tau,\Sigma}(B)) = 0$;*

(iii) *if Y is closed and uniformly (w.r.t. $\sigma \in \Sigma$) attracts B, then $\omega_{\tau,\Sigma}(B) \subset Y$.*

Furthermore, if $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, satisfies the translation identity (2.3)–(2.4), then

(iv) *$\omega_{\tau,\Sigma}(B) \equiv \omega_{0,\Sigma}(B)$; that is, $\omega_{\tau,\Sigma}(B)$ is independent of $\tau \in \mathbb{R}$.*

Proof. (i) For any fixed $\tau \in \mathbb{R}$, and then for any $t_n \in \mathbb{R}^\tau$, $t_n \to \infty$, $\sigma_n \in \Sigma$, and $x_n \in B$, by the definition of uniform (w.r.t. $\sigma \in \Sigma$) asymptotic compactness we know that $\{U_{\sigma_n}(t_n,\tau)x_n\}_{n=1}^\infty$ is precompact in X, and without loss of generality, we assume that

$$U_{\sigma_n}(t_n,\,\tau)x_n \to y.$$

Then by the definition of ω-limit set we know that $y \in \omega_{\tau,\,\Sigma}(B)$, which implies that $\omega_{\tau,\,\Sigma}(B)$ is nonempty.

For any $y_m \in \omega_{\tau,\,\Sigma}(B)$, $m = 1, 2, \ldots$, we will show that $\{y_m\}_{m=1}^\infty$ is precompact in X. By the definition, for each $m \in \mathbb{N}$, there exist $t_m \in \mathbb{R}^\tau$, $t_m \geqslant m$, $\sigma_m \in \Sigma$, and $x_m \in B$ such that

$$\rho(U_{\sigma_m}(t_m,\,\tau)x_m,\,y_m) \leqslant \frac{1}{m},$$

where $\rho(\cdot,\cdot)$ is the metric on X.

Therefore, by the assumption of uniform (w.r.t. $\sigma \in \Sigma$) asymptotic compactness again, we have that $\{U_{\sigma_m}(t_m,\tau)x_m\}_{m=1}^\infty$ is precompact in X, and without loss of generality, we assume that $\{U_{\sigma_m}(t_m,\tau)x_m\}_{m=1}^\infty$ is a Cauchy sequence in X. Then, from

$$\begin{aligned}
&\rho(y_n,\,y_m)\\
&\leqslant \rho(y_n,\,U_{\sigma_n}(t_n,\,\tau)x_n) + \rho(U_{\sigma_n}(t_n,\,\tau)x_n,\,U_{\sigma_m}(t_m,\,\tau)x_m) + \rho(U_{\sigma_m}(t_m,\,\tau)x_m,\,y_m)\\
&\leqslant \frac{1}{n} + \rho(U_{\sigma_n}(t_n,\,\tau)x_n,\,U_{\sigma_m}(t_m,\,\tau)x_m) + \frac{1}{m},
\end{aligned}$$

we know that $\{y_m\}_{m=1}^\infty$ is also a Cauchy sequence in X. Moreover, from the definition we obviously have that $\omega_{\tau,\,\Sigma}(B)$ is closed in X.

Hence, $\omega_{\tau,\Sigma}(B)$ is compact in X.

(ii) If (ii) is not true, then there exist $\varepsilon_0 > 0$, $\sigma_n \in \Sigma$, $x_n \in B$, and $t_n \in \mathbb{R}^\tau$ with $t_n \geqslant n$, such that

$$\operatorname{dist}(U_{\sigma_n}(t_n, \tau)x_n, \omega_{\tau,\Sigma}(B)) \geqslant \varepsilon_0, \quad n = 1, 2, \ldots.$$

However, the uniform (w.r.t. $\sigma \in \Sigma$) asymptotic compactness implies that $\{U_{\sigma_n}(t_n, \tau)x_n\}_{n=1}^{\infty}$ is precompact in X; that is, $\{U_{\sigma_n}(t_n, \tau)x_n\}_{n=1}^{\infty}$ has a convergent subsequence which converges to some point of $\omega_{\tau,\Sigma}(B)$. This is a contradiction.

(iii) For all $y \in \omega_{\tau,\Sigma}(B)$, there are $\sigma_n \in \Sigma$, $x_n \in B$, and $t_n \in \mathbb{R}^\tau$ with $t_n \to \infty$ such that $U_{\sigma_n}(t_n, \tau)x_n \to y$. From the assumption that Y uniformly attracts B, obviously, we have

$$\operatorname{dist}(U_{\sigma_n}(t_n, \tau)x_n, Y) \to 0 \quad \text{as } n \to \infty.$$

At the same time, the closeness of Y implies $y \in Y$. Hence, $\omega_{\tau,\Sigma}(B) \subset Y$.

(iv) For any fixed $\tau \in \mathbb{R}$ and $\sigma \in \Sigma$, from the translation identity (2.3) we know (e.g., see [28, 32]) that for any $\tau_0 \in \mathbb{R}$ there is a $\sigma' \in \Sigma$ such that

$$U_\sigma(t, \tau) = U_{\sigma'}(t - \tau + \tau_0, \tau_0) \quad \forall\, t \geqslant \tau.$$

Combining this with (2.4), we have that for any $t \geqslant \tau$,

$$\bigcup_{\sigma\in\Sigma}\bigcup_{s\geqslant t} U_\sigma(s, \tau)B = \bigcup_{\sigma\in\Sigma}\bigcup_{s\geqslant t} U_\sigma(s - \tau, 0)B.$$

Therefore, we have

$$\omega_{\tau,\Sigma}(B) = \bigcap_{t\geqslant\tau}\overline{\bigcup_{\sigma\in\Sigma}\bigcup_{s\geqslant t} U_\sigma(s, \tau)B} = \bigcap_{t\geqslant 0}\overline{\bigcup_{\sigma\in\Sigma}\bigcup_{s\geqslant t} U_\sigma(s, 0)B} = \omega_{0,\Sigma}(B). \qquad \blacksquare$$

Theorem 3.4. *Let X be a complete metric space and $\{U_\sigma(t, \tau)\}$, $\sigma \in \Sigma$, be a family of processes on X which satisfies the translation identity (2.3)–(2.4). Then, $\{U_\sigma(t, \tau)\}$, $\sigma \in \Sigma$, has a compact uniform (w.r.t. $\sigma \in \Sigma$) attractor $\mathscr{A}_\Sigma$ in X and satisfies*

$$\mathscr{A}_\Sigma = \omega_{0,\Sigma}(B_0) = \omega_{\tau,\Sigma}(B_0) = \bigcup_{B\in\mathcal{B}(X)} \omega_{\tau,\Sigma}(B) \quad \forall\, \tau \in \mathbb{R}$$

if and only if $\{U_\sigma(t, \tau)\}$, $\sigma \in \Sigma$,

(i) *has a bounded uniformly (w.r.t. $\sigma \in \Sigma$) absorbing set B_0; and*

(ii) *is uniformly (w.r.t. $\sigma \in \Sigma$) asymptotically compact.*

Proof. The necessity follows from the definition of uniform (w.r.t. $\sigma \in \Sigma$) attractor and the compactness of $\mathscr{A}_\Sigma$.

Now we prove the sufficiency. For any fixed $\tau \in \mathbb{R}$ and any $B \in \mathcal{B}(X)$, we know that there is a $T = T(\tau, B)$ such that

$$\bigcup_{\sigma\in\Sigma}\bigcup_{t\geqslant T} U_\sigma(t, \tau)B \subset B_0.$$

Combining this with the equivalent characterization (Proposition 3.2) of the ω-limit set, and Lemma 3.3, we have

$$\omega_{\tau,\Sigma}(B) \subset \omega_{\tau,\Sigma}(B_0) = \omega_{0,\Sigma}(B_0) \tag{3.2}$$

and $\omega_{\tau,\Sigma}(B)$; of course, $\omega_{0,\Sigma}(B_0)$ uniformly (w.r.t. $\sigma \in \Sigma$) attracts B.

Moreover, (3.2) implies that $\bigcup_{B\in\mathcal{B}(X)} \omega_{\tau,\Sigma}(B) \subset \omega_{0,\Sigma}(B_0)$, and from $B_0 \in \mathcal{B}(X)$ we obtain $\bigcup_{B\in\mathcal{B}(X)} \omega_{\tau,\Sigma}(B) = \omega_{0,\Sigma}(B_0)$.

The minimality and closeness follow immediately from (iii) of Lemma 3.3, and the compactness follows from (i) of Lemma 3.3. ■

3.2. Structure of the uniform attractor. We describe the structure of the uniform attractor by means of its kernel sections.

Hereafter, we assume that X is a Banach space with norm $\|\cdot\|_X$ and Σ is a complete metric space with metric $d(\cdot,\cdot)$.

Let $\{U(t,\tau)|t \geqslant \tau,\ \tau \in \mathbb{R}\} = \{U(t,\tau)\}$ be a process acting in a Banach space X, and let $\mathcal{K}$ be the kernel of the process $\{U(t,\tau)\}$. We recall (e.g., see [11]) that the kernel $\mathcal{K}$ consists of all bounded complete trajectories of the process, i.e.,

$$\mathcal{K} = \{u(\cdot)\ |\|u(t)\|_X \leqslant C_u,\ U(t,\tau)u(\tau) = u(t)\ \ \forall\, t \geqslant \tau,\ \tau \in \mathbb{R}\},$$

and $\mathcal{K}(s)$ denotes the kernel section at a time moment $s \in \mathbb{R}$:

$$\mathcal{K}(s) = \{u(s)\ |u(\cdot) \in \mathcal{K}\}, \quad \mathcal{K}(s) \subset X.$$

As mentioned in [11], since the invariance of the global attractor of a semigroup is replaced by the minimality in the definition of the uniform attractor of a family of processes, the existence of uniform attractor does not need any continuity for the processes. However, in order to obtain the structure of the uniform attractor, the continuity may be necessary to some extent.

3.2.1. Norm-to-weak continuous processes. In [9, 11], in order to obtain the structure of the uniform attractor, the authors assume that the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, is $(X \times \Sigma, X)$-continuous; see Theorem 5.1 in Chapter IV of [11].

Now, as noticed in [41], in order to obtain the invariance of the global attractor of a semigroup for an autonomous system, we need only the norm-to-weak continuity. In this part, we will generalize these results to nonautonomous systems.

Definition 3.5. *A family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, is said to be norm-to-weak continuous if, for any fixed t and $\tau \in \mathbb{R}$ with $t \geqslant \tau$ and for any $\{x_n\} \subset X$ and $\{\sigma_n\} \subset \Sigma$, we have*

$$\left.\begin{array}{r} x_n \overset{\|\cdot\|_X}{\to} x \\ \sigma_n \overset{d}{\to} \sigma \end{array}\right\} \Rightarrow U_{\sigma_n}(t,\tau)x_n \rightharpoonup U_\sigma(t,\tau)x \quad \textit{weakly in } X.$$

For convenience, we also use the following notation.

Definition 3.6. *A semigroup* $\{S(t)\}_{t\geqslant 0}$: $X\times\Sigma\to X\times\Sigma$ *is to be called skew productively norm-to-weak continuous if, for any fixed* $t\geqslant 0$ *and for any* $\{x_n\}\subset X$ *and* $\{\sigma_n\}\subset\Sigma$, *we have*

$$\Pi_1 S(t)(x_n,\,\sigma_n)\overset{weak}{\to}\Pi_1 S(t)(x,\,\sigma),$$
$$\Pi_2 S(t)(x_n,\,\sigma_n)\overset{d}{\to}\Pi_2 S(t)(x,\,\sigma)$$

provided that $x_n\overset{\|\cdot\|_X}{\to}x$ *and* $\sigma_n\overset{d}{\to}\sigma$, *where* Π_1 *and* Π_2 *are the canonical projectors from* $X\times\Sigma$ *to* X *and* Σ, *respectively. Denote such continuity by*

$$S(t)(x_n,\,\sigma_n)\overset{s-w}{\to}S(t)(x,\,\sigma).$$

Following from the definition of a skew productively norm-to-weak continuous semigroup, we have the following result.

Proposition 3.7. *Assume that* $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, *is a family norm-to-weak continuous processes in* X *and the translation semigroup* $\{T(t)\}_{t\geqslant 0}$ *is continuous (w.r.t. the metric* $d(\cdot,\cdot)$*) in* Σ. *Then, the semigroup* $\{S(t)\}_{t\geqslant 0}$ *corresponding to* $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, *defined by* (2.5) *and acting on* $X\times\Sigma$, *is skew productively norm-to-weak continuous.*

3.2.2. Kernel sections of the uniform attractor.

Theorem 3.8. *Let* X *be a Banach space and* Σ *be a compact metric space. Assume that a family of processes* $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, *satisfies the translation identity* (2.3)–(2.4) *as well as the following conditions:*

(i) *The translation semigroup* $\{T(t)\}_{t\geqslant 0}$ *is continuous on* Σ;
(ii) $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, *is norm-to-weak continuous on* X;
(iii) $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, *has a bounded uniformly (w.r.t.* $\sigma\in\Sigma$*) absorbing set* B_0 *in* X;
(iv) $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, *is uniformly (w.r.t.* $\sigma\in\Sigma$*) asymptotically compact in* X.

Then, $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, *has a uniform (w.r.t.* $\sigma\in\Sigma$*) attractor* $\mathscr{A}_\Sigma$ *satisfying*

$$\mathscr{A}_\Sigma=\omega_{0,\,\Sigma}(B_0)=\bigcup_{\sigma\in\Sigma}\mathcal{K}_\sigma(s)\quad\forall\ s\in\mathbb{R},\tag{3.3}$$

where $\mathcal{K}_\sigma(s)$ *is the section at* $t=s$ *of the kernel* $\mathcal{K}_\sigma$ *of the process* $\{U_\sigma(t,\,\tau)\}$ *with symbol* σ.

The proof of this theorem is similar to what was done in Chepyzhov and Vishik [11, Theorem 5.1, Chapter IV]; the different continuity assumptions can be adjusted to those in [41] for autonomous systems, so we omit them here.

Similar to [11, 28], the following results give a method for obtaining the structure of the uniform (w.r.t. $\tau\in\mathbb{R}$) attractor of a process $\{U_{\sigma_0}(t,\tau)\}$, $\tau\in\mathbb{R}$, via the structure of the uniform (w.r.t. $\sigma\in\Sigma_0$) attractor for the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma_0$.

Since our processes are norm-to-weak continuous, we first give a simple lemma about metrizability. We recall (e.g., see Diestel [17, p. 18]) that a set $F\subset X^*$ is called total if $f(x)=0$ for each $f\in F$ implies $x=0$.

Lemma 3.9. *If* K *is a (relatively) weakly compact subset in a Banach space* X *and* K *is countable, then* $\overline{K}^{weak}$ *is metrizable, where* $\overline{K}^{weak}$ *means the weak closure of* K *in* X.

Proof. Denote $Y=\overline{span\{K\}}$.

From the convexity of $span\{K\}$ we know Y is weakly closed in X. Therefore, $K = K \cap Y$ is (relatively) weakly compact in the separable Banach space Y. Since the dual of a separable Banach space contains a countable total set, we know that $\overline{K}^{weak}$ is metrizable in Y, and from that Y is a closed subspace of X, we get $\overline{K}^{weak}$ is metrizable in X. ■

Theorem 3.10. *Let Σ_0 be a parameter set, Σ is a completion of Σ_0 w.r.t. some metric $d(\cdot,\cdot)$, and the translation semigroup $\{T(t)\}_{t\geqslant 0}$ also satisfies the translation identity (2.3)–(2.4) on Σ_0. Furthermore, assume that the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, satisfies all of the assumptions in Theorem 3.8. Then, both families of processes $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$ and $\sigma\in\Sigma_0$, have compact uniform (w.r.t. $\sigma\in\Sigma$ and $\sigma\in\Sigma_0$, respectively) attractors $\mathscr{A}_\Sigma$ and $\mathscr{A}_{\Sigma_0}$, respectively, and moreover,*

$$\mathscr{A}_{\Sigma_0} = \mathscr{A}_\Sigma = \omega_{0,\Sigma}(B_0) = \bigcup_{\sigma\in\Sigma} \mathcal{K}_\sigma(s) \quad \forall\ s\in\mathbb{R}.$$

Proof. The existence is an immediate consequence of Theorem 3.8, and obviously, we have

$$\mathscr{A}_{\Sigma_0} \subset \mathscr{A}_\Sigma = \omega_{0,\Sigma}(B_0) = \bigcup_{\sigma\in\Sigma} \mathcal{K}_\sigma(s) \quad \forall\ s\in\mathbb{R}.$$

Now we prove $\omega_{0,\Sigma_0}(B_0) = \omega_{0,\Sigma}(B_0)$. For any $y\in\omega_{0,\Sigma}(B_0)$, from Proposition 3.2, we know that there exist $x_n\in B_0$, $t_n\to\infty$, and $\sigma_n\in\Sigma$ such that

$$U_{\sigma_n}(t_n,0)x_n \to y \quad \text{as } n\to\infty. \tag{3.4}$$

On the other hand, from the assumption that Σ is the completion of Σ_0 we know that there exists $\{\sigma_m^{(n)}\}\subset\Sigma_0$ which satisfies $\sigma_m^{(n)} \overset{d}{\to} \sigma_n$ as $m\to\infty$ for each $n\in\mathbb{N}$. Therefore, due to the norm-to-weak continuity of the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, we have

$$U_{\sigma_m^{(n)}}(t_n,0)x_n \rightharpoonup U_{\sigma_n}(t_n,0)x_n \quad \text{as } m\to\infty \tag{3.5}$$

for each $n\in\mathbb{N}$. Denote $K = \{U_{\sigma_m^{(n)}}(t_n,0)x_n \mid n,m\in\mathbb{N}\}$; then K is countable and thanks to the condition (iv) of Theorem 3.8 we know that K is also relatively weakly compact in X. Consequently, from Lemma 3.9 we have that $\overline{K}^{weak}$ is metrizable.

Hence, combining (3.4) and (3.5), we can obtain that there exists $\sigma_n'\in\Sigma_0$ for each $n\in\mathbb{N}$ such that

$$U_{\sigma_n'}(t_n,0)x_n \rightharpoonup y \quad \text{in } X \text{ as } n\to\infty.$$

Then noticing the uniform asymptotic compactness again and the uniqueness of limits, we have $y\in\omega_{0,\Sigma_0}(B_0)$. ■

4. A criterion for verifying the uniform asymptotic compactness. In this section, we present a technical method to verify the uniform asymptotic compactness (given in Definition 3.1) for the family of processes generated by the nonautonomous hyperbolic type of evolutionary equations. This criterion is partially motivated by the methods in [13, 14, 15, 16, 24, 38]

for autonomous systems. In [15], the authors present a general abstract framework for *autonomous* wave equations. Here, the following results and proof are similar to those in [24, 38] for autonomous cases.

Definition 4.1. *Let X be a Banach space, B be a bounded subset of X, and Σ be a symbol (or parameter) space. We call a function $\phi(\cdot,\cdot;\cdot,\cdot)$, defined on $(X\times X)\times(\Sigma\times\Sigma)$, to be a contractive function on $B\times B$ if for any sequence $\{x_n\}_{n=1}^{\infty}\subset B$ and any $\{\sigma_n\}\subset\Sigma$, there is a subsequence $\{x_{n_k}\}_{k=1}^{\infty}\subset\{x_n\}_{n=1}^{\infty}$ and $\{\sigma_{n_k}\}_{k=1}^{\infty}\subset\{\sigma_n\}_{n=1}^{\infty}$ such that*

$$\lim_{k\to\infty}\lim_{l\to\infty}\phi(x_{n_k},x_{n_l};\sigma_{n_k},\sigma_{n_l})=0.$$

We denote the set of all contractive functions on $B\times B$ by $Contr(B,\Sigma)$.

Theorem 4.2. *Let $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, be a family of processes that satisfies the translation identity (2.3)–(2.4) on Banach space X and has a bounded uniformly (w.r.t. $\sigma\in\Sigma$) absorbing set $B_0\subset X$. Moreover, assume that for any $\varepsilon>0$ there exist $T=T(B_0,\varepsilon)$ and $\phi_T\in Contr(B_0,\Sigma)$ such that*

$$\|U_{\sigma_1}(T,0)x-U_{\sigma_2}(T,0)y\|\leqslant\varepsilon+\phi_T(x,y;\sigma_1,\sigma_2)\quad\forall\ x,\ y\in B_0,\ \forall\ \sigma_1,\ \sigma_2\in\Sigma.$$

Then $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, is uniformly (w.r.t. $\sigma\in\Sigma$) asymptotically compact in X.

Proof. For any fixed $\tau\in\mathbb{R}$, let $\{x_n\}_{n=1}^{\infty}$ be a bounded sequence of X, $\sigma_n\in\Sigma$ and $t_n\geqslant\tau$ satisfy $t_n\to\infty$ as $n\to\infty$. We need to show that

$$\{U_{\sigma_n}(t_n,\tau)x_n\}_{n=1}^{\infty}\text{ is precompact in }X.$$

Thanks to the translation identity (2.3)–(2.4), we know that for any fixed $\tau\in\mathbb{R}$ and $\sigma\in\Sigma$ we can find $\sigma'\in\Sigma$ such that

$$U_{\sigma'}(t+\tau,\tau)x=U_\sigma(t,0)x\qquad\forall\ t\geqslant0\text{ and }x\in X.\tag{4.1}$$

Therefore, we need only to show that $\{U_{\sigma_n}(t_n,0)x_n\}_{n=1}^{\infty}$ is precompact in X.

In the following, we will prove that $\{U_{\sigma_n}(t_n,0)x_n\}_{n=1}^{\infty}$ has a Cauchy subsequence via a diagonal method.

Taking $\varepsilon_m>0$ with $\varepsilon_m\to0$ as $m\to\infty$.

At first, for ε_1, by the assumptions, there exist $T_1=T_1(\varepsilon_1)$ and $\phi_1\in Contr(B_0,\Sigma)$ such that

$$\|U_{\sigma_1}(T_1,0)x-U_{\sigma_2}(T_1,0)y\|\leqslant\varepsilon_1+\phi_1(x,y;\sigma_1,\sigma_2)\qquad\forall\ x,y\in B_0\text{ and }\sigma_1,\sigma_2\in\Sigma.\tag{4.2}$$

Since $t_n\to\infty$, for such fixed T_1, without loss of generality, we assume that $t_n\gg T_1$ is so large that $U_{\sigma_n}(t_n-T_1,0)x_n\in B_0$ for each $n\in\mathbb{N}$.

Similar to (4.1), for each $n\in\mathbb{N}$, there is a $\sigma_n'\in\Sigma$ such that

$$U_{\sigma_n'}(T_1,0)=U_{\sigma_n}(t_n,t_n-T_1).\tag{4.3}$$

Let $y_n = U_{\sigma_n}(t_n - T_1, 0)x_n$; then from (4.2) and (4.3) we have

$$\begin{aligned}
&\|U_{\sigma_n}(t_n,0)x_n - U_{\sigma_m}(t_m,0)x_m\| \\
&\quad = \|U_{\sigma_n}(t_n, t_n - T_1)U_{\sigma_n}(t_n - T_1, 0)x_n - U_{\sigma_m}(t_m, t_m - T_1)U_{\sigma_n}(t_m - T_1, 0)x_m\| \\
&\quad = \|U_{\sigma_n}(t_n, t_n - T_1)y_n - U_{\sigma_m}(t_m, t_m - T_1)y_m\| \\
&\quad = \|U_{\sigma'_n}(T_1, 0)y_n - U_{\sigma'_m}(T_1, 0)y_m\| \\
&\quad \leqslant \varepsilon_1 + \phi_1(y_n, y_m; \sigma'_n, \sigma'_m). \qquad (4.4)
\end{aligned}$$

Due to the definition of $Contr(B_0, \Sigma)$ and $\phi_1 \in Contr(B_0, \Sigma)$, we know that $\{y_n\}_{n=1}^{\infty}$ has a subsequence $\{y_{n_k}^{(1)}\}_{k=1}^{\infty}$ and $\{\sigma'_n\}_{n=1}^{\infty}$ has a subsequence $\{\sigma'^{(1)}_{n_k}\}_{k=1}^{\infty}$ such that

$$\lim_{k\to\infty}\lim_{l\to\infty}\phi_1(y_{n_k}^{(1)}, y_{n_l}^{(1)}; \sigma'^{(1)}_{n_k}, \sigma'^{(1)}_{n_l}) \leqslant \frac{\varepsilon_1}{2}. \qquad (4.5)$$

And similar to the autonomous cases (e.g., see [24, 38]), we have

$$\begin{aligned}
&\lim_{k\to\infty}\sup_{p\in\mathbb{N}}\|U_{\sigma^{(1)}_{n_{k+p}}}(t^{(1)}_{n_{k+p}}, 0)x^{(1)}_{n_{k+p}} - U_{\sigma^{(1)}_{n_k}}(t^{(1)}_{n_k}, 0)x^{(1)}_{n_k}\| \\
&\quad \leqslant \lim_{k\to\infty}\sup_{p\in\mathbb{N}}\limsup_{l\to\infty}\|U_{\sigma^{(1)}_{n_{k+p}}}(t^{(1)}_{n_{k+p}}, 0)x^{(1)}_{n_{k+p}} - U_{\sigma^{(1)}_{n_l}}(t^{(1)}_{n_l}, 0)x^{(1)}_{n_l}\| \\
&\qquad + \limsup_{k\to\infty}\limsup_{l\to\infty}\|U_{\sigma^{(1)}_{n_k}}(t^{(1)}_{n_k}, 0)x^{(1)}_{n_k} - U_{\sigma^{(1)}_{n_l}}(t^{(1)}_{n_l}, 0)x^{(1)}_{n_l}\| \\
&\quad \leqslant \varepsilon_1 + \lim_{k\to\infty}\sup_{p\in\mathbb{N}}\lim_{l\to\infty}\phi_1(y^{(1)}_{n_{k+p}}, y^{(1)}_{n_l}; \sigma'^{(1)}_{n_{k+p}}, \sigma'^{(1)}_{n_l}) \\
&\qquad + \varepsilon_1 + \lim_{k\to\infty}\lim_{l\to\infty}\phi_1(y^{(1)}_{n_k}, y^{(1)}_{n_l}; \sigma'^{(1)}_{n_k}, \sigma'^{(1)}_{n_l}),
\end{aligned}$$

which, combined with (4.4) and (4.5), implies that

$$\lim_{k\to\infty}\sup_{p\in\mathbb{N}}\|U_{\sigma^{(1)}_{n_{k+p}}}(t^{(1)}_{n_{k+p}}, 0)x^{(1)}_{n_{k+p}} - U_{\sigma^{(1)}_{n_k}}(t^{(1)}_{n_k}, 0)x^{(1)}_{n_k}\| \leqslant 4\varepsilon_1.$$

Therefore, there is a K_1 such that

$$\|U_{\sigma^{(1)}_{n_k}}(t^{(1)}_{n_k}, 0)x^{(1)}_{n_k} - U_{\sigma^{(1)}_{n_l}}(t^{(1)}_{n_l}, 0)x^{(1)}_{n_l}\| \leqslant 5\varepsilon_1 \quad \forall\ k, l \geqslant K_1.$$

By induction, we obtain that, for each $m \geqslant 1$, there is a subsequence $\{U_{\sigma^{(m+1)}_{n_k}}(t^{(m+1)}_{n_k}, 0)\cdot x^{(m+1)}_{n_k}\}_{k=1}^{\infty}$ of $\{U_{\sigma^{(m)}_{n_k}}(t^{(m)}_{n_k}, 0)x^{(m)}_{n_k}\}_{k=1}^{\infty}$ and certain K_{m+1} such that

$$\|U_{\sigma^{(m+1)}_{n_k}}(t^{(m+1)}_{n_k}, 0)x^{(m+1)}_{n_k} - U_{\sigma^{(m+1)}_{n_l}}(t^{(m+1)}_{n_l}, 0)x^{(m+1)}_{n_l}\| \leqslant 5\varepsilon_{m+1} \quad \forall\ k, l \geqslant K_{m+1}.$$

Now, we consider the diagonal subsequence $\{U_{\sigma^{(k)}_{n_k}}(t^{(k)}_{n_k}, 0)x^{(k)}_{n_k}\}_{k=1}^{\infty}$. Since for each $m \in \mathbb{N}$, $\{U_{\sigma^{(k)}_{n_k}}(t^{(k)}_{n_k}, 0)x^{(k)}_{n_k}\}_{k=m}^{\infty}$ is a subsequence of $\{U_{\sigma^{(m)}_{n_k}}(t^{(m)}_{n_k}, 0)x^{(m)}_{n_k}\}_{k=1}^{\infty}$, then

$$\|U_{\sigma^{(k)}_{n_k}}(t^{(k)}_{n_k}, 0)x^{(k)}_{n_k} - U_{\sigma^{(l)}_{n_l}}(t^{(l)}_{n_l}, 0)x^{(l)}_{n_l}\| \leqslant 5\varepsilon_m \quad \forall\ k, l \geqslant \max\{m, K_m\},$$

which, combined with $\varepsilon_m \to 0$ as $m \to \infty$, implies that $\{U_{\sigma^{(k)}_{n_k}}(t^{(k)}_{n_k}, 0)x^{(k)}_{n_k}\}_{k=m}^{\infty}$ is a Cauchy sequence in X. This shows that $\{U_{\sigma_n}(t_n, 0)x_n\}_{n=1}^{\infty}$ is precompact in X. ■

5. Application to wave equation.

5.1. Mathematical setting. Similar to the autonomous cases (e.g., see [14]), applying the Galerkin approximation method, we have the following existence and uniqueness results (e.g., see [22, 27]), and the time-dependent terms make no essential complications.

Theorem 5.1. *Let Ω be a bounded domain of $\mathbb{R}^3$ with smooth boundary, h and f satisfy (1.4)–(1.8), and $g \in L^\infty(\mathbb{R}; L^2(\Omega))$. Then the nonautonomous initial boundary value problem (1.1)–(1.3) has a unique solution $u(t)$ satisfying $(u(t), u_t(t)) \in \mathcal{C}(\mathbb{R}^\tau; H_0^1(\Omega) \times L^2(\Omega))$ and $\partial_{tt}u(t) \in L^2_{loc}(\mathbb{R}^\tau; H^{-1}(\Omega))$ for any initial data $(u^{0\tau}, u^{1\tau}) \in H_0^1(\Omega) \times L^2(\Omega)$. Moreover, the solution satisfies the energy inequalities.*

For convenience, hereafter let $|\cdot|$ and $\|\cdot\|$ be the norms of $L^2(\Omega)$ and $H_0^1(\Omega)$, respectively, and C a general positive constant, which may be different in different estimates.

We use the notation in Chepyzhov and Vishik [11]: Let $y(t) = (u(t), u_t(t))$, $y_\tau = (u^{0\tau}, u^{1\tau})$, and $X = H_0^1(\Omega) \times L^2(\Omega)$ with finite energy norm

$$\|y\|_X = \{\|u\|^2 + |u_t|^2\}^{\frac{1}{2}}.$$

Let $A_{\sigma(t)}(u, v) = (v, \Delta u - f(u) - h(v) + \sigma(t))$. Then the nonautonomous system (1.1)–(1.3) can be rewritten in the operator form

$$\partial_t y = A_{\sigma(t)}(y), \qquad y|_{t=\tau} = y_\tau, \tag{5.1}$$

where $\sigma(s) = g(x, s)$ is symbol of (5.1).

We now define the symbol space for (5.1). Taking a fixed symbol $\sigma_0(s) = g_0(x, s)$, $g_0 \in L^\infty(\mathbb{R}; L^2(\Omega)) \cap W_b^{1,r}(\mathbb{R}; L^r(\Omega))$ for some $r > \frac{6}{5}$. Set

$$\Sigma_0 = \{(x, t) \mapsto g_0(x, t + h) : \ h \in \mathbb{R}\} \tag{5.2}$$

and

$$\Sigma \text{ is the } *\text{-weakly closure of } \Sigma_0 \text{ in } L^\infty(\mathbb{R}; L^2(\Omega)) \cap W_b^{1,r}(\mathbb{R}; L^r(\Omega)). \tag{5.3}$$

Then we have the following simple properties.

Proposition 5.2.

(i) *Σ is bounded in $L^\infty(\mathbb{R}; L^2(\Omega)) \cap W_b^{1,r}(\mathbb{R}; L^r(\Omega))$, and for any $\sigma \in \Sigma$, the following estimate holds:*

$$\|\sigma\|_{L^\infty(\mathbb{R}; L^2(\Omega)) \cap W_b^{1,r}(\mathbb{R}; L^r(\Omega))} \leqslant \|g_0\|_{L^\infty(\mathbb{R}; L^2(\Omega)) \cap W_b^{1,r}(\mathbb{R}; L^r(\Omega))}.$$

(ii) *The translation semigroup $\{T(h) | h \geqslant 0\}$ acting on Σ is invariant in Σ; that is,*

$$T(h)\Sigma = \Sigma \quad \forall\ h \in \mathbb{R}^+.$$

Thus, from Theorem 5.1, we know that (1.1)–(1.3) is well posed for all $\sigma(s) \in \Sigma$ and generates a family of processes $\{U_\sigma(t, \tau)\}$, $\sigma \in \Sigma$, given by the formula $U_\sigma(t, \tau)y^\tau = y(t)$, where $y(t)$ is the solution of (1.1)–(1.8), and $\{U_\sigma(t, \tau)\}$, $\sigma \in \Sigma$, satisfies (2.1)–(2.2). At the same time, by the unique solvability, we know that $\{U_\sigma(t, \tau)\}$, $\sigma \in \Sigma$, satisfies the translation identity (2.3).

In what follows, we denote by $\{U_\sigma(t, \tau)\}$, $\sigma \in \Sigma$, the family of processes generated by (5.1)–(5.3).

5.2. Bounded uniformly (w.r.t. $\sigma \in \Sigma$) absorbing set. We begin with the following result on the existence of a bounded uniformly (w.r.t. $\sigma \in \Sigma$) absorbing set. Its proof is essentially established in Haraux [22], and for the reader's convenience, we replicate it here and make only a few minor changes for our problem.

Theorem 5.3. *Under the assumptions of Theorem 5.1, the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, corresponding to (5.1) has a bounded (in X) uniformly (w.r.t. $\sigma \in \Sigma$) absorbing set B_0, i.e., there exists a positive constant ρ, which depends on $\|g_0\|_{L^\infty(\mathbb{R};L^2(\Omega))}$ and the coefficients in (1.6)–(1.8), such that for any bounded subset $B \subset X$ and any $\tau \in \mathbb{R}$, there is a $T = T(B)$ such that for any $t - \tau \geqslant T$, $\sigma \in \Sigma$, and $(u^{0\tau}, u^{1\tau}) \in B$,*

$$\|U_\sigma(t,\tau)(u^{0\tau}, u^{1\tau})\|_X \leqslant \rho.$$

Proof. Since $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, satisfies the translation identity, we need only to prove Theorem 5.3 for the cases $\tau \equiv 0$. Moreover, from the definition of Σ we know that for all $\sigma \in \Sigma$,

$$\|\sigma\|_{L^\infty(\mathbb{R};L^2(\Omega))} \leqslant \|g_0\|_{L^\infty(\mathbb{R};L^2(\Omega))}.$$

Hence, without loss of generality, in the remainder of the proof, we will not point out the difference in symbols and will denote different σ by g.

For any $\varepsilon \geqslant 0$, we set

$$E_\varepsilon(t) = \frac{1}{2}\|u(t)\|^2 + \frac{1}{2}|u_t(t)|^2 + \int_\Omega F(u(x))dx + \varepsilon\langle u_t(t), u(t)\rangle. \tag{5.4}$$

Then we have $E_\varepsilon(t) \to E_0(t)$ as $\varepsilon \to 0$. Moreover, there exist $C_0, C_1 \geqslant 0$ such that

$$\frac{C_0}{2}(\|u(t)\|^2 + |u_t(t)|^2) - C_1 \leqslant E_0(t). \tag{5.5}$$

By differentiating (5.4) with time t, we obtain that

$$\begin{aligned}
&\frac{d}{dt}(E_\varepsilon(t))\\
&= \langle u_{tt} - \triangle u, u_t\rangle + \langle f(u), u_t(t)\rangle + \varepsilon|u_t(t)|^2 + \varepsilon\langle u_{tt}(t), u_t(t)\rangle\\
&= \langle g, u_t\rangle - \langle h(u_t), u_t\rangle + \varepsilon|u_t(t)|^2 - \varepsilon\|u\|^2 - \varepsilon\langle f(u), u\rangle + \varepsilon\langle g, u\rangle - \varepsilon\langle h(u_t), u\rangle. \qquad (5.6)
\end{aligned}$$

It is obvious that (1.5) implies that

$$\langle h(u_t), u_t\rangle \geqslant \alpha|u_t|^2 - C|\Omega|,$$

and from (1.7) and (1.8) we know that there are $\lambda_1 > \delta > 0$ and $C > 0$ such that

$$\langle f(u), u\rangle \geqslant \delta\int_\Omega F(u)dx - C.$$

Hence we get the inequality

$$\begin{aligned}
E'_\varepsilon(t) \leqslant &\left(2\varepsilon - \frac{\alpha}{2}\right)|u_t(t)|^2 - \frac{\varepsilon}{2}\|u(t)\|^2 - \varepsilon\delta\int_\Omega F(u)dx\\
&- \frac{1}{2}\langle h(u_t), u_t\rangle + \varepsilon\|u(t)\|\|h(u_t(t))\|_{H^{-1}} + C,
\end{aligned}$$

where C depends on $\|g\|_{L^\infty(\mathbb{R},L^2(\Omega))}$.

On the other hand, from (1.4)–(1.6) we have (e.g., see the lemma in [22]) that there is a constant K such that

$$\|h(v)\|_{H^{-1}} \leqslant K(1+\langle h(v),v\rangle) \quad \forall\, v\in H_0^1(\Omega).$$

Denote

$$w(t)\triangleq 1+\langle h(u_t(t)),u_t(t)\rangle (>0).$$

Then, by taking ε small enough, we obtain that, for all $t\geqslant 0$,

$$\begin{aligned} E_\varepsilon'(t) &\leqslant -\gamma\varepsilon E_\varepsilon(t)+(K\varepsilon\|u(t)\|-1/2)w(t)+C \\ &\leqslant -\gamma\varepsilon E_\varepsilon(t)+(N\varepsilon\sqrt{E_\varepsilon(t)}-1/2)w(t)+C, \end{aligned} \tag{5.7}$$

where $N,C>0$ depending only on f, g, h, and Ω (not on the initial data) and $\gamma>0$.

Now choose $\varepsilon>0$ so small that

$$E_\varepsilon(0)<\left(\frac{1}{2N\varepsilon}\right)^2-\frac{C}{\gamma\varepsilon}. \tag{5.8}$$

Then,

$$E_\varepsilon(t)<\left(\frac{1}{2N\varepsilon}\right)^2 \qquad \forall\, t\geqslant 0. \tag{5.9}$$

If (5.9) is not true, let $t_0=\inf\{t\geqslant 0,\ E_\varepsilon(t)\geqslant(\frac{1}{2N\varepsilon})^2\}$; then $E_\varepsilon(t_0)=(\frac{1}{2N\varepsilon})^2$, and for all $t\in[0,t_0]$, we have

$$E_\varepsilon(t)\leqslant\left(\frac{1}{2N\varepsilon}\right)^2. \tag{5.10}$$

Therefore, from (5.7) and (5.10) we can obtain that

$$E_\varepsilon(t_0)\leqslant e^{-\gamma\varepsilon t_0}E_\varepsilon(0)+\frac{C}{\gamma\varepsilon}<\left(\frac{1}{2N\varepsilon}\right)^2. \tag{5.11}$$

This is a contradiction and means that (5.9) is indeed satisfied.

Combining (5.7) and (5.9), by use of the uniform Gronwall lemma, we obtain that

$$E_\varepsilon(t)\leqslant e^{-\gamma\varepsilon t}E_\varepsilon(0)+\frac{C}{\gamma\varepsilon}. \tag{5.12}$$

Finally, we notice that for every bounded set $B\subset X$, we assume the bounds of B (in X) are E (>0); then by taking $1/\varepsilon=4N\sqrt{E}$ we can obtain $\frac{1}{(2N\varepsilon)^2}-\frac{C}{\gamma\varepsilon}\geqslant E$ for any E large enough. It follows from (5.12) that there exist $M>0$ (independent of the initial data) and $T=T(B)$ such that

$$E_0(t)\leqslant M(1+\sqrt{E_0(0)}) \quad \forall\, t\geqslant T \text{ and } (u^{0\tau},u^{1\tau})\in B. \tag{5.13}$$

Without loss of generality, assume $M > 1$; then from (5.13) we have that for any bounded set $B \subset X$, there is a $T = T(B)$ such that

$$E_0(t) \leqslant 4M^2 + 1 \quad \forall\, t \geqslant T \text{ and } (u^{0\tau}, u^{1\tau}) \in B.$$

Combining this with (5.5), we know that Theorem 5.3 is true. ∎

Remark 5.4. *If $g \in W_b^{1,1}(\mathbb{R}; H^{-1})$, then we can also obtain the same results about the existence of a bounded uniformly absorbing set. For the proof we need only to replace $E_\varepsilon(t)$ by a new energy functional $G_\varepsilon(t) = E_\varepsilon(t) - \langle g, u\rangle$, replace the constant $\frac{C}{\gamma\varepsilon}$ in (5.8), (5.11), and (5.12) by another constant $C^\star$, where*

$$C^\star = \frac{1}{1 - e^{-\gamma\varepsilon}} \|g_t\|_{L_b^1(\mathbb{R};H^{-1})} + \|g\|_{L^\infty(\mathbb{R};H^{-1})} + 1,$$

and note that $C^\star\varepsilon^2 \to 0$ as $\varepsilon \to 0$.

5.3. Uniform (w.r.t. $\sigma \in \Sigma$) asymptotic compactness. The main result in this subsection is summarized in the following theorem.

Theorem 5.5. *Let Ω be a bounded domain in $\mathbb{R}^3$ with smooth boundary, and h and f satisfy (1.4)–(1.8). If $g_0 \in L^\infty(\mathbb{R}; L^2(\Omega)) \cap W_b^{1,r}(\mathbb{R}; L^r(\Omega))$ for some $r > \frac{6}{5}$ and Σ is defined by (5.3), then the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, corresponding to (5.1) or (1.1), is uniformly (w.r.t. $\sigma \in \Sigma$) asymptotically compact in $H_0^1(\Omega) \times L^2(\Omega)$.*

The idea for the proof is similar to that in Chueshov and Lasiecka [13, 14, 15] and Khanmamedov [24]; see also [38] for linear damping and autonomous cases.

Hereafter, we always assume that the hypotheses of Theorem 5.1 hold and denote by B_0 the bounded uniformly absorbing set obtained in Theorem 5.3.

5.3.1. Preliminaries. Note that condition (1.6) implies that

$$|h(s)|^{\frac{1}{p}} \leqslant C(1 + |s|);$$

therefore, we have

$$|h(s)|^{\frac{p+1}{p}} = |h(s)|^{\frac{1}{p}} \cdot |h(s)| \leqslant C(1 + |s|)|h(s)| \leqslant C|h(s)| + Ch(s) \cdot s.$$

Combining Young's inequality and (1.4), we obtain that

$$|h(s)|^{\frac{p+1}{p}} \leqslant C(1 + h(s) \cdot s) \quad \forall\, s \in \mathbb{R}, \tag{5.14}$$

where the constant C is independent of s. Moreover, we recall the following result.

Lemma 5.6 (see [19, 24]). *Let h satisfy (1.4) and (1.5). Then for any $\delta > 0$, there exists a constant C_δ, depending on δ, such that*

$$|u - v|^2 \leqslant \delta + C_\delta(h(u) - h(v))(u - v) \quad \text{for any } u, v \in \mathbb{R}.$$

Proposition 5.7. *Let $s_i \in \mathbb{R}$ $(i = 1, 2, \dots)$, $g \in W_b^{1,r}(\mathbb{R}; L^r(\Omega))$ for some $r > \frac{6}{5}$, $\{u_n(t) |\ t \geqslant 0,\ n = 1, 2, \dots\}$ is bounded in $H_0^1(\Omega)$, and for any $T_1 > 0$, $\{u_{n_t}(t) |\ n = 1, 2, \dots\}$ is bounded*

in $L^\infty(0,T_1;L^2(\Omega))$. *Then for any* $T>0$, *there exist subsequences* $\{u_{n_k}\}_{k=1}^\infty$ *of* $\{u_n\}_{n=1}^\infty$ *and* $\{s_{n_k}\}_{k=1}^\infty$ *of* $\{s_n\}_{n=1}^\infty$ *such that*

$$\lim_{k\to\infty}\lim_{l\to\infty}\int_0^T\int_s^T\int_\Omega (g(x,\tau+s_{n_k})-g(x,\tau+s_{n_l}))(u_{n_k}-u_{n_l})_t(\tau)dxd\tau ds=0.$$

Proof. Since $\{u_n(t)|\ t\geqslant 0,\ n=1,2,\dots\}$ is bounded in $H_0^1(\Omega)$ and for any $T_1>0$, $\{u_{n_t}(t)|\ n=1,2,\dots\}$ is bounded in $L^\infty(0,T_1;L^2(\Omega))$, then for any $T>0$, without loss of generality (at most by passing subsequence), we assume

$$u_n(T)\to u_0\quad\text{in}\quad L^2(\Omega)$$

and

$$u_n\to v\quad\text{in}\quad L^k(0,T;\,L^k(\Omega))\qquad\left(\text{this requires } r>\frac{6}{5}\right),$$

where $k<6$.

Note that

$$\begin{aligned}&(g(x,t+s_i)-g(x,t+s_j))w_t(t)\\&\quad=\frac{d}{dt}((g(x,t+s_i)-g(x,t+s_j))w(t))-(g_t(x,t+s_i)-g_t(x,t+s_j))w(t)\end{aligned}$$

for any $w\in W^{1,2}_{loc}(\mathbb{R};L^2(\Omega))$; then by use of the Hölder inequality, we obtain that

$$\begin{aligned}&\lim_{n\to\infty}\lim_{m\to\infty}\int_0^T\int_s^T\int_\Omega (g(x,\tau+s_n)-g(x,\tau+s_m))(u_n-u_m)_t(\tau)dxd\tau ds\\&\leqslant\lim_{n\to\infty}\lim_{m\to\infty}2MT^{\frac{2r+1}{r}}\left(\int_\Omega|u_n(T)-u_m(T)|^{\frac{r}{r-1}}dx\right)^{\frac{r-1}{r}}\\&\quad+\lim_{n\to\infty}\lim_{m\to\infty}2MT^{\frac{r+1}{r}}\left(\int_0^T\int_\Omega|u_n(s)-u_m(s)|^{\frac{r}{r-1}}dxds\right)^{\frac{r-1}{r}}\\&\quad+\lim_{n\to\infty}\lim_{m\to\infty}T\int_0^T\int_\Omega|(g_t(x,s+s_n)-g_t(x,s+s_m))(u_n(s)-u_m(s))|dxds\\&=\lim_{n\to\infty}\lim_{m\to\infty}T\int_0^T\int_\Omega|(g_t(x,s+s_n)-g_t(x,s+s_m))(u_n(s)-u_m(s))|dxds\\&\leqslant\lim_{n\to\infty}\lim_{m\to\infty}T\left(\int_0^T\int_\Omega|g_t(x,s+s_n)-g_t(x,s+s_m)|^r\right)^{\frac{1}{r}}\left(\int_0^T\int_\Omega|u_n(s)-u_m(s)|^{\frac{r}{r-1}}\right)^{\frac{r-1}{r}}\\&=0,\end{aligned}$$

where the constant M depends only on $\|g\|_{W_b^{1,r}(\mathbb{R};L^r(\Omega))}$. ∎

Remark 5.8. *If* $g\in W_b^{1,1+\delta}(\mathbb{R};H^{-1+\delta})$ *with some* $\delta>0$, *then by the compact embedding*

$$W^{1,1+\delta}(0,T;\,H^{-1+\delta})\hookrightarrow W^{1,1}(0,T;\,H^{-1})\ \text{for any}\ T>0,$$

we can obtain Proposition 5.7 *directly.*

5.3.2. A priori estimates. The main purpose of this part is to establish (5.22)–(5.23), which will be used to obtain the asymptotic compactness. In the following, we deal only with the strong solutions; the generalized solution case then follows easily by a density argument.

For any $(u_0^i, v_0^i) \in B_0$, let $(u_i(t), u_{i_t}(t))$ be the corresponding solution to σ_i w.r.t. initial data (u_0^i, v_0^i), $i = 1, 2$; that is, $(u_i(t), u_{i_t}(t))$ is the solution of the following equation:

$$\begin{cases} u_{tt} + h(u_t) - \Delta u + f(u(t)) = \sigma_i(x,t), \\ (u(0),\, u_t(0)) = (u_0^i,\, v_0^i), \quad u|_{\partial\Omega} = 0. \end{cases} \tag{5.15}$$

For convenience, we denote

$$g_i(t) = \sigma_i(x,\, t), \quad h_i(t) = h(u_{i_t}(t)), \quad t \geqslant 0, \quad i = 1, 2, \quad \text{and} \quad w(t) = u_1(t) - u_2(t).$$

Then $w(t)$ satisfies

$$\begin{cases} w_{tt} + h_1(t) - h_2(t) - \Delta w + f(u_1(t)) - f(u_2(t)) = g_1(t) - g_2(t), \\ (w(0),\, w_t(0)) = (u_0^1,\, v_0^1) - (u_0^2,\, v_0^2), \quad w|_{\partial\Omega} = 0. \end{cases} \tag{5.16}$$

Set

$$E_w(t) = \frac{1}{2}\int_\Omega |w(t)|^2 + \frac{1}{2}\int_\Omega |\nabla w(t)|^2.$$

Step 1. Multiplying (5.16) by $w_t(t)$ and integrating over $[s, T] \times \Omega$, we obtain

$$\begin{aligned} &E_w(T) + \int_s^T\int_\Omega (h_1(\tau) - h_2(\tau))w_t(\tau)dxd\tau + \int_s^T\int_\Omega (f(u_1(\tau)) - f(u_2(\tau)))w_t(\tau)dxd\tau \\ &\leqslant \int_s^T\int_\Omega (g_1(\tau) - g_2(\tau))w_t(\tau)dxd\tau + E_w(s), \end{aligned} \tag{5.17}$$

where $0 \leqslant s \leqslant T$. Then

$$\begin{aligned} \int_s^T\int_\Omega (h_1(\tau) - h_2(\tau))w_t(\tau)dxd\tau \leqslant E_w(s) &+ \int_s^T\int_\Omega (g_1(\tau) - g_2(\tau))w_t(\tau)dxd\tau \\ &- \int_s^T\int_\Omega (f(u_1(\tau)) - f(u_2(\tau)))w_t(\tau)dxd\tau. \end{aligned}$$

Combining this with Lemma 5.6, we get that, for any $\delta > 0$,

$$\begin{aligned} \int_s^T\int_\Omega |w_t(\tau)|^2 dxd\tau \leqslant |T - s|\delta \cdot mes(\Omega) + C_\delta E_w(s) + C_\delta \int_s^T\int_\Omega (g_1 - g_2)w_t dxd\tau \\ - C_\delta \int_s^T\int_\Omega (f(u_1(\tau)) - f(u_2(\tau)))w_t x d\tau. \end{aligned} \tag{5.18}$$

Step 2. Multiplying (5.16) by $w(t)$ and integrating over $[0,T]\times\Omega$, we get that

$$\begin{aligned}&\int_0^T\int_\Omega|\nabla w(s)|^2dxds+\int_\Omega w_t(T)\cdot w(T)dx\\&\leqslant\int_0^T\int_\Omega|w_t(s)|^2dxds-\int_0^T\int_\Omega(h_1-h_2)wdxds+\int_\Omega w_t(0)\cdot w(0)dx\\&\quad-\int_0^T\int_\Omega(f(u_1(s))-f(u_2(s)))wdxds+\int_0^T\int_\Omega(g_1-g_2)wdxds.\end{aligned}\tag{5.19}$$

Therefore, from (5.18) and (5.19), we have

$$\begin{aligned}2\int_0^T E_w(s)ds&\leqslant 2\delta T\,mes(\Omega)+2C_\delta E_w(0)+2C_\delta\int_0^T\int_\Omega(g_1-g_2)w_tdxds\\&\quad-2C_\delta\int_0^T\int_\Omega(f(u_1(s))-f(u_2(s)))w_tdxds-\int_\Omega w_t(T)w(T)+\int_\Omega w_t(0)w(0)\\&\quad-\int_0^T\int_\Omega(h_1-h_2)w-\int_0^T\int_\Omega(f(u_1(s))-f(u_2(s)))w+\int_0^T\int_\Omega(g_1-g_2)w.\end{aligned}$$

Integrating (5.17) over $[0,T]$ w.r.t. s, we have that

$$\begin{aligned}&TE_w(T)+\int_0^T\int_s^T\int_\Omega(h_1(\tau)-h_2(\tau))w_t(\tau)dxd\tau ds\\&\leqslant-\int_0^T\int_s^T\int_\Omega(f(u_1(\tau))-f(u_2(\tau)))w_tdxd\tau ds\\&\qquad+\int_0^T\int_s^T\int_\Omega(g_1-g_2)w_tdxd\tau ds+\int_0^T E_w(s)ds\\&\leqslant-\int_0^T\int_s^T\int_\Omega(f(u_1(\tau))-f(u_2(\tau)))w_tdxd\tau ds+\int_0^T\int_s^T\int_\Omega(g_1-g_2)w_tdxd\tau ds\\&\quad+\delta T\,mes(\Omega)+C_\delta E_w(0)+C_\delta\int_0^T\int_s^T\int_\Omega(g_1-g_2)w_tdxds\\&\quad-C_\delta\int_0^T\int_\Omega(f(u_1(s))-f(u_2(s)))w_tdxds-\frac12\int_\Omega w_t(T)w(T)+\frac12\int_\Omega w_t(0)w(0)\\&\quad-\frac12\int_0^T\int_\Omega(h_1-h_2)w-\frac12\int_0^T\int_\Omega(f(u_1(s))-f(u_2(s)))w+\frac12\int_0^T\int_\Omega(g_1-g_2)w.\end{aligned}$$

Step 3. We need to deal with $\int_0^T\int_\Omega(h_1-h_2)w$. Multiplying (5.15) by $u_{i_t}(t)$ and combining this with the existence of bounded uniformly absorbing set, we have

$$\int_0^T\int_\Omega h(u_{i_t})u_{i_t}\leqslant M_T,$$

where the constant M_T depends on T (which is different from the autonomous cases) and the bounds of B_0. Then, noticing (5.14), we obtain that

$$\int_0^T\int_\Omega|h(u_{i_t})|^{\frac{p+1}{p}}dxds\leqslant M_T.\tag{5.20}$$

Therefore, using the Hölder inequality, from (5.20) we have

$$\left|\int_0^T\int_\Omega (h_1-h_2)w\right| \leqslant 2M_T^{\frac{p}{p+1}}\left(\int_0^T\int_\Omega |w|^{p+1}\right)^{\frac{1}{p+1}}. \tag{5.21}$$

Remark 5.9. *To some extent,* (5.21) *requires that the growth order of* h *is strictly less than* 5.

Set

$$\begin{aligned}
&\phi_{\delta,T}((u_0^1,v_0^1),(u_0^2,v_0^2);\sigma_1,\sigma_2)\\
&=-\int_0^T\int_s^T\int_\Omega (f(u_1(\tau))-f(u_2(\tau)))w_t dxd\tau ds+(1+C_\delta)\int_0^T\int_s^T\int_\Omega (g_1-g_2)w_t dxd\tau ds\\
&\quad -C_\delta\int_0^T\int_\Omega (f(u_1(s))-f(u_2(s)))w_t dxds-\frac{1}{2}\int_0^T\int_\Omega (h_1-h_2)wdxds\\
&\quad -\frac{1}{2}\int_0^T\int_\Omega (f(u_1(s))-f(u_2(s)))wdxds+\frac{1}{2}\int_0^T\int_\Omega (g_1-g_2)wdxds,
\end{aligned} \tag{5.22}$$

$$C_M=\delta T\, mes(\Omega)+C_\delta E_w(0)-\frac{1}{2}\int_\Omega w_t(T)w(T)d+\frac{1}{2}\int_\Omega w_t(0)w(0)dx.$$

Then we have

$$E_w\leqslant \frac{C_M}{T}+\frac{1}{T}\phi_{\delta,T}((u_0^1,v_0^1),(u_0^2,v_0^2);\sigma_1,\sigma_2). \tag{5.23}$$

5.3.3. Uniform asymptotic compactness.

Proof of Theorem 5.5. Since the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, has a bounded uniformly absorbing set, for any fixed $\varepsilon>0$, we can first choose $\delta\leqslant\frac{\varepsilon}{2\,mes(\Omega)}$ and then let T be so large that

$$\frac{C_M}{T}\leqslant\varepsilon.$$

Hence, thanks to Theorem 4.2, we need only to verify that $\phi_{\delta,T}(\cdot,\cdot;\,\cdot,\cdot)\in Contr(B_0,\Sigma)$ for each fixed T.

At first, we can observe from the proof procedure of Theorem 5.3 that for any fixed T, we have

$$\bigcup_{\sigma\in\Sigma}\bigcup_{t\in[0,T]} U_\sigma(t,0)B_0 \text{ is bounded in } E_0, \tag{5.24}$$

and the bound depends on T.

Let (u_n,u_{n_t}) be the solutions corresponding to initial data $(u_0^n,v_0^n)\in B_0$ w.r.t. symbol $\sigma_n\in\Sigma$, $n=1,2,\ldots$. Then, from (5.24), without loss of generality (at most by passing

subsequence), we assume that

$$u_n \to u \quad \star\text{-weakly in } L^\infty(0,T;\, L^6(\Omega)), \tag{5.25}$$

$$u_n \to u \quad \text{in } L^{p+1}(0,T;\, L^{p+1}(\Omega)), \tag{5.26}$$

$$u_{n_t} \to u_t \quad \star\text{-weakly in } L^\infty(0,T;\, L^2(\Omega)), \tag{5.27}$$

$$u_n \to u \quad \text{in } L^2(0,T;\, L^2(\Omega)), \tag{5.28}$$

and

$$u_n(0) \to u(0) \ \text{ and } \ u_n(T) \to u(T) \quad \text{in } L^4(\Omega), \tag{5.29}$$

where we used the compact embeddings $H_0^1 \hookrightarrow L^4$ and $H_0^1 \hookrightarrow L^{p+1}$ (since $p < 5$).

Now, we will deal with each term corresponding to that in (5.22) one by one.

First, from (5.21), we have

$$\left|\int_0^T\int_\Omega (h(u_{n_t}(s)) - h(u_{m_t}(s)))(u_n(s) - u_m(s))dxds\right| \leqslant 2M_T^{\frac{p}{p+1}} \left(\int_0^T\int_\Omega |u_n(s) - u_m(s)|^{p+1}\right)^{\frac{1}{p+1}},$$

where M_T depends on T and the bound of B_0 in $H_0^1 \times L^2$. Therefore, from (5.26) we can get

$$\lim_{n\to\infty}\lim_{m\to\infty}\int_0^T\int_\Omega (h(u_{n_t}(s)) - h(u_{m_t}(s)))(u_n(s) - u_m(s))dxds = 0. \tag{5.30}$$

Second, from Proposition 5.7 and (5.29), we have

$$\lim_{n\to\infty}\lim_{m\to\infty}\int_0^T\int_\Omega (g_n(x,\, s) - g_m(x,\, s))(u_{n_t}(s) - u_{m_t}(s))dxds = 0 \tag{5.31}$$

and

$$\lim_{n\to\infty}\lim_{m\to\infty}\int_0^T\int_s^T\int_\Omega (g_n(x,\, \tau) - g_m(x,\, \tau))(u_{n_t}(\tau) - u_{m_t}(\tau))dxd\tau ds = 0. \tag{5.32}$$

At the same time, from the growth condition (1.7) and (5.28), we can get easily that

$$\lim_{n\to\infty}\lim_{m\to\infty}\int_0^T\int_\Omega (f(u_n(s)) - f(u_m(s)))(u_n(s) - u_m(s))dxds = 0. \tag{5.33}$$

Finally, since

$$\begin{aligned}&\int_0^T\int_\Omega (u_{n_t}(s) - u_{m_t}(s))(f(u_n(s)) - f(u_m(s)))dxds\\ &\quad = \int_\Omega F(u_n(T)) - \int_\Omega F(u_n(0)) + \int_\Omega F(u_m(T)) - \int_\Omega F(u_m(0))\\ &\qquad - \int_0^T\int_\Omega u_{n_t} f(u_m(s)) - \int_0^T\int_\Omega u_{m_t} f(u_n(s)),\end{aligned}$$

then, by use of (5.25), (5.27), (5.29), and (1.7), taking first $m \to \infty$ and then $n \to \infty$, we obtain that

$$\begin{aligned}
&\lim_{n\to\infty}\lim_{m\to\infty}\int_0^T\int_\Omega (u_{n_t}(s)-u_{m_t}(s))(f(u_n(s))-f(u_m(s)))dxds\\
&\quad=\int_\Omega F(u(T))-\int_\Omega F(u(0))+\int_\Omega F(u(T))-\int_\Omega F(u(0))\\
&\qquad-\int_0^T\int_\Omega u_t f(u(s))-\int_0^T\int_\Omega u_t f(u(s))\\
&\quad=0. \qquad (5.34)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\int_s^T\int_\Omega (u_{n_t}(\tau)-u_{m_t}(\tau))(f(u_n(\tau))-f(u_m(\tau)))dxd\tau\\
&\quad=\int_\Omega F(u_n(T))-\int_\Omega F(u_n(s))+\int_\Omega F(u_m(T))-\int_\Omega F(u_m(s))\\
&\qquad-\int_s^T\int_\Omega u_{n_t} f(u_m(\tau))-\int_s^T\int_\Omega u_{m_t} f(u_n(\tau)).
\end{aligned}$$

At the same time, $|\int_s^T\int_\Omega (u_{n_t}(\tau)-u_{m_t}(\tau))(f(u_n(\tau))-f(u_m(\tau)))dxd\tau|$ is bounded for each fixed T; then by the Lebesgue dominated convergence theorem we have

$$\begin{aligned}
&\lim_{n\to\infty}\lim_{m\to\infty}\int_0^T\int_s^T\int_\Omega (u_{n_t}(\tau)-u_{m_t}(\tau))(f(u_n(\tau))-f(u_m(\tau)))dxd\tau ds\\
&=\int_0^T\left(\lim_{n\to\infty}\lim_{m\to\infty}\int_s^T\int_\Omega (u_{n_t}(\tau)-u_{n_t}(\tau))(f(u_n(\tau))-f(u_m(\tau)))dxd\tau\right)ds\\
&=\int_0^T 0ds=0. \qquad (5.35)
\end{aligned}$$

Hence, combining (5.30)–(5.35), we get that $\phi_{\delta,T}(\cdot,\cdot;\cdot,\cdot)\in Contr(B_0,\Sigma)$, and then this completes the proof of Theorem 5.5. ■

Remark 5.10. *If $g_0\in L^\infty(\mathbb{R};L^2(\Omega))\cap L^2_c(\mathbb{R};L^2(\Omega))$ (e.g., $g_0\in L^\infty(\mathbb{R};L^2(\Omega))$ and is a time periodic, quasi-periodic, or almost periodic function in $L^2_{loc}(\mathbb{R};L^2(\Omega))$), then we can obtain (5.31) and (5.32) directly from the definition of $L^2_c(\mathbb{R};L^2(\Omega))$. That is, we do not need the preliminaries in Proposition 5.7, and Theorem 5.5 on uniform asymptotic compactness still holds.*

Remark 5.11. *From Remarks 5.4 and 5.8, we know that Theorem 5.5 also holds if $g_0\in W_b^{1,1+\delta}(\mathbb{R};H^{-1+\delta})$ for some $\delta>0$.*

5.4. Existence of uniform attractor.

Theorem 5.12. *Let Ω be a bounded domain in $\mathbb{R}^3$ with smooth boundary, and h and f satisfy (1.4)–(1.8). If $g_0\in L^\infty(\mathbb{R};L^2(\Omega))\cap W_b^{1,r}(\mathbb{R};L^r(\Omega))$ for some $r>\frac{6}{5}$ and Σ is defined by (5.3), then the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma\in\Sigma$, corresponding to (5.1) or (1.1) has a compact uniform (w.r.t. $\sigma\in\Sigma$) attractor $\mathscr{A}_\Sigma$.*

Proof. From Theorems 5.3 and 5.5 we know that the conditions of Theorem 3.4 are all satisfied. ■

Remark 5.13. *If $g_0 \in L^\infty(\mathbb{R}; L^2(\Omega))$ and g_0 is a time periodic, quasi-periodic, or almost periodic function in $L^2_{loc}(\mathbb{R}; L^2(\Omega))$, then the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma$, corresponding to (5.1) or (1.1) has a compact uniform (w.r.t. $\sigma \in \Sigma$) attractor $\mathscr{A}_\Sigma$. Similarly, if $g_0 \in W^{1,1+\delta}_b(\mathbb{R}; H^{-1+\delta})$ for some $\delta > 0$, then the same result about the existence of a compact uniform (w.r.t. $\sigma \in \Sigma$) attractor holds.*

5.5. Structure of uniform attractor. In this subsection, we will consider the structure of a uniform attractor by applying Theorems 3.8 and 3.10.

For this purpose, we need some continuities for the processes.

Now, assume that $g_0 \in W^{1,1+\delta}_b(\mathbb{R}; H^{-1+\delta})$ for some $\delta > 0$, and define

$$\Sigma_0' = \{(x,t) \mapsto g_0(x, t+h) : \ h \in \mathbb{R}\}$$

and

(5.36) Σ' is the closure of Σ_0' w.r.t. the local weak convergence topology of $W^{1,1+\delta}_{loc}(\mathbb{R}; H^{-1+\delta})$.

Then, by the results of Chepyzhov and Vishik [11], we see that Σ' with the local weak convergence topology of $W^{1,1+\delta}_{loc}(\mathbb{R}; H^{-1+\delta})$ forms a sequentially compact and metrizable complete space, and we denote the equivalent metric by $d_1(\cdot,\cdot)$. Thus (Σ', d_1) is a compact metric space. Moreover, from [11, Lemma 4.1, Chapter V] we have the following lemma.

Lemma 5.14. *The translation semigroup $\{T(t)\}_{t\geqslant 0}$ is invariant and continuous in Σ' w.r.t. the local weak convergence topology of $W^{1,1+\delta}_{loc}(\mathbb{R}; H^{-1+\delta})$—equivalently, w.r.t. the metric d_1.*

On the other hand, from the assumptions on $h(\cdot)$, $f(\cdot)$, and g_0, we can (e.g., using the methods similar to [14, 19]) show that the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma'$: $(X, \|\cdot\|_X) \times (\Sigma', d_1) \mapsto (X, \|\cdot\|_X)$, is continuous (and obviously is norm-to-weak continuous).

Therefore, applying Theorem 3.10, from Remark 5.13 and Lemma 5.14, we have the following result.

Theorem 5.15. *Let Ω be a bounded domain in $\mathbb{R}^3$ with smooth boundary, and h and f satisfy (1.4)–(1.8). If $g_0 \in W^{1,1+\delta}_b(\mathbb{R}; H^{-1+\delta})$ for some $\delta > 0$ and Σ' is defined by (5.36), then the family of processes $\{U_\sigma(t,\tau)\}$, $\sigma \in \Sigma'$, corresponding to (5.1) has a compact uniform (w.r.t. $\sigma \in \Sigma'$) attractor $\mathscr{A}_{\Sigma'}$. Moreover,*

$$\mathscr{A}_{\Sigma_0'} = \mathscr{A}_{\Sigma'} = \bigcup_{\sigma\in\Sigma'} \mathcal{K}_\sigma(0).$$

6. Some discussions. As indicated by some authors (for example, [3, 5, 29, 32, 33, 39]), the energy method for proving asymptotic compactness requires that the solutions satisfy some *energy equality* to some extent. For our problem, since we assume only that the nonlinear function h satisfies (1.4)–(1.6), we can obtain only the inequality estimate (5.20). This brings an obstacle for us to verify the energy equality and apply the energy method; see [13, 14]. Fortunately, these energy inequalities (5.17)–(5.19) are enough for us to obtain the asymptotic

compactness. Note that if we take other (possibly stronger) assumptions for h (other than (1.4)–(1.6)), e.g., the standard two-sided growth assumption

$$h(v)v \geqslant -C + C_0|v|^{p+1} \text{ for any } v \in \mathbb{R},$$

then it is possible to obtain the energy equality, and the energy method applies.

Moreover, in section 3, we present a slightly improved result about the kernel section, from continuous processes to "weakly" continuous processes (i.e., norm-to-weak continuous processes). For our wave model here, under the assumptions on the nonlinear functions and forcing term, we can prove that, indeed, the corresponding process is continuous in the strong topology (i.e., with the usual topology on the energy state space and weak topology on the hull of the external force). However, the theory on attractors for norm-to-weak continuous processes maybe useful for some other systems. For example, as mentioned in Robinson [35], for a class of autonomous reaction-diffusion equations with a polynomial growth nonlinearity, the continuity of solution semigroups in H_0^1 is unknown if we do not take any restriction on the growth order of nonlinearity. The same situation is also possible for some nonautonomous systems.

Acknowledgment. The authors would like to thank the referees for very helpful comments and suggestions.

REFERENCES

[1] J. Arrieta, A. N. Carvalho, and J. K. Hale, *A damped hyperbolic equation with critical exponents*, Comm. Partial Differential Equations, 17 (1992), pp. 841–866.
[2] A. V. Babin and M. I. Vishik, *Attractors of Evolution Equations*, North–Holland, Amsterdam, 1992.
[3] J. M. Ball, *Global attractors for damped semilinear wave equations*, Discrete Contin. Dyn. Syst., 10 (2004), pp. 31–52.
[4] V. P. Bongolan-Walsh, D. Cheban, and J. Duan, *Recurrent motions in the nonautonomous Navier-Stokes system*, Discrete Contin. Dyn. Syst. Ser. B, 3 (2003), pp. 255–262.
[5] T. Caraballo, P. E. Kloeden, and J. Real, *Pullback and forward attractors for a damped wave equation with delays*, Stoch. Dyn., 4 (2004), pp. 405–423.
[6] D. N. Cheban and J. Duan, *Almost periodic motions and global attractors of the non-autonomous Navier-Stokes equations*, J. Dynam. Differential Equations, 16 (2004), pp. 1–34.
[7] V. V. Chepyzhov and A. Miranville, *Trajectory and global attractors of dissipative hyperbolic equations with memory*, Commun. Pure Appl. Anal., 4 (2005), pp. 115–142.
[8] V. V. Chepyzhov and M. I. Vishik, *A Hausdorff dimension estimate for kernel sections of nonautonomous evolution equations*, Indiana Univ. Math. J., 42 (1993), pp. 1057–1076.
[9] V. V. Chepyzhov and M. I. Vishik, *Attractors of nonautonomous dynamical systems and their dimension*, J. Math. Pures Appl. (9), 73 (1994), pp. 279–333.
[10] V. V. Chepyzhov and M. I. Vishik, *Non-autonomous evolutionary equations with translation-compact symbols and their attractors*, C. R. Acad. Sci. Paris Sér. I Math., 321 (1995), pp. 153–158.
[11] V. V. Chepyzhov and M. I. Vishik, *Attractors for Equations of Mathematical Physics*, Amer. Math. Soc. Colloq. Publ. 49, AMS, Providence, RI, 2002.
[12] L. Chierchia and J. You, *KAM tori for 1D wave equations with periodic boundary conditions*, Comm. Math. Phys., 211 (2000), pp. 497–525.
[13] I. Chueshov and I. Lasiecka, *Attractors for second-order evolution equations with nonlinear damping*, J. Dynam. Differential Equations, 16 (2004), pp. 469–512.
[14] I. Chueshov and I. Lasiecka, *Long-time behavior of second order evolution equations with nonlinear damping*, Mem. Amer. Math. Soc., to appear.

[15] I. Chueshov and I. Lasiecka, *Long Time Dynamics of Semilinear Wave Equation with Nonlinear Interior/Boundary Damping and Sources of Critical Exponents*, Comtemp. Math., AMS, Providence, RI, to appear.
[16] I. Chueshov and I. Lasiecka, *Long-time dynamics of von Karman semi-flows with nonlinear boundary/interior damping*, J. Differential Equations, 233 (2007), pp. 42–86.
[17] J. Diestel, *Sequences and Series in Banach Spaces*, Grad. Texts in Math. 92, Springer-Verlag, New York, 1984.
[18] A. Eden, V. Kalantarov, and A. Miranville, *Finite-dimensional attractors for a general class of nonautonomous wave equations*, Appl. Math. Lett., 13 (2000), pp. 17–22.
[19] E. Feireisl, *Global attractors for damped wave equations with supercritical exponent*, J. Differential Equations, 116 (1995), pp. 431–447.
[20] J.-M. Ghidaglia and A. Marzocchi, *Longtime behaviour of strongly damped wave equations, global attractors and their dimension*, SIAM J. Math. Anal., 22 (1991), pp. 879–895.
[21] J. K. Hale, *Asymptotic Behavior of Dissipative Systems*, AMS, Providence, RI, 1988.
[22] A. Haraux, *Recent results on semilinear wave equations with dissipation*, in Semigroups, Theory and Applications, Vol. I, Pitman Res. Notes Math. Ser. 141, Longman, Harlow, UK, 1986, pp. 150–157.
[23] A. Haraux, *Systèmes dynamiques dissipatifs et applications*, Masson, Paris, 1991.
[24] A. Kh. Khanmamedov, *Global attractors for von Karman equations with nonlinear interior dissipation*, J. Math. Anal. Appl., 318 (2006), pp. 92–101.
[25] A. Kh. Khanmamedov, *Existence of a global attractor for the plate equation with a critical exponent in an unbounded domain*, Appl. Math. Lett., 18 (2005), pp. 827–832.
[26] O. A. Ladyzhenskaya, *Attractors for Semigroups and Evolution Equations*, Lezioni Lincee, Cambridge University Press, Cambridge, UK, 1991.
[27] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Paris, 1969.
[28] S. S. Lu, H. Q. Wu, and C. K. Zhong, *Attractors for nonautonomous 2D Navier-Stokes equations with normal external forces*, Discrete Contin. Dyn. Syst., 13 (2005), pp. 701–719.
[29] G. Lukaszewicz and W. Sadowski, *Uniform attractor for 2D magneto-micropolar fluid flow in some unbounded domains*, Z. Angew. Math. Phys., 55 (2004), pp. 247–257.
[30] Q. F. Ma, S. H. Wang, and C. K. Zhong, *Necessary and sufficient conditions for the existence of global attractors for semigroups and applications*, Indiana Univ. Math. J., 51 (2002), pp. 1541–1557.
[31] A. Majda, *Introduction to PDEs and Waves for the Atmosphere and Oceans*, AMS, Providence, RI, 2003.
[32] I. Moise, R. Rosa, and X. Wang, *Attractors for noncompact nonautonomous systems via energy equations*, Discrete Contin. Dyn. Syst., 10 (2004), pp. 473–496.
[33] I. Moise, R. Rosa, and X. Wang, *Attractors for non-compact semigroup via energy equations*, Nonlinearity, 11 (1998), pp. 1369–1393.
[34] V. Pata and S. Zelik, *A remark on the damped wave equation*, Commun. Pure Appl. Anal., 5 (2006), pp. 609–614.
[35] J. C. Robinson, *Infinite-Dimensional Dynamical Systems. An Introduction to Dissipative Parabolic PDEs and the Theory of Global Attractors*, Cambridge University Press, Cambridge, UK, 2001.
[36] G. R. Sell and Y. You, *Dynamics of Evolutionary Equations*, Springer-Verlag, New York, 2002.
[37] C. Y. Sun, M. H. Yang, and C. K. Zhong, *Global attractors for the wave equation with nonlinear damping*, J. Differential Equations, 227 (2006), pp. 427–443.
[38] C. Y. Sun, M. H. Yang, and C. K. Zhong, *Global attractors for hyperbolic equations with critical exponent in locally uniform spaces*, J. Differential Equations, submitted.
[39] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag, New York, 1997.
[40] S. V. Zelik, *Asymptotic regularity of solutions of a nonautonomous damped wave equation with a critical growth exponent*, Commun. Pure Appl. Anal., 3 (2004), pp. 921–934.

[41] C. K. Zhong, M. H. Yang, and C. Y. Sun, *The existence of global attractors for the norm-to-weak continuous semigroup and application to the nonlinear reaction-diffusion equations*, J. Differential Equations, 223 (2006), pp. 367–399.

[42] S. F. Zhou and L. S. Wang, *Kernel sections for damped non-autonomous wave equations with critical exponent*, Discrete Contin. Dyn. Syst., 9 (2003), pp. 399–412.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 2, pp. 319–347

Nonlinear Convective Instability of Turing-Unstable Fronts near Onset: A Case Study*

Anna Ghazaryan† and Björn Sandstede‡

Abstract. Fronts are traveling waves in spatially extended systems that connect two different spatially homogeneous rest states. If the rest state behind the front undergoes a supercritical Turing instability, then the front will also destabilize. On the linear level, however, the front will be only convectively unstable since perturbations will be pushed away from the front as it propagates. In other words, perturbations may grow, but they can do so only behind the front. The goal of this paper is to prove for a specific model system that this behavior carries over to the full nonlinear system.

Key words. fronts, Turing bifurcation, nonlinear stability, convective instability

AMS subject classifications. Primary, 37L15; Secondary, 35B32, 35B35, 35K45, 35K57, 37L10

DOI. 10.1137/060670262

1. Introduction.

Propagating fronts are of interest in many different applications. In this manuscript, we are interested in the transition from stable to convectively unstable fronts: An initial perturbation to a convectively unstable front grows in time in any translation invariant norm but is simultaneously transported either to the right or to the left toward infinity in such a way that it decays pointwise at every fixed point in space as time goes to infinity [3]. It is worth pointing out that a convective instability can justifiably be viewed as a form of stability since perturbations decay pointwise: Consequently, we will use the terms "convective instability" and "convective stability" as synonyms. Whether or not an instability is convective depends strongly on the coordinate frame in which we measure the growth of perturbations. A natural reference frame for fronts is the comoving frame in which the front becomes stationary. Our aim is to show for a model system that the convective nature of certain front instabilities can be captured analytically.

The general situation can be described as follows. Consider a reaction-diffusion system

$$\partial_t U = D\partial_x^2 U + F(U;\alpha), \qquad x \in \mathbb{R}, \quad t > 0, \quad U \in \mathbb{R}^N, \tag{1.1}$$

with a control parameter $\alpha \in \mathbb{R}$, where D is a diagonal matrix with strictly positive entries, and F is a smooth function. We assume that the system exhibits a front, i.e., a nonlinear wave $U(x,t) = U_h(x-ct)$ that travels with positive speed $c > 0$ and connects two different homogeneous rest states $U_\pm$ so that $U_h(\xi) \to U_\pm$ as $\xi \to \pm\infty$. Once we specified that the

*Received by the editors September 19, 2006; accepted for publication by T. Kaper December 23, 2006; published electronically April 24, 2007. This work was supported by the NSF under grants DMS-9971703 and DMS-0203854.
http://www.siam.org/journals/siads/6-2/67026.html
†Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599 (ghazarya@email.unc.edu).
‡Department of Mathematics, University of Surrey, Guildford, GU2 7XH, UK (b.sandstede@surrey.ac.uk). This author gratefully acknowledges a Royal Society Wolfson Research Merit Award.

front velocity c is positive so that the front travels toward $x = \infty$, we may refer to the rest states U_+ and U_- as being, respectively, ahead of and behind the front.

We say that a front is stable if every solution that starts near the front converges to the front, or one of its spatial translates, as time goes to infinity. A sufficient criterion for stability is that the spectrum of the linearization $\mathcal{L}$ of (1.1) about the front, computed in the comoving frame $\xi = x - ct$, lies in the left half-plane except for a simple eigenvalue at the origin, which arises due to translational symmetry. Fronts become unstable when a subset of the spectrum of $\mathcal{L}$ crosses the imaginary axis. The effect of such instabilities on the dynamics near a given front depends on which part of the spectrum crosses the imaginary axis. If isolated eigenvalues cross the imaginary axis, then the problem can be analyzed using Lyapunov–Schmidt reduction or, alternatively, center-manifold theory. The two generic bifurcations that occur are saddle-nodes and Hopf bifurcations. At a Hopf bifurcation, a unique modulated front bifurcates, i.e., a solution that becomes time periodic in an appropriate comoving coordinate frame. It is also possible that part of the essential spectrum crosses the imaginary axis. The boundary of the essential spectrum of $\mathcal{L}$ coincides with the spectra of the asymptotic rest states $U_\pm$, and we concentrate here exclusively on Turing bifurcations of one of the asymptotic rest states: Turing bifurcations lead to stationary spatially periodic patterns whose deviation from the rest state is small; these patterns are commonly referred to as Turing patterns. If the rest state U_+ ahead of the front destabilizes, then there exists a continuum of modulated fronts which connect the rest state U_- behind the front with the Turing patterns ahead of the front [23]. The bifurcating modulated fronts are spectrally stable provided the periodic patterns are spectrally stable [22, 23]. For a certain model system that shares the main features of general reaction-diffusion systems, they have also been shown to be nonlinearly stable [10]. It should be emphasized that nonlinear stability does not follow from spectral stability, because the essential spectrum of the linearization about the modulated front touches the imaginary axis. The proof of nonlinear stability in [10] is based on exponential weights [4, 8, 9, 24] to handle the essential spectrum and on renormalization techniques [2, 7, 8] to take care of the nonlinearity.

In this paper, we are interested in the case where the rest state U_- behind the front destabilizes. It has been proved in [23] that modulated fronts that connect the Turing patterns behind the front to the rest state ahead cannot exist in this situation (see also [25] for related formal results). Thus, while the front is linearly unstable, there are no stable coherent front structures nearby. Numerical simulations and formal arguments give the following picture: The Turing bifurcation behind the front leads to stationary patterns. In the frame that moves with the front (which has speed c), we therefore expect, at least on the linear level, that initial perturbations to the front are transported with speed c to the left toward $x = -\infty$. In other words, the front can be thought of as pushing any perturbation away to the left. On the nonlinear level, we expect that growth saturates at the Turing pattern. Numerical simulations indeed show that initial data near the linearly unstable front evolve to a superposition of two fronts that move with different speeds, namely, a small Turing front which connects the Turing patterns far to the left with the unstable rest state U_-, and the primary linearly unstable front which travels faster and leaves the Turing front behind in its wake; see Figure 1. Sherratt [25] investigated in great detail the dynamics in the wake of convectively unstable fronts using formal arguments. Our goal is to make the above picture rigorous, at least for a model system similar to that considered in [10]: Our approach involves a priori estimates that we are

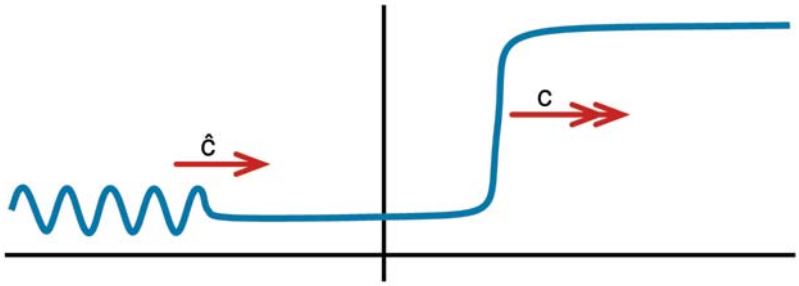


Figure 1. *A schematic illustration of the expected dynamics near a convectively unstable front is shown: The speeds satisfy $\hat{c} < c$.*

currently able to establish only in specific cases using restrictive tools such as the maximum principle and energy methods. We nevertheless believe that our general approach to nonlinear convective stability will apply more widely, which is why we carry out this case study. We certainly expect the overall phenomenon to be general for supercritical Turing bifurcations.

We consider the system

$$\begin{aligned} \partial_t u_1 &= \partial_x^2 u_1 + \frac{1}{2}(u_1 - c)(1 - u_1^2) + \gamma_1 u_2^2, \\ \partial_t u_2 &= -(1 + \partial_x^2)^2 u_2 + \alpha u_2 - u_2^3 - \gamma_2 u_2(1 + u_1), \end{aligned} \tag{1.2}$$

where $x \in \mathbb{R}$, $t \geq 0$, and $U = (u_1, u_2)$. The parameters $\gamma_1 \in \mathbb{R}$, $\gamma_2 > 0$, and $c \in (0, 1)$ are fixed, while the parameter α is a bifurcation parameter which varies near zero. For every α, the system (1.2) admits the traveling-wave solution

$$U_h(x - ct) = \begin{pmatrix} h(x - ct) \\ 0 \end{pmatrix}, \qquad h(\xi) = \tanh \frac{\xi}{2},$$

which connects the rest state $U_- = (-1, 0)$ at $x = -\infty$ with the rest state $U_+ = (1, 0)$ at $x = \infty$. The idea of considering the Chafee–Infante equation coupled to the Swift–Hohenberg equation is adopted from [10], where a similar system has been used to investigate the nonlinear stability of modulated fronts which bifurcate when the rest state ahead of the primary front becomes unstable.

A standard bifurcation argument (see, for instance, [4]) has been used in [11] to show that spatially periodic equilibria bifurcate at $\alpha = 0$ from the rest state U_-. More precisely, assume that the parameters γ_1 and γ_2 satisfy

$$\gamma_1 \gamma_2 > -\frac{3(1 + c)(5 + c)}{11 + 3c}; \tag{1.3}$$

then (1.2) has spatially periodic equilibria U_{per} for $\alpha > 0$ sufficiently close to zero which are given by

$$U_{\mathrm{per}}(x) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \sqrt{\frac{\alpha}{a_0}} \begin{pmatrix} 0 \\ \cos x \end{pmatrix} + \mathrm{O}(\alpha), \quad a_0 = \frac{3}{4} + \frac{\gamma_1 \gamma_2}{2} \left(\frac{1}{1 + c} + \frac{1}{2(5 + c)} \right). \tag{1.4}$$

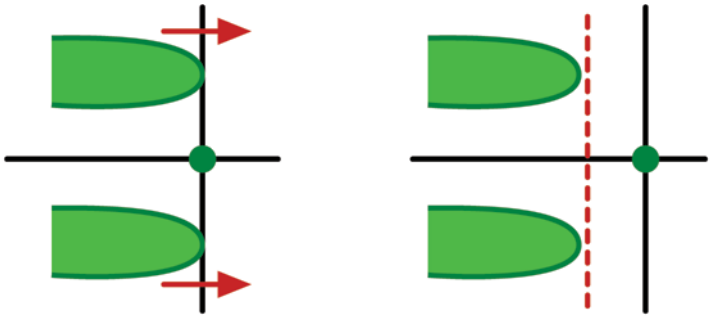


Figure 2. *A schematic picture of the spectrum of the front U_h of (1.2) in the comoving frame $\xi = x - ct$ is shown in the complex plane $\mathbb{C}$ for $\alpha = 0$ in spaces with (right) and without (left) exponential weight $e^{\beta\xi}$ with $\beta > 0$. Upon increasing α, the spectrum moves in the direction indicated by the arrows on spaces without exponential weights but stays to the left of the imaginary axis in spaces with exponential weight.*

In particular, the bifurcation is supercritical provided (1.3) holds. The following result shows that the periodic patterns are nonlinearly stable with respect to perturbations in the space $H^2(2)$ defined to be the set of L^2-functions for which the norm

$$\|U\|_{H^2(2)} := \left(\sum_{j=0}^{2} \int_{\mathbb{R}} |\partial_x^j U(x)|^2 (1+x^2)^2 \, \mathrm{d}x \right)^{\frac{1}{2}}$$

is finite.

Theorem 1 (see [11, Theorem 3.2]). *Assume that $\gamma_2 > 0$ and $c \in (0,1)$ are fixed and that (1.3) is met. For each $\alpha > 0$ sufficiently small, there are positive numbers K and δ such that, for every $V_0 \in H^2(2)$ with $\|V_0\|_{H^2(2)} \le \delta$, (1.2) with initial data $U_{\mathrm{per}} + V_0$ has a unique global solution $U(t) = U_{\mathrm{per}} + V(t)$, and $\|V(t)\|_{C^0} \le K(1+t)^{-1/2}$ for $t \ge 0$.*

The proof of the preceding theorem is essentially identical to that of [10, Theorem 2.4], where a slightly different system was studied, and we therefore refer the reader to [11] for details.

The front U_h exists for all values of α but it will be spectrally unstable for $\alpha > 0$, since part of its essential spectrum will then lie in the open right half-plane. To repeat the reasoning outlined above, we might expect that waves bifurcating from the front at $\alpha = 0$ resemble a pattern obtained by gluing together the front U_h and the Turing patterns U_{per} that emerge in its wake. Such waves would be time-periodic, rather than stationary, in a frame that moves with the front. It was shown though in [23] that, for small $\alpha > 0$, such waves cannot bifurcate. Thus, it is natural to ask how perturbations of the front will evolve in time for $\alpha > 0$. We shall see that the spectrum of the front can actually be moved into the left half-plane in the comoving frame $\xi = x - ct$, provided it is computed in an exponentially weighted function space with norm $\|e^{\beta\xi}U(\xi)\|$ for some appropriate $\beta > 0$; see Figure 2 for an illustration. Thus, if perturbations are localized ahead of the front, while being allowed to grow behind the front, then they will decay exponentially in time as $t \to \infty$. The main result of this paper asserts that the same statement is true for the full nonlinear problem: The front is only convectively unstable for $\alpha > 0$ in that perturbations are pushed away from the front toward its wake.

The results on nonlinear convective instability of the front U_h are formulated in the spaces $H^1_{\mathrm{ul}}(\mathbb{R},\mathbb{R}^2)$ of uniformly local functions (see [16, section 3.1]) whose definition we recall in section 2. These spaces contain, in particular, all differentiable bounded functions such as fronts or periodic solutions. The results furthermore utilize the smooth weight functions

$$\rho_\beta(x) := \begin{cases} \mathrm{e}^{\beta x}, & x \le -1, \\ 1, & x \ge 1, \end{cases} \tag{1.5}$$

defined for $\beta > 0$, with $\rho'_\beta(x) \ge 0$ for all x.

Theorem 2. *Assume that $\gamma_2 > 0$ and $c \in (0,1)$ are fixed, that (1.3) is met, and further that either $\gamma_1 \ge 0$ or else $\gamma_2 < \gamma_1 + \sqrt{2} < \sqrt{2}$. There are then positive constants α_*, β_*, ε_*, K, and Λ_* so that the following is true for all (α,ε) with $|\alpha| < \alpha_*$ and $0 < \varepsilon < \varepsilon_*$: For every function $V_0 = (v_1^0, v_2^0)$ with*

$$\|v_1^0\|_{H^1_{\mathrm{ul}}} \le \varepsilon^2, \qquad \|v_2^0\|_{H^1_{\mathrm{ul}}} \le \varepsilon, \qquad \|\rho_{\beta_*} V_0\|_{H^1_{\mathrm{ul}}} \le \varepsilon^2,$$

(1.2) with initial data $U_0 = U_h + V_0$ has a unique global solution $U(t)$, which can be expressed as

$$U(x,t) = U_h(x - ct - q(t)) + V(x,t)$$

for an appropriate real-valued function $q(t)$, and there is a $q_* \in \mathbb{R}$ so that

$$\|V(\cdot,t)\|_{H^1_{\mathrm{ul}}} + |q(t)| \le K\left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}}, \quad \|\rho_{\beta_*}(\cdot - ct)V(\cdot,t)\|_{H^1_{\mathrm{ul}}} + |q(t) - q_*| \le K\mathrm{e}^{-\Lambda_* t} \tag{1.6}$$

for $t \ge 0$.

Upon setting $\eta_* := \epsilon_*^2/2$ and $K_* := K[\varepsilon_* + \sqrt{\alpha_*}]^{\frac{1}{2}}/\eta_*$, we obtain the following slightly weaker, but also less technical, corollary of Theorem 2 which we formulate in the comoving frame.

Corollary 1. *Under the assumptions of Theorem 2, there are positive constants α_*, β_*, η_*, K_*, and Λ_* so that the following is true for all α with $|\alpha| < \alpha_*$: For each function V_0 with $\|V_0\|_{H^1_{\mathrm{ul}}} \le \eta_*$, (1.2) with initial data $U_0 = U_h + V_0$ has a unique global solution $U(t)$, which can be expressed as*

$$U(x,t) = U_h(x - ct - q(t)) + V(x - ct, t)$$

for an appropriate real-valued function $q(t)$, and there is a $q_* \in \mathbb{R}$ so that

$$\|V(\cdot,t)\|_{H^1_{\mathrm{ul}}} + |q(t)| \le K_*\eta_*, \qquad \|\rho_{\beta_*}(\cdot)V(\cdot,t)\|_{H^1_{\mathrm{ul}}} + |q(t) - q_*| \le K_*\mathrm{e}^{-\Lambda_* t}$$

for $t \ge 0$.

Thus, the conclusion of the preceding theorem and corollary is that the dynamical behavior of the front does not change at all near $\alpha = 0$ provided we measure perturbations in the weighted norm: Perturbations stay bounded in the C^0-norm and decay exponentially to zero as $t \to \infty$ when they are multiplied by $\mathrm{e}^{\beta_*(x-ct)}$ for some appropriate $\beta_* > 0$, so that the front is nonlinearly stable in this norm for all values of α near zero. Note that our results say nothing about the detailed dynamics behind the front. Indeed, our approach, outlined in

detail below, relies only on a priori estimates and does not take the specific dynamics behind the front into account.

We comment briefly on the scalings in ε and α that appear in Theorem 2. The components v_1 and v_2 of the perturbation V scale with different powers in ε because the instability manifests itself on a linear level only in the v_2-component, while the v_1-component is affected only through the quadratic nonlinearity. The estimates (1.6) for the perturbation V are certainly not optimal as we expect solutions to saturate at order $|\alpha|^{\frac{1}{2}}$. The weaker estimates (1.6) are an artifact of our method which requires a supercritical bifurcation, but not necessarily its genericity, and which consequently will not yield sharp estimates.

As already mentioned, nonlinear stability of the front U_h in the weighted spaces cannot be inferred from spectral stability because the nonlinearity does not map the weighted spaces into themselves. Indeed, if we define $W = \mathrm{e}^{\beta x}V$ and use $W = (w_1, w_2)$ as the new dependent variable, then we would like to find bounds for W in C^0 or H^1_{ul}. If we transform the equation for the initial perturbation V to the new weighted variable W, then the nonlinear term u_1^n becomes

$$\mathrm{e}^{\beta x}[\mathrm{e}^{-\beta x}w_1]^n = \mathrm{e}^{(1-n)\beta x}w_1^n,$$

which is unbounded as $x \to -\infty$ for $n > 1$. To overcome this difficulty, we use a method introduced originally in [19] in the Hamiltonian context. If we can obtain a priori estimates for the solution in the space without weight, for instance, in C^0 or H^1_{ul}, and show that it stays bounded and sufficiently small, then the nonlinear terms u_j^n, written as $u_j^n = u_j^{n-1}u_j$, become $u_j^{n-1}w_j$ when transformed to the weighted functions W, which are now well behaved due to the a priori estimates for u_j. This interplay of the spatially uniform norm and the exponentially weighted norm is the key for the proof of nonlinear stability of the front. An example of a successful application of this technique has also been given independently in [1], where a reaction-diffusion-convection system is considered that has essential spectra up to the imaginary axis for all values of the bifurcation parameter while an isolated pair of simple eigenvalues crosses the imaginary axis at the bifurcation point.

The plan of the paper is as follows. In section 2, we discuss the spectral stability of the front and state several auxiliary results that we need later. Section 3 contains the proof of Theorem 2. Numerical simulations and some further implications of our results are given in section 4, and we end with conclusions and a discussion in section 5.

2. Linear convective instability. We begin by introducing the spaces $L^2_{\mathrm{ul}}(\mathbb{R})$ from [16] in which we shall work. Pick any positive and bounded function $\sigma \in C^2(\mathbb{R})$ for which $\int_{\mathbb{R}} \sigma(x)\,\mathrm{d}x = 1$ and $|\sigma'(x)|, |\sigma''(x)| \le \sigma(x)$ for $x \in \mathbb{R}$: We may, for instance, set $\sigma(x) = \frac{1}{\pi}\operatorname{sech} x$. For each $0 < b < 1$, we define $\sigma_b(x) := \sigma(bx)$ and record that $\int_{\mathbb{R}} \sigma_b(x)\,\mathrm{d}x = 1/b$.

Using the weight function σ, we define the Banach space L^2_{ul} of uniformly local weighted L^2 functions to be

$$L^2_{\mathrm{ul}}(\mathbb{R}) = \left\{ u \in L^2_{\mathrm{loc}}(\mathbb{R}) :\ \|u\|^2_{L^2_{\mathrm{ul}}} := \sup_{y\in\mathbb{R}} \int_{\mathbb{R}} \sigma(x+y)|u(x)|^2\,\mathrm{d}x < \infty \right.$$
$$\left. \text{and } \|T_y u - u\|_{L^2_{\mathrm{ul}}} \to 0 \text{ as } y \to 0 \right\},$$

where $[T_y u](x) := u(x+y)$ is the translation operator. We denote the associated Sobolev spaces by $H^k_{\mathrm{ul}}(\mathbb{R})$ and remark that different choices for σ result in the same spaces with equivalent norms. We collect various properties of these spaces in the following lemma.

Lemma 2.1 (see [16, Lemmas 3.1 and 3.8]). *There is a constant K_0 with the following properties:*

(i) *H^1_{ul} is an algebra and embeds continuously into C^0_{unif} with $\|u\|^2_{C^0} \le K_0 \|u\|_{L^2_{\mathrm{ul}}} \|u\|_{H^1_{\mathrm{ul}}}$ for all $u \in H^1_{\mathrm{ul}}$.*

(ii) *For each $0 < b < 1$, let $\sigma_b(x) := \sigma(bx)$; then $\|u\|^2_{L^2_{\mathrm{ul}}(\sigma)} \le K_0(1+b)\|u\|^2_{L^2_{\mathrm{ul}}(\sigma_b)}$ for all $u \in L^2_{\mathrm{ul}}(\sigma)$.*

(iii) *We have $-\int_{\mathbb{R}} \sigma_b u(1+\partial_x^2)^2 u \,\mathrm{d}x \le \frac{7b^2}{2} \int_{\mathbb{R}} \sigma_b u^2 \,\mathrm{d}x$ for all $u \in H^4_{\mathrm{ul}}$.*

We return now to the partial differential equation (1.2). Upon transforming (1.2) into the comoving coordinate $\xi = x - ct$, we obtain the system

$$\begin{aligned} \partial_t u_1 &= \partial_\xi^2 u_1 + c\partial_\xi u_1 + \frac{1}{2}(u_1 - c)(1 - u_1^2) + \gamma_1 u_2^2, \\ \partial_t u_2 &= -(1+\partial_\xi^2)^2 u_2 + c\partial_\xi u_2 + \alpha u_2 - u_2^3 - \gamma_2 u_2(1+u_1). \end{aligned} \tag{2.1}$$

The linearization of (2.1) about a stationary solution of the form $U_* = (u_*, 0)$ is given by the diagonal operator

$$\mathcal{L}_0[U_*] := \begin{pmatrix} \partial_\xi^2 + c\partial_\xi + \frac{1}{2}(1 + 2cu_* - 3u_*^2) & 0 \\ 0 & -[1+\partial_\xi^2]^2 + c\partial_\xi + \alpha - \gamma_2(1+u_*) \end{pmatrix}. \tag{2.2}$$

The operator $\mathcal{L}_0[U_*]$ is sectorial on $\mathcal{X}_0 := H^1_{\mathrm{ul}} \times H^1_{\mathrm{ul}}$ with dense domain $H^3_{\mathrm{ul}} \times H^5_{\mathrm{ul}}$. We shall also consider (2.2) in exponentially weighted spaces: For $\beta > 0$, we defined in (1.5) the weight function

$$\rho_\beta(\xi) = \begin{cases} \mathrm{e}^{\beta\xi}, & \xi \le -1, \\ 1, & \xi \ge 1, \end{cases}$$

where $\rho'_\beta(\xi) \ge 0$ for all ξ. We then set

$$W(\xi) := \rho_\beta(\xi) V(\xi)$$

so that W satisfies $W_t = \mathcal{L}_\beta[U_*]W$ with

$$\mathcal{L}_\beta[U_*] = \rho_\beta \mathcal{L}_0[U_*] \rho_\beta^{-1}$$

whenever V satisfies $V_t = \mathcal{L}_0[U_*]V$. It is easy to check that the operator $\mathcal{L}_\beta[U_*]$ is again sectorial on $\mathcal{X}_0$. From now on, we shall denote by $\mathcal{L}_\beta := \mathcal{L}_\beta[U_h]$ the linearized operator belonging to the front U_h.

Proposition 2.2. *Given $\gamma_2 > 0$ and $c \in (0,1)$, there are positive numbers α_0 and β_0 and a strictly positive function $\Lambda_0(\beta)$ defined for $0 < \beta < \beta_0$ so that the following holds for $|\alpha| \le \alpha_0$: The spectrum of $\mathcal{L}_\beta$ satisfies*

$$\mathrm{spec}(\mathcal{L}_\beta) = \{0\} \cup \Sigma \qquad \textit{with} \qquad \mathrm{Re}\,\Sigma \le -\Lambda_0(\beta),$$

and $\lambda = 0$ *is the simple eigenvalue of* $\mathcal{L}_\beta$ *and* $\mathcal{L}_0$*. Furthermore, the spectrum of* $\mathcal{L}_0$ *satisfies*

$$\begin{aligned} \operatorname{spec}(\mathcal{L}_0) &= \{0\} \cup \Sigma \text{ with } \operatorname{Re}\Sigma < 0 \text{ for } \alpha < 0, \\ \operatorname{spec}(\mathcal{L}_0) \cap \{\lambda : \operatorname{Re}\lambda \geq 0\} &= \{0\} \cup \{\pm \mathrm{i}\} \text{ for } \alpha = 0, \end{aligned}$$

and $\operatorname{spec}(\mathcal{L}_0) \cap \{\lambda : \operatorname{Re}\lambda > 0\} \neq \emptyset$ *for* $\alpha > 0$.

In particular, the front U_h is orbitally stable for $\alpha < 0$ due to [12, section 5.1], while it is spectrally unstable for $\alpha > 0$.

Proof. For each β, the spectrum of $\mathcal{L}_\beta$ on $\mathcal{X}_0$ is the disjoint union of the essential spectrum and the point spectrum, where the latter consists, by definition, of all isolated eigenvalues with finite multiplicity. It follows from [12, appendix to section 5] or [17] that the essential spectrum of $\mathcal{L}_\beta$ on $\mathcal{X}_0$ is, for any $\beta \geq 0$, bounded to the right by the essential spectra of the asymptotic operators

$$\mathcal{L}_\beta^- := \mathcal{L}_\beta[U_-] = \begin{pmatrix} (\partial_\xi - \beta)^2 + c(\partial_\xi - \beta) - (1+c) & 0 \\ 0 & -[1 + (\partial_\xi - \beta)^2]^2 + c(\partial_\xi - \beta) + \alpha \end{pmatrix}$$

and $\mathcal{L}_0^+ := \mathcal{L}_0[U_+]$. Indeed, the weight function ρ_β is equal to one for $\xi \geq 1$ and therefore has no effect on the asymptotic coefficients when $\xi \to \infty$.

Thus, to determine the rightmost elements in the essential spectrum of $\mathcal{L}_0$, it suffices to compute the essential spectra of the operators $\mathcal{L}_0^\pm$ on the space $\mathcal{X}_0$. On account of multiplier theory [16, Lemma 3.3], these spectra can be calculated using the Fourier transform: A complex number λ is in the spectrum of $\mathcal{L}_0^\pm$ if and only if there are a vector $V_0 \in \mathbb{C}^2$ and a number $k \in \mathbb{R}$ so that

$$\lambda \mathrm{e}^{\mathrm{i}k\xi} V_0 = \mathcal{L}_0^\pm \mathrm{e}^{\mathrm{i}k\xi} V_0,$$

that is, if and only if

$$\det \begin{pmatrix} -k^2 + \mathrm{i}kc + \frac{1}{2}(1 + 2cu_\pm - 3u_\pm^2) - \lambda & 0 \\ 0 & -(1-k^2)^2 + \mathrm{i}kc + \alpha - \gamma_2(1+u_\pm) - \lambda \end{pmatrix} = 0,$$

where $u_\pm = \pm 1$. In particular, we see that the spectrum of $\mathcal{L}_0^+$ is given by

$$\begin{aligned} \operatorname{spec}(\mathcal{L}_0^+) = \big\{ &\lambda \in \mathbb{C};\ \lambda = \lambda_1^+(k) := -k^2 + \mathrm{i}kc - (1-c) \text{ or} \\ &\lambda = \lambda_2^+(k) := -(1-k^2)^2 + \mathrm{i}kc + \alpha - 2\gamma_2 \text{ for some } k \in \mathbb{R} \big\} \end{aligned}$$

and therefore lies in the left half-plane and is uniformly bounded away from the imaginary axis for all α near zero. Similarly, the spectrum of the operator $\mathcal{L}_0^-$ associated with the rest state behind the front is given by

$$\begin{aligned} \operatorname{spec}(\mathcal{L}_0^-) = \big\{ &\lambda \in \mathbb{C};\ \lambda = \lambda_1^-(k) := -k^2 + \mathrm{i}kc - (1+c) \text{ or} \\ &\lambda = \lambda_2^-(k) := -(1-k^2)^2 + \mathrm{i}kc + \alpha \text{ for some } k \in \mathbb{R} \big\}. \end{aligned} \tag{2.3}$$

It lies in the left half-plane, bounded away from the imaginary axis, except for the curve $\lambda = \lambda_2^-(k)$ which crosses into the right half-plane for $\alpha \geq 0$ and $k \in \mathbb{R}$ near $k_\mathrm{c} = \pm 1$.

The spectrum of $\mathcal{L}_\beta^-$ can be computed either analogously or, more directly, by replacing k with $k+\mathrm{i}\beta$ in the above expression for $\operatorname{spec}(\mathcal{L}_0^-)$. The rightmost part of the spectrum of $\mathcal{L}_\beta^-$ is therefore given by the linear dispersion curve

$$\lambda = \lambda_2^-(k+\mathrm{i}\beta) = -(1-k^2+\beta^2)^2 + \alpha - c\beta + 4\beta^2k^2 + \mathrm{i}[ck - 4\beta k(1-k^2-\beta^2)]$$

for $k \in \mathbb{R}$, and we have

$$\max_{k\in\mathbb{R}} \operatorname{Re}\lambda_2^-(k+\mathrm{i}\beta) = \alpha - c\beta + 4\beta^2(1+2\beta^2), \tag{2.4}$$

which is achieved at $k = \pm\sqrt{1+3\beta^2}$. Choosing $\beta = \frac{c}{8}$, we obtain the bound

$$\Lambda_{\mathrm{ess}}^- = \alpha - \frac{c^2}{16}\left(1 - \frac{c^2}{32}\right) < 0$$

for the maximal real part of $\operatorname{spec}(\mathcal{L}_\beta^-)$, which is strictly negative for fixed $c \in (0,1)$ and $|\alpha| \leq \frac{c^2}{32}$.

In summary, the essential spectrum of $\mathcal{L}_0$ lies in the open left half-plane for $\alpha = 0$, touches the imaginary axis at $\lambda = \pm\mathrm{i}$ when $\alpha = 0$, and crosses into the right half-plane for $\alpha > 0$.

Having discussed the essential spectrum, we now turn to the point spectrum. The situation here is similar to the one considered in [10]. The eigenfunctions associated with isolated eigenvalues of $\mathcal{L}_0$ necessarily decay exponentially as $|\xi| \to \infty$. The origin $\lambda = 0$ is always in the point spectrum of $\mathcal{L}_0$ with eigenfunction $U_h'(\xi) = (h_\xi(\xi), 0)$.

For $\alpha < 0$, any isolated eigenvalue λ of $\mathcal{L}_0$ satisfies either $\operatorname{Re}\lambda < 0$ or $\lambda = 0$. To prove this claim, we assume that there is an eigenvalue λ with eigenfunction $V = (v_1, v_2)$ which therefore satisfies the decoupled system

$$\lambda v_1 = \partial_\xi^2 v_1 + c\partial_\xi v_1 + \frac{1}{2}(1 + 2ch - 3h^2)v_1, \tag{2.5}$$

$$\lambda v_2 = -(1+\partial_\xi^2)^2 v_2 + c\partial_\xi v_2 + \alpha v_2 - \gamma_2(1+h)v_2. \tag{2.6}$$

We see that $\lambda = 0$ is an eigenvalue of (2.5) with positive eigenfunction $h_\xi(\xi) = \frac{1}{2}\operatorname{sech}^2\frac{\xi}{2}$. Sturm–Liouville theory implies that $\lambda = 0$ is simple for (2.5) and that all other eigenvalues of (2.5) are strictly negative. To analyze (2.6), we multiply by $\bar{v}_2$ and integrate over $\mathbb{R}$ to obtain

$$\operatorname{Re}\lambda \|v_2\|_{L^2}^2 \leq -\|(1+\partial_\xi^2)v_2\|_{L^2}^2 + \alpha\|v_2\|_{L^2}^2 \leq \alpha\|v_2\|_{L^2}^2,$$

where we used that $\gamma_2(1+h(\xi)) \geq 0$. Thus, either $\operatorname{Re}\lambda \leq \alpha$ or $v_2 = 0$, which proves the claim.

Next, we consider the isolated eigenvalues of $\mathcal{L}_\beta$ for $0 < \beta < \beta_0$ for an appropriate $\beta_0 > 0$. We claim that there are no eigenvalues on or to the right of the imaginary axis for all α with $|\alpha|$ sufficiently small, except for a simple eigenvalue at the origin. To prove this claim, we first record that eigenfunctions associated with isolated eigenvalues of $\mathcal{L}_\beta$ in the closed right half-plane decay exponentially as $\xi \to \infty$ with a rate that does not depend on the eigenvalue. In particular, there is a $\beta_0 > 0$ so that the following is true for each $0 < \beta \leq \beta_0$: if $(w_1, w_2) = \rho_\beta(v_1, v_2)$ is an L^2-eigenfunction of $\mathcal{L}_\beta$ belonging to an eigenvalue λ

with $\operatorname{Re}\lambda \geq 0$, then $e^{\beta\xi}(v_1, v_2)$ will also be in L^2. Thus, it suffices to prove the claim for the operator $e^{\beta\xi}\mathcal{L}_0(\partial_\xi)e^{-\beta\xi} = \mathcal{L}_0(\partial_\xi - \beta)$: the associated eigenvalue problem is given by

$$\lambda w_1 = (\partial_\xi - \beta)^2 w_1 + c(\partial_\xi - \beta)w_1 + \frac{1}{2}(1 + 2ch - 3h^2)w_1, \tag{2.7}$$

$$\lambda w_2 = -(1 + (\partial_\xi - \beta)^2)^2 w_2 + c(\partial_\xi - \beta)w_2 + \alpha w_2 - \gamma_2(1 + h)w_2. \tag{2.8}$$

Multiplying (2.8) by $\bar{w}_2$ and integrating over $\mathbb{R}$, we obtain

$$\operatorname{Re}\lambda \|w_2\|_{L^2} \leq (\alpha - c\beta)\|w_2\|_{L^2} \leq 0,$$

and therefore either $\operatorname{Re}\lambda \leq \alpha - c\beta < 0$ or $w_2 = 0$. It remains to consider (2.7), which has an eigenvalue at the origin with bounded positive eigenfunction $e^{\beta\xi}h_\xi(\xi)$. Sturm–Liouville theory implies again that all other eigenvalues are strictly negative: In fact, the largest negative eigenvalue is equal to $-\frac{3}{4}(1 - c^2)$. Thus, for eigenvalues λ of (2.7)–(2.8), we have $\lambda = 0$ or $\operatorname{Re}\lambda \leq \max\{\alpha - c\beta, -\frac{3}{4}(1 - c^2)\} < 0$, and as mentioned above the same statement holds for the eigenvalues of $\mathcal{L}_\beta$.

Finally, we remark that solutions V of (2.5)–(2.6) and W of (2.7)–(2.8) are in one-to-one correspondence via $W(\xi) = e^{\beta\xi}V(\xi)$. This shows that $\mathcal{L}_0$ cannot have any isolated eigenvalues in the closed right half-plane except at $\lambda = 0$. ∎

In the nonlinear stability analysis of our model, we need the semigroup estimates for the operators

$$\mathcal{A}_1 := \partial_x^2 - (1 + c), \qquad \mathcal{A}_2 := -(1 + \partial_x^2)^2$$

provided by the following lemma, which is a straightforward application of multiplier theory [16, Lemma 3.3].

Lemma 2.3. *The operators $\mathcal{A}_1$ and $\mathcal{A}_2$ are sectorial and thus generate holomorphic semigroups $e^{\mathcal{A}_1 t}$ and $e^{\mathcal{A}_2 t}$. There is a positive constant K_0 with*

$$\begin{aligned} \|e^{\mathcal{A}_1 t}\|_{L^2_{\mathrm{ul}} \to H^1_{\mathrm{ul}}} &\leq K_0(1 + t^{-\frac{1}{4}})e^{-t}, & \|e^{\mathcal{A}_1 t}\|_{H^1_{\mathrm{ul}} \to H^1_{\mathrm{ul}}} &\leq K_0 e^{-t}, \\ \|e^{\mathcal{A}_2 t}\|_{L^2_{\mathrm{ul}} \to H^s_{\mathrm{ul}}} &\leq K_0(1 + t^{-\frac{s}{4}}), & \|e^{\mathcal{A}_2 t}\|_{H^1_{\mathrm{ul}} \to H^1_{\mathrm{ul}}} &\leq K_0 \end{aligned}$$

uniformly in $t > 0$.

3. Nonlinear convective instability. This section contains the proof of Theorem 2. We want to show that the front is convectively stable in the comoving frame for initial perturbations which are small in H^1_{ul}. In the comoving frame $\xi = x - ct$, the front is a stationary solution of

$$\begin{aligned} \partial_t u_1 &= \partial_\xi^2 u_1 + c\partial_\xi u_1 + \frac{1}{2}(u_1 - c)(1 - u_1^2) + \gamma_1 u_2^2, \\ \partial_t u_2 &= -(1 + \partial_\xi^2)^2 u_2 + c\partial_\xi u_2 + \alpha u_2 - u_2^3 - \gamma_2 u_2(1 + u_1). \end{aligned} \tag{3.1}$$

The proof of Theorem 2 is divided into two parts. First we show that suitable a priori estimates imply the nonlinear stability of the front in appropriate exponentially weighted norms imposed in the comoving frame. Afterward, we establish these a priori estimates. We focus exclusively on the case $\gamma_1 \geq 0$ and refer the reader to [11] for the modifications that are necessary for the case $\gamma_2 < \gamma_1 + \sqrt{2} < \sqrt{2}$. Recall that $\gamma_2 > 0$ and $c \in (0, 1)$ are both fixed.

3.1. A priori estimates imply nonlinear stability. We expect that initial data close to the front will converge to an appropriate translate of the front but not necessarily to the primary front U_h itself. To capture this behavior, we introduce a time-dependent spatial shift function $q(t)$ in the argument of the front U_h and write solutions to (3.1) as

$$U(\xi,t)=\begin{pmatrix} u_1(\xi,t)\\ u_2(\xi,t)\end{pmatrix}=\begin{pmatrix} h(\xi-q(t))\\ 0\end{pmatrix}+\begin{pmatrix} v_1(\xi,t)\\ v_2(\xi,t)\end{pmatrix}, \tag{3.2}$$

with $h(\xi)=\tanh\frac{\xi}{2}$. We may assume that $q(0)=0$ since our system is translationally invariant. The decomposition (3.2) can be made unique by requiring that the perturbation $V=(v_1,v_2)$ be "perpendicular," in an appropriate way that we specify below, to the one-dimensional subspace spanned by the derivative of the front.

The perturbation $V=(v_1,v_2)$ of the front satisfies the system

$$\begin{aligned}\partial_t v_1 &= \partial_\xi^2 v_1 + c\partial_\xi v_1 + \frac{1}{2}[1-3h^2(\xi-q(t))+2ch(\xi-q(t))]v_1+\frac{1}{2}[c-3h(\xi-q(t))]v_1^2\\ &\quad -\frac{1}{2}v_1^3+\dot q(t)h_\xi(\xi-q(t))+\gamma_1 v_2^2,\\ \partial_t v_2 &= -(1+\partial_\xi^2)^2 v_2+c\partial_\xi v_2+\alpha v_2-v_2^3-\gamma_2(1+h(\xi-q(t)))v_2-\gamma_2 v_1 v_2\end{aligned} \tag{3.3}$$

with initial data $v_1(\xi,0)=v_1^0(\xi)$, $v_2(\xi,0)=v_2^0(\xi)$, and $q(0)=0$. Using the notation

$$\begin{aligned}\mathcal{A}&=\begin{pmatrix}\partial_\xi^2+c\partial_\xi & 0\\ 0 & -(1+\partial_\xi^2)^2+c\partial_\xi+\alpha\end{pmatrix},\\ \mathcal{R}(\xi)&=\begin{pmatrix}\mathcal{R}_1 & 0\\ 0 & \mathcal{R}_2\end{pmatrix}=\begin{pmatrix}\frac{1}{2}[1-3h^2(\xi)+2ch(\xi)] & 0\\ 0 & -\gamma_2(1+h(\xi))\end{pmatrix},\\ \mathcal{N}(V)&=\begin{pmatrix}\mathcal{N}_1(V)\\ \mathcal{N}_2(V)\end{pmatrix}=\begin{pmatrix}\frac{1}{2}[c-3h(\xi-q(t))]v_1(\xi,t)-\frac{1}{2}v_1^2(\xi,t) & \gamma_1 v_2(\xi,t)\\ 0 & -v_2^2(\xi,t)-\gamma_2 v_1(\xi,t)\end{pmatrix},\end{aligned}$$

system (3.3) becomes

$$\partial_t V=\mathcal{A}V+\mathcal{R}(\xi-q(t))V+\mathcal{N}(V)V+\dot q(t)h_\xi(\xi-q(t))e_1,\qquad e_1=(1,0). \tag{3.4}$$

Next, we introduce the weighted solution $W=(w_1,w_2)$ via

$$W(\xi,t)=\rho_\beta(\xi)V(\xi,t),$$

with ρ_β as defined in (1.5), which satisfies the system

$$\partial_t W=\mathcal{L}_\beta W+[\mathcal{R}(\xi-q(t))-\mathcal{R}(\xi)]W+\mathcal{N}(V)W+\dot q(t)h_\xi(\xi-q(t))\rho_\beta(\xi)e_1 \tag{3.5}$$

with $\mathcal{L}_\beta=\rho_\beta\mathcal{A}\rho_\beta^{-1}+\mathcal{R}(\xi)$ being the linearization of the front U_h discussed in section 2 whenever $V(\xi,t)$ satisfies (3.4).

Throughout the remainder of the proof, we fix β with $0<\beta<\beta_0$ as in Proposition 2.2: We then know that $\lambda=0$ is a simple isolated eigenvalue of $\mathcal{L}_\beta$ with eigenfunction $\rho_\beta(\xi)\partial_\xi U_h$ and the rest of the spectrum has real part less than Λ_0 with Λ_0 from Proposition 2.2. We define

$\mathcal{P}^{\mathrm{c}}_\beta : H^1_{\mathrm{ul}} \times H^1_{\mathrm{ul}} \to H^1_{\mathrm{ul}} \times H^1_{\mathrm{ul}}$ to be the spectral projection onto the one-dimensional eigenspace of $\mathcal{L}_\beta$ corresponding to the zero eigenvalue and denote by $\mathcal{P}^{\mathrm{s}}_\beta = 1 - \mathcal{P}^{\mathrm{c}}_\beta$ the complementary projection onto the stable eigenspace.

Lemma 3.1. *For $0 < \beta < \beta_0$, there are constants $K_0 > 0$ and $\alpha_0 > 0$ such that the following is true for any α with $|\alpha| < \alpha_0$. The spectral projection $\mathcal{P}^{\mathrm{c}}_\beta$ is given by*

$$\mathcal{P}^{\mathrm{c}}_\beta W = \begin{pmatrix} P^{\mathrm{c}}_\beta & 0 \\ 0 & 0 \end{pmatrix} W = \langle \psi^{\mathrm{c}}_1, W_1 \rangle_{L^2} \rho_\beta \partial_\xi U_h, \tag{3.6}$$

where

$$\psi^{\mathrm{c}}_1(\xi) = \frac{\mathrm{e}^{c\xi} \rho_\beta(\xi) h_\xi(\xi)}{\int_{\mathbb{R}} \mathrm{e}^{c\zeta} h_\xi(\zeta)^2 \, \mathrm{d}\zeta},$$

and we have

$$\| \mathrm{e}^{\mathcal{P}^{\mathrm{c}}_\beta \mathcal{L}_\beta t} \|_{H^1_{\mathrm{ul}}} \le K_0 \mathrm{e}^{-\Lambda_0 t}, \qquad t \ge 0, \tag{3.7}$$

with Λ_0 as in Proposition 2.2.

Proof. It is easy to check that, in the space of bounded functions, the kernel of the operator adjoint to $\mathcal{L}_\beta$ is spanned by $(h_\xi(\xi)\mathrm{e}^{c\xi}\rho_\beta(\xi), 0)$. Upon normalizing this function, we end up with the expression (3.6) for the center projection. The estimate (3.7) is a consequence of Proposition 2.2 once we observe that the constant K_0 does not depend on α despite the presence of α in the definition of $\mathcal{L}_\beta$. Indeed, when $\alpha = 0$, the spectrum of $\mathcal{P}^{\mathrm{c}}_\beta \mathcal{L}_\beta$ belongs to $\{\lambda \in \mathbb{C} : \operatorname{Re}\lambda < -\beta c\}$, and an estimate of the form (3.7) holds for some K_0. The operator for $\alpha \neq 0$ is a bounded perturbation of order $\mathrm{O}(\alpha)$ of the $\alpha = 0$ operator, and [18, Theorem 1.1] implies that K_0 can be chosen to be independent of α for α sufficiently close to zero. ∎

To fix $q(t)$, we require that $\mathcal{P}^{\mathrm{c}}_\beta W(t) = 0$ for all t for which the decomposition (3.2) exists. In other words, we require that $W(t) \in \operatorname{Range}(\mathcal{P}^{\mathrm{s}}_\beta)$ for all t. Applying the projections $\mathcal{P}^{\mathrm{c}}_\beta$ and $\mathcal{P}^{\mathrm{s}}_\beta$ to (3.5), we obtain the evolution system

$$\partial_t V = \mathcal{A} V + \mathcal{R}(\xi - q(t))V + \mathcal{N}(V)V + \dot{q}(t) h_\xi(\xi - q(t)) e_1, \tag{3.8}$$

$$\partial_t W = \mathcal{P}^{\mathrm{s}}_\beta \mathcal{L}_\beta W + \mathcal{P}^{\mathrm{s}}_\beta \left([\mathcal{R}(\xi - q(t)) - \mathcal{R}(\xi)]W + \mathcal{N}(V)W + \dot{q}(t) h_\xi(\xi - q(t)) \rho_\beta(\xi) e_1 \right), \tag{3.9}$$

$$\dot{q}(t) = -\frac{\langle \psi^{\mathrm{c}}_1, [\mathcal{R}_1(\xi - q(t)) - \mathcal{R}_1(\xi)]W_1 + \mathcal{N}_1(V)W \rangle_{L^2}}{\langle \psi^{\mathrm{c}}_1, h_\xi(\xi - q(t))\rho_\beta(\xi) \rangle_{L^2}} \tag{3.10}$$

for $V = (v_1, v_2)$, $W = (w_1, w_2)$, and q. It is easy to see that the linear parts of the right-hand sides in (3.8)–(3.9) are sectorial operators on $H^1_{\mathrm{ul}}(\mathbb{R}, \mathbb{R}^2)$ with dense domain $H^3_{\mathrm{ul}} \times H^5_{\mathrm{ul}}$. The nonlinearity is smooth from $\mathcal{Y} := H^1_{\mathrm{ul}}(\mathbb{R}, \mathbb{R}^2) \times H^1_{\mathrm{ul}}(\mathbb{R}, \mathbb{R}^2) \times \mathbb{R}$ into itself, and there is a constant K_1 such that

$$\| \mathcal{R}_1(\cdot - q) - \mathcal{R}_1(\cdot) \|_{H^1_{\mathrm{ul}}} + \| \mathcal{N}(V) \|_{H^1_{\mathrm{ul}}} \le K_1 (|q| + \|V\|_{H^1_{\mathrm{ul}}}) \tag{3.11}$$

and

$$|\dot{q}| \le K_1 (|q| + \|V\|_{H^1_{\mathrm{ul}}}) \|W\|_{H^1_{\mathrm{ul}}} \tag{3.12}$$

for all $(V, W, q) \in \mathcal{Y}$ with norm less than one, say. We therefore have the methods introduced in [12] at our disposal which give local existence and uniqueness of solutions for initial data in $\mathcal{Y}$ as well as continuous dependence on initial conditions, thus proving local existence and uniqueness of the decomposition (3.2).

These arguments also allow us to claim that, for each given $0 < \eta_0 \leq 1$, there exist a $\delta_0 > 0$ and a time $T > 0$ such that the decomposition (3.2) exists for $0 \leq t < T$ with

$$|q(t)| + \|V(t)\|_{H^1_{\mathrm{ul}}} \leq \eta_0 \tag{3.13}$$

provided $\|V(0)\|_{H^1_{\mathrm{ul}}} \leq \delta_0$. Let $T_{\max} = T_{\max}(\eta_0)$ be the maximal time for which (3.13) holds.

Lemma 3.2. *Pick Λ with $0 < \Lambda < \Lambda_0$ and $\hat{\eta}_0 > 0$ so that*

$$\frac{2K_0K_1(1+K_0)}{\Lambda_0 - \Lambda}\hat{\eta}_0 < 1; \tag{3.14}$$

then there are positive constants K_2 and K_3 that are independent of α such that for any $0 < \eta_0 \leq \hat{\eta}_0$ we have

$$\|W(t)\|_{H^1_{\mathrm{ul}}} \leq K_2 \mathrm{e}^{-\Lambda t}\|W(0)\|_{H^1_{\mathrm{ul}}}, \qquad |q(t)| \leq K_3\|W(0)\|_{H^1_{\mathrm{ul}}}$$

for all $0 \leq t < T_{\max}(\eta_0)$ and any solution that satisfies (3.13). If $T_{\max}(\eta_0) = \infty$, then there is a $q_* \in \mathbb{R}$ with

$$|q(t) - q_*| \leq \frac{K_1K_2}{\Lambda}\mathrm{e}^{-\Lambda t}\|W(0)\|_{H^1_{\mathrm{ul}}} \tag{3.15}$$

for $t \geq 0$.

Thus, to complete the proof of Theorem 2 once the lemma has been proved, it suffices to establish a priori estimates which guarantee that $V(t)$ stays so small that $T_{\max} = \infty$ for our particular choice (3.14) of $\hat{\eta}_0$.

Proof. The variation-of-constants formula applied to (3.9) gives

$$\begin{aligned} W(t) = {} & \mathrm{e}^{\mathcal{P}^{\mathrm{s}}_\beta \mathcal{L}_\beta t} W(0) \\ & + \int_0^t \mathrm{e}^{\mathcal{P}^{\mathrm{s}}_\beta \mathcal{L}_\beta (t-s)} \mathcal{P}^{\mathrm{s}}_\beta \left[(\mathcal{R}(\xi - q(s)) - \mathcal{R}(\xi) + \mathcal{N}(V(s)))\, W(s) + \dot{q}(s) h_\xi(\xi - q(s))\rho_\beta(\xi) e_1 \right] \mathrm{d}s. \end{aligned}$$

The estimates (3.7) and (3.11) give

$$\begin{aligned} \|W(t)\|_{H^1_{\mathrm{ul}}} \leq {} & K_0 \mathrm{e}^{-\Lambda_0 t}\|W(0)\|_{H^1_{\mathrm{ul}}} \\ & + K_0 \int_0^t \mathrm{e}^{-\Lambda_0(t-s)} \left[K_1\eta_0 \|W(s)\|_{H^1_{\mathrm{ul}}} + |\dot{q}(s)|\, \|h_\xi(\xi - q(s))\rho_\beta(\xi)\|_{H^1_{\mathrm{ul}}} \right] \mathrm{d}s, \end{aligned}$$

which due to (3.12) and $\|h_\xi(\xi - q(s))\rho_\beta(\xi)\|_{H^1_{\mathrm{ul}}} \leq K_0$ implies

$$\|W(t)\|_{H^1_{\mathrm{ul}}} \leq K_0 \mathrm{e}^{-\Lambda_0 t}\|W(0)\|_{H^1_{\mathrm{ul}}} + 2K_0K_1(1+K_0)\eta_0 \int_0^t \mathrm{e}^{-\Lambda_0(t-s)}\|W(s)\|_{H^1_{\mathrm{ul}}}\, \mathrm{d}s \tag{3.16}$$

for $0 < t < T_{\max}$. Let

$$M(T) := \sup_{0\le t\le T} \mathrm{e}^{\Lambda t}\|W(t)\|_{H^1_{\mathrm{ul}}},$$

where $0 \le T \le T_{\max}$ with $T < \infty$. Equation (3.16) gives

$$\begin{aligned}\mathrm{e}^{\Lambda t}\|W(t)\|_{H^1_{\mathrm{ul}}} &\le K_0\mathrm{e}^{-(\Lambda_0-\Lambda)t}\|W(0)\|_{H^1_{\mathrm{ul}}} + 2K_0K_1(1+K_0)\eta_0\int_0^t \mathrm{e}^{-(\Lambda_0-\Lambda)(t-s)}\mathrm{e}^{\Lambda s}\|W(s)\|_{H^1_{\mathrm{ul}}}\,\mathrm{d}s\\ &\le K_0\|W(0)\|_{H^1_{\mathrm{ul}}} + 2K_0K_1(1+K_0)\eta_0 M(T)\int_0^t \mathrm{e}^{-(\Lambda_0-\Lambda)(t-s)}\,\mathrm{d}s,\end{aligned}$$

from which we conclude that

$$M(T) \le K_0\|W(0)\|_{H^1_{\mathrm{ul}}} + \frac{2K_0K_1(1+K_0)\eta_0}{\Lambda_0-\Lambda}M(T) \le K_0\|W(0)\|_{H^1_{\mathrm{ul}}} + \frac{2K_0K_1(1+K_0)\hat\eta_0}{\Lambda_0-\Lambda}M(T).$$

The choice (3.14) of $\hat\eta_0$ shows that there is a constant K_2 such that

$$\sup_{0\le t\le T} \mathrm{e}^{\Lambda t}\|W(t)\|_{H^1_{\mathrm{ul}}} \le K_2\|W(0)\|_{H^1_{\mathrm{ul}}}$$

and therefore

$$\|W(t)\|_{H^1_{\mathrm{ul}}} \le K_2\mathrm{e}^{-\Lambda t}\|W(0)\|_{H^1_{\mathrm{ul}}} \tag{3.17}$$

for $0 \le t \le T$ as desired. From (3.12) and (3.17), we conclude that

$$|\dot q(t)| \le 2K_1\mathrm{e}^{-\Lambda t}\|W(0)\|_{H^1_{\mathrm{ul}}} \tag{3.18}$$

for $0 \le t \le T$. To obtain an estimate for $q(t)$, we write

$$q(t) = q(s) + \int_s^t q'(\tau)\,\mathrm{d}\tau \tag{3.19}$$

and, setting $s = 0$ and using (3.18), we obtain

$$|q(t)| \le \int_0^t |\dot q(s)|\,\mathrm{d}s \le 2K_1K_2\|W(0)\|_{H^1_{\mathrm{ul}}}\int_0^t \mathrm{e}^{-\Lambda s}\,\mathrm{d}s \le \frac{2K_1K_2}{\Lambda}\|W(0)\|_{H^1_{\mathrm{ul}}}.$$

Setting $K_3 = 2K_1K_2/\Lambda$, we get the desired estimate

$$|q(t)| \le K_3\|W(0)\|_{H^1_{\mathrm{ul}}} \tag{3.20}$$

for $0 \le t \le T$.

Finally, if $T_{\max} = \infty$, then (3.17), (3.12), and (3.20) are valid for all times since the constants K_2 and K_3 do not depend upon T or η_0. Thus, (3.18) implies that the limit $q_* = \lim_{t\to\infty} q(t)$ exists, and (3.20) shows that $|q_*| \le K_3\|W(0)\|_{H^1_{\mathrm{ul}}}$. We can therefore take the limit $s \to \infty$ in (3.19) and get

$$q(t) = q_* + \int_\infty^t q'(\tau)\,\mathrm{d}\tau,$$

which, together with (3.18), gives the estimate (3.15). ∎

3.2. Establishing the necessary a priori estimates. To complete the proof of Theorem 2, it suffices to prove that, for sufficiently small $0 < \eta_0 \le 1$, there exists a $\delta_0 > 0$ such that

$$|q(t)| + \|V(t)\|_{H^1_{\mathrm{ul}}} \le \eta_0$$

for all $t \ge 0$ provided $\|V(0)\|_{H^1_{\mathrm{ul}}} \le \delta_0$. Throughout this section, we consider initial data $q(0) = 0$ and $V(0) \in H^1_{\mathrm{ul}}$ for which $W(0) = \rho_\beta V(0) \in H^1_{\mathrm{ul}}$.

Proposition 3.3. *There exists a constant $\varepsilon_0 > 0$ such that, if $0 < \varepsilon \le \varepsilon_0$ and*

$$\|v_1(0)\|_{H^1_{\mathrm{ul}}} \le \varepsilon^2, \qquad \|v_2(0)\|_{H^1_{\mathrm{ul}}} \le \varepsilon, \qquad \|W(0)\|_{H^1_{\mathrm{ul}}} \le \varepsilon^2,$$

then

$$|q(t)| + \|V(t)\|_{H^1_{\mathrm{ul}}} \le \left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}}$$

for $t \ge 0$ and, in particular, $T_{\max}(\eta_0) = \infty$ for $\eta_0 > 0$ sufficiently small.

Theorem 2 follows now from Proposition 3.3. Indeed, the proposition implies that (3.13) holds for all $t > 0$ so that (3.17) and (3.20) are valid for all positive times. In the remainder of this section, we prove Proposition 3.3.

Recall that v_1 satisfies (3.3), which we write as

$$\begin{aligned}\partial_t v_1 = {} & \partial_\xi^2 v_1 + c\partial_\xi v_1 - (1+c)v_1 + \tilde{\mathcal{R}}_1(\xi - q(t))\rho_\beta(\xi)^{-1} w_1 + \dot{q}(t) h_\xi(\xi - q(t)) \\ & + \frac{1}{2}[c - 3h(\xi - q(t))]v_1^2 - \frac{1}{2}v_1^3 + \gamma_1 v_2^2,\end{aligned} \tag{3.21}$$

where $\tilde{\mathcal{R}}_1(\xi) = [\frac{3}{2}(1 - h(\xi)) + c][1 + h(\xi)]$ and $w_1 = \rho_\beta v_1$.

We claim that the term $\tilde{\mathcal{R}}_1(\xi - q(t))\rho_\beta(\xi)^{-1}$ is bounded in H^1_{ul} which will allow us to control the term linear in w_1 in (3.21) using the estimate (3.17) for w_1. To show that $\tilde{\mathcal{R}}_1(\xi - q(t))\rho_\beta(\xi)^{-1}$ is bounded, we recall that $q(t)$ is bounded on $[0, T]$ on account of (3.20), $0 < \frac{3}{2}(1 - h(\xi - q(t))) \le 3 + c$, and

$$0 < [1 + h(y)]\rho_\beta(\xi)^{-1} = \left[1 + \tanh\frac{y}{2}\right] \mathrm{e}^{-\beta\xi}, \qquad \xi < -1,$$

is bounded, which, taken together, proves boundedness in H^1_{ul} as claimed. Using this result, we find that there is a positive constant K_4 such that

$$\tilde{\mathcal{G}}_1(\xi, q, W) := \tilde{\mathcal{R}}_1(\xi - q)\rho_\beta(\xi)^{-1} w_1 + \dot{q} h_x(\xi - q)$$

satisfies

$$\|\tilde{\mathcal{G}}_1(\cdot, q, W)\|_{H^1_{\mathrm{ul}}} \le K_4 \|W\|_{H^1_{\mathrm{ul}}}$$

for all $(q, V) = (q, \rho_\beta^{-1} W)$ with norm less than one, say. For any solution (q, V) satisfying (3.13) with η_0 as in Lemma 3.2, Lemma 3.2 then implies that

$$\|\tilde{\mathcal{G}}_1(\cdot, q(t), W(t))\|_{L^2_{\mathrm{ul}}} \le K_4 \|W(t)\|_{H^1_{\mathrm{ul}}} \le K_2 K_4 \mathrm{e}^{-\Lambda t} \|W(0)\|_{H^1_{\mathrm{ul}}} \le K_2 K_4 \|W(0)\|_{H^1_{\mathrm{ul}}} \tag{3.22}$$

for $t \in [0, T_{\max})$. In particular, we have

$$\sup_{0\le t<T_{\max}} \|\tilde{\mathcal{G}}_1(\cdot, q(t), W(t))\|_{L^2_{\mathrm{ul}}} \le \frac{1}{20K_0^2\sqrt{\pi}} \tag{3.23}$$

provided

$$\|W(0)\|_{H^1_{\mathrm{ul}}} \le \frac{1}{20K_0^2K_2K_4\sqrt{\pi}}. \tag{3.24}$$

Finally, the nonlinear term

$$\tilde{\mathcal{N}}_1(q, v_1) = \frac{1}{2}[c - 3h(\xi - q)]v_1^2 - \frac{1}{2}v_1^3$$

can be estimated by

$$\|\tilde{\mathcal{N}}_1(q, v_1)\|_{H^1_{\mathrm{ul}}} \le \frac{5}{2}\delta_1\|v_1\|^2_{H^1_{\mathrm{ul}}} \le \frac{1}{2K_0}\|v_1\|_{H^1_{\mathrm{ul}}} \tag{3.25}$$

for all $(q, v_1) \in \mathbb{R} \times H^1_{\mathrm{ul}}$ with $|q| \le 1$ and $\|v_1\|_{H^1_{\mathrm{ul}}} \le \frac{1}{5K_0}$.

Since the H^1_{ul}-norm is invariant under translations, we may as well consider (3.21) in the laboratory frame (x, t) in which it becomes

$$\partial_t v_1 = \mathcal{A}_1 v_1 + \tilde{\mathcal{G}}_1(x - ct, q(t), W(t)) + \tilde{\mathcal{N}}_1(q(t), v_1) + \gamma_1 v_2^2, \tag{3.26}$$

where $\mathcal{A}_1 = \partial_x^2 - (1 + c)$. The coupling term $\gamma_1 v_2^2$ makes it difficult to obtain estimates for v_1 without dealing with v_2 at the same time. Thus, our goal is to compare v_1 to the solution $\bar{v}_1$ of the equation

$$\partial_t \bar{v}_1 = \mathcal{A}_1 \bar{v}_1 + \tilde{\mathcal{G}}_1(x - ct, q(t), W(t)) + \tilde{\mathcal{N}}_1(q(t), \bar{v}_1) \tag{3.27}$$

with initial condition

$$\bar{v}_1(x, 0) = v_1(x, 0)$$

for which estimates are easier to obtain. As a first step toward estimating $\bar{v}_1$, we state the following lemma.

Lemma 3.4. *There exists a constant $K_5 > 0$ with the following property. Consider the equation*

$$\partial_t \bar{v}_1 = \mathcal{A}_1 \bar{v}_1 + G(x, t) + \tilde{\mathcal{N}}_1(q(t), \bar{v}_1), \tag{3.28}$$

where $G(x, t)$ is a given function with

$$\sup_{0\le t\le T_1} \|G(\cdot, t)\|_{L^2_{\mathrm{ul}}} < \frac{1}{20K_0^2\sqrt{\pi}}$$

for some $T_1 > 0$, and solve it with an initial condition $\bar{v}_1(0)$ with $\|\bar{v}_1(0)\|_{H^1_{\mathrm{ul}}} \le \frac{1}{20K_0^2}$; then the solution $\bar{v}_1$ of (3.28) exists for $t \in [0, T_1]$ and

$$\|\bar{v}_1(t)\|_{H^1_{\mathrm{ul}}} \le K_5(\|\bar{v}_1(0)\|_{H^1_{\mathrm{ul}}} + \sup_{0\le s\le T_1} \|G(\cdot, s)\|_{L^2_{\mathrm{ul}}})$$

for $0 \le t \le T_1$.

Proof. Since $\mathcal{A}_1$ is sectorial on H^1_{ul} (see Lemma 2.3) and the initial condition satisfies $\|\bar{v}_1(0)\|_{H^1_{\mathrm{ul}}} \le \frac{1}{20K_0^2}$, we see that there is a maximal number T_2 with $0 < T_2 \le T_1$ such that the solution to the initial-value problem (3.28) exists on $[0, T_2)$ with $\|\bar{v}_1(t)\|_{H^1_{\mathrm{ul}}} \le \frac{1}{5K_0}$ for $t \in [0, T_2)$. We claim that $T_2 = T_1$. Indeed, the variation-of-constants formula for $\bar{v}_1$ reads

$$\bar{v}_1(t) = \mathrm{e}^{\mathcal{A}_1 t}\bar{v}_1(0) + \int_0^t \mathrm{e}^{\mathcal{A}_1(t-s)}G(\cdot, s)\,\mathrm{d}s + \int_0^t \mathrm{e}^{\mathcal{A}_1(t-s)}\tilde{\mathcal{N}}_1(q(s), \bar{v}_1(s))\,\mathrm{d}s.$$

Using Lemma 2.3 and (3.25), we obtain

$$\begin{aligned}\|\bar{v}_1(t)\|_{H^1_{\mathrm{ul}}} &\le K_0\mathrm{e}^{-t}\|\bar{v}_1(0)\|_{H^1_{\mathrm{ul}}} + K_0 \sup_{0\le s\le T_1} \|G(\cdot, s)\|_{L^2_{\mathrm{ul}}} \int_0^t \mathrm{e}^{-(t-s)}(t-s)^{-\frac{1}{2}}\,\mathrm{d}s + \frac{1}{2}\sup_{0\le s\le t}\|\bar{v}_1(s)\|_{H^1_{\mathrm{ul}}} \\ &\le K_0\mathrm{e}^{-t}\|\bar{v}_1(0)\|_{H^1_{\mathrm{ul}}} + K_0\sqrt{\pi}\sup_{0\le s\le T_1}\|G(\cdot, s)\|_{L^2_{\mathrm{ul}}} + \frac{1}{2}\sup_{0\le s\le t}\|\bar{v}_1(s)\|_{H^1_{\mathrm{ul}}}\end{aligned}$$

for $0 \le t \le T_2$. Using the assumptions on $\bar{v}_1(0)$ and G, we find that $\|\bar{v}_1(T_2)\|_{H^1_{\mathrm{ul}}} \le \frac{1}{5K_0}$ from which we conclude that $T_2 = T_1$ as claimed. The above inequality then gives

$$\sup_{0\le t\le T_1}\|\bar{v}_1(t)\|_{H^1_{\mathrm{ul}}} \le 2K_0\sqrt{\pi}(\|\bar{v}_1(0)\|_{H^1_{\mathrm{ul}}} + \sup_{0\le t\le T_1}\|G(\cdot, t)\|_{L^2_{\mathrm{ul}}}),$$

which completes the proof of the lemma. ∎

To apply the preceding lemma to (3.27) on the time interval $[0, T_{\max})$, we need to prove that

$$\|\tilde{\mathcal{G}}_1(\cdot, q(t), W(t))\|_{L^2_{\mathrm{ul}}} < \frac{1}{20K_0^2\sqrt{\pi}}$$

on $[0, T_{\max})$. Equation (3.23) shows that this estimate holds for any solution (q, V) that satisfies (3.13) with η_0 as in Lemma 3.2 provided $W(0)$ satisfies (3.24). In this case, we therefore have

$$\|\bar{v}_1(t)\|_{H^1_{\mathrm{ul}}} \le K_5(\|\bar{v}_1(0)\|_{H^1_{\mathrm{ul}}} + K_2K_4\|W(0)\|_{H^1_{\mathrm{ul}}}) \tag{3.29}$$

for $t \in [0, T_{\max})$.

We shall now use the preceding estimate for $\bar{v}_1$ to obtain estimates for v_1 on the interval $[0, T_{\max})$, where $T_{\max}$ is the maximal time for which the inequality (3.13) holds for some η_0 satisfying (3.14) and for all initial conditions for which $\|V_0\|_{H^1_{\mathrm{ul}}}$ and $\|W_0\|_{H^1_{\mathrm{ul}}}$ are small enough.

Lemma 3.5. *Assume that $\gamma_1 \geq 0$. There are positive numbers K_7, α_0, and ε_0 such that the following is true for all (α, ε) with $|\alpha| < \alpha_0$ and $0 < \varepsilon < \varepsilon_0$: If $(V, W, q) = (v_1, v_2, w_1, w_2, q)$ satisfies (3.8)–(3.10) with initial data for which*

$$\|v_1(0)\|_{H^1_{\mathrm{ul}}} \leq \varepsilon^2, \qquad \|v_2(0)\|_{H^1_{\mathrm{ul}}} \leq \varepsilon, \qquad \|W(0)\|_{H^1_{\mathrm{ul}}} \leq \varepsilon^2, \tag{3.30}$$

then the v_2-component of the solution satisfies

$$\|v_2(t)\|_{L^2_{\mathrm{ul}}} \leq K_7 \left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}}$$

for all t with $0 < t < T_{\max}$.

Proof. Using (3.30), we infer from (3.29) that $\|\bar{v}_1(t)\|_{C^0} \leq K_0 \|\bar{v}_1(t)\|_{H^1_{\mathrm{ul}}} \leq K_6 \varepsilon^2$, where $K_6 = K_0 K_2 K_4 K_5$ does not depend on ε, and therefore

$$\bar{v}_1(x,t) \geq -K_6 \varepsilon^2$$

for all $x \in \mathbb{R}$ and $0 < t < T_{\max}$. Next, (3.26) and (3.27) together with the assumption $\gamma_1 \geq 0$ show that

$$\begin{aligned} &\partial_t \bar{v}_1 - \mathcal{A}_1 \bar{v}_1 - \tilde{\mathcal{G}}_1(x - ct, q(t), W(t)) - \tilde{\mathcal{N}}_1(q(t), \bar{v}_1) \overset{(3.27)}{=} 0 \\ &\leq \gamma_1 v_2^2 \overset{(3.26)}{=} \partial_t v_1 - \mathcal{A}_1 v_1 - \tilde{\mathcal{G}}_1(x - ct, q(t), W(t)) - \tilde{\mathcal{N}}_1(q(t), v_1). \end{aligned}$$

The comparison principle [4, Theorem 25.1 in section VII] gives $\bar{v}_1(x,t) \leq v_1(x,t)$ for $0 \leq t < T_{\max}$ and $x \in \mathbb{R}$, and therefore

$$v_1(x,t) \geq \bar{v}_1(x,t) \geq -K_6 \varepsilon^2 \tag{3.31}$$

for $0 \leq t < T_{\max}$ and $x \in \mathbb{R}$. Having established the lower pointwise bound (3.31) for v_1, we return to the equation

$$\partial_t v_2 = -(1 + \partial_x^2)^2 v_2 + \alpha v_2 - v_2^3 - \gamma_2 [1 + h(x - ct - q(t))] v_2 - \gamma_2 v_1 v_2$$

for v_2, written in the laboratory frame. Using Lemma 2.1(iii) and the bounds (3.31), $\gamma_2 \geq 0$, and $[1 + h(y)] \geq 0$ for all $y \in \mathbb{R}$, we obtain

$$\begin{aligned} \frac{1}{2} \partial_t \|v_2\|^2_{L^2_{\mathrm{ul}}(\sigma_b)} &\leq \left[\frac{7}{2} b^2 + |\alpha|\right] \int_{\mathbb{R}} \sigma_b v_2^2 \, \mathrm{d}x - \int_{\mathbb{R}} \sigma_b v_2^4 \, \mathrm{d}x - \int_{\mathbb{R}} \sigma_b \gamma_2 [1 + h(x - ct - q(t))] v_2^2 \, \mathrm{d}x \\ &\quad + \gamma_2 K_6 \varepsilon^2 \int_{\mathbb{R}} \sigma_b v_2^2 \, \mathrm{d}x \\ &\leq \left[\frac{7}{2} b^2 + K_6 \gamma_2 \varepsilon^2 + |\alpha|\right] \int_{\mathbb{R}} \sigma_b v_2^2 \, \mathrm{d}x - \int_{\mathbb{R}} \sigma_b v_2^4 \, \mathrm{d}x. \end{aligned}$$

Next, we record that

$$\int_{\mathbb{R}} a \sigma_b v_2^2 \, \mathrm{d}x - \int_{\mathbb{R}} \sigma_b v_2^4 \, \mathrm{d}x \leq \frac{a^2}{b} - \int_{\mathbb{R}} a \sigma_b v_2^2 \, \mathrm{d}x$$

for any constant $a > 0$ since

$$\int 2a\sigma_b v_2^2 - \int \sigma_b v_2^4 \le \int \sqrt{2}a\sigma_b^{1/2}\sqrt{2}\sigma_b^{1/2}v_2^2 - \int \sigma_b v_2^4 \le \int a^2\sigma_b + \int \sigma_b v_2^4 - \int \sigma_b v_2^4 \le \frac{a^2}{b}.$$

Therefore,

$$\frac{1}{2}\partial_t \|v_2\|^2_{L^2_{\mathrm{ul}}(\sigma_b)} \le \frac{[7b^2 + 2(K_6\gamma_2\varepsilon^2 + |\alpha|)]^2}{4b} - \frac{1}{2}[7b^2 + 2(K_6\gamma_2\varepsilon^2 + |\alpha|)]\|v_2\|^2_{L^2_{\mathrm{ul}}(\sigma_b)}.$$

This is a differential inequality of the form $\frac{1}{2}f'(t) \le d_1 - d_2 f(t)$ for which Gronwall's estimate [13, Theorem 1.5.7] gives

$$f(t) \le \mathrm{e}^{-2d_2 t}f(0) + \frac{d_1}{d_2}(1 - \mathrm{e}^{-2d_2 t}) \le f(0) + \frac{d_1}{d_2}$$

for $d_2 > 0$. In our case, this estimate becomes

$$\|v_2(t)\|^2_{L^2_{\mathrm{ul}}(\sigma_b)} \le \|v_2(0)\|^2_{L^2_{\mathrm{ul}}(\sigma_b)} + \frac{7b^2 + 2(K_6\gamma_2\varepsilon^2 + |\alpha|)}{2b} \le \frac{2K_0^2\varepsilon^2 + [7b^2 + 2(K_6\gamma_2\varepsilon^2 + |\alpha|)]}{2b},$$

where we used that

$$\|v_2(0)\|^2_{L^2_{\mathrm{ul}}(\sigma_b)} \le \frac{1}{b}\|v_2(0)\|^2_{L^2_{\mathrm{ul}}(\sigma)} \le \frac{K_0^2}{b}\|v_2(0)\|^2_{H^1_{\mathrm{ul}}} \le \frac{K_0^2\varepsilon^2}{b}.$$

Setting $b = \sqrt{\varepsilon^2 + |\alpha|}$ and using Lemma 2.1(ii), we finally get

$$\|v_2(t)\|^2_{L^2_{\mathrm{ul}}} \le K_7^2\sqrt{\varepsilon^2 + |\alpha|}, \qquad 0 \le t < T_{\max},$$

for an appropriate constant K_7 that depends only on K_0, K_6, and γ_2 but not on α, ε, or t. ∎

Next, we estimate v_1 and v_2 in the H^1_{ul}-norm.

Lemma 3.6. *Assume that $\gamma_1 \ge 0$. There are positive numbers K_8, α_0, and ε_0 such that the following is true for all (α, ε) with $|\alpha| < \alpha_0$ and $0 < \varepsilon < \varepsilon_0$: If $(V, W, q) = (v_1, v_2, w_1, w_2, q)$ satisfies (3.8)–(3.10) with initial data for which (3.30) holds, then*

$$\|V(t)\|_{H^1_{\mathrm{ul}}} \le K_8\left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}} \tag{3.32}$$

uniformly in $t \in [0, T_{\max})$.

Proof. First, we shall estimate $\|v_1\|_{H^1_{\mathrm{ul}}}$. Assumption (3.30) allows us to apply Lemma 3.4, and we conclude

$$\|v_2(t)\|_{L^2_{\mathrm{ul}}} \le K_7\left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}}$$

on $[0, T_{\max})$. Furthermore, from the definition of $T_{\max}$, we know that $\|V(t)\|_{H^1_{\mathrm{ul}}} \le \eta_0$ on $[0, T_{\max})$. Taken together, these estimates show that

$$\|v_2^2(t)\|_{L^2_{\mathrm{ul}}} \le \eta_0 K_0 K_7\left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}} \tag{3.33}$$

on $[0, T_{\max})$. To find an estimate for $\|v_1(t)\|_{H^1_{\mathrm{ul}}}$, we wish to apply Lemma 3.4 to (3.26) which means that we have to set

$$G(x,t) := \tilde{\mathcal{G}}_1(x - ct, q(t), W(t)) + \gamma_1 v_2(x,t)^2$$

in Lemma 3.4. The estimates (3.22) and (3.33) give

$$\|G(\cdot,t)\|_{L^2_{\mathrm{ul}}} \le K_9 \left(\|W(0)\|_{H^1_{\mathrm{ul}}} + \left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}} \right)$$

for some constant K_9 that does not depend on α, ε, or t. Lemma 3.4 and (3.30) now show that

$$\|v_1(t)\|_{H^1_{\mathrm{ul}}} \le K_5 K_9 \left(\|v_1(0)\|_{H^1_{\mathrm{ul}}} + \|W(0)\|_{H^1_{\mathrm{ul}}} + \left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}} \right) \le K_5 K_9 \left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}}$$

on $[0, T_{\max})$.

Next, we employ energy methods to estimate $\|v_2(t)\|_{H^1_{\mathrm{ul}}}$, and we begin by collecting the bounds

$$\|V(t)\|_{H^1_{\mathrm{ul}}} \le \eta_0, \qquad \|v_2(t)\|_{L^2_{\mathrm{ul}}} \le K_7 \left[\varepsilon + \sqrt{|\alpha|}\right]^{\frac{1}{2}}, \qquad \|W(t)\|_{H^1_{\mathrm{ul}}} \le K_2 \|W(0)\|_{H^1_{\mathrm{ul}}} \le K_2 \varepsilon^2,$$

which we established so far on the time interval $[0, T_{\max})$. We write the equation for the v_2-component in the form

$$\partial_t v_2 = -(1 + \partial_x^2)^2 v_2 + \alpha v_2 - v_2^3 + \mathcal{R}_2(x - ct - q(t)) \rho_\beta(x - ct)^{-1} w_2 - \gamma_2 v_1 v_2,$$

where we recall that $\mathcal{R}_2(y - q(t))\rho_\beta(y)^{-1}$ is bounded in H^1_{ul}, and consider also the equivalent integral equation

$$\begin{aligned} v_2(t) = \mathrm{e}^{\mathcal{A}_2 t} v_2(0) + \int_0^t \mathrm{e}^{\mathcal{A}_2 (t-s)} \big[& \alpha v_2(s) - v_2^3(s) - \gamma_2 v_1(s) v_2(s) \\ & + \mathcal{R}_2(x - cs - q(s)) \rho_\beta(x - cs)^{-1} w_2(s) \big] \, \mathrm{d}s \end{aligned} \tag{3.34}$$

with $\mathcal{A}_2 = -(1 + \partial_x^2)^2$. Lemma 2.3 shows that

$$\|\mathrm{e}^{\mathcal{A}_2 t}\|_{H^1_{\mathrm{ul}}} \le K_0, \qquad \|\mathrm{e}^{\mathcal{A}_2 t}\|_{L^2_{\mathrm{ul}} \to H^1_{\mathrm{ul}}} \le K_0 t^{-\frac{1}{4}}, \qquad t > 0,$$

and applying these estimates together with (3.17) to (3.34) gives

$$\begin{aligned} \|v_2(t)\|_{H^1_{\mathrm{ul}}} \le K_0 \|v_2(0)\|_{H^1_{\mathrm{ul}}} + K_0 \int_0^t (t-s)^{-\frac{1}{4}} \Big[& |\alpha| \|v_2(s)\|_{L^2_{\mathrm{ul}}} + \|v_2(s)\|_{L^2_{\mathrm{ul}}}^2 \|v_2(s)\|_{H^1_{\mathrm{ul}}} \\ & + \|v_1(s)\|_{H^1_{\mathrm{ul}}} \|v_2(s)\|_{L^2_{\mathrm{ul}}} + \|W(0)\|_{H^1_{\mathrm{ul}}} \Big] \, \mathrm{d}s, \end{aligned}$$

where we used that

$$\|v_2^3\|_{L^2_{\mathrm{ul}}} \le K_0 \|v_2\|_{L^2_{\mathrm{ul}}} \|v_2\|_{C^0}^2 \le K_0^2 \|v_2\|_{L^2_{\mathrm{ul}}}^2 \|v_2\|_{H^1_{\mathrm{ul}}}$$

on account of Lemma 2.1(i).

We set $T_1 := \min\{T_{\max}, 1\}$. For $0 < t < T_1$, we then have

$$\|v_2(t)\|_{H^1_{\mathrm{ul}}} \le K_0 K_7 \left[\|v_2(0)\|_{H^1_{\mathrm{ul}}} + \int_0^t (t-s)^{-\frac{1}{4}} \left(|\alpha| \left[\varepsilon + \sqrt{|\alpha|} \right]^{\frac{1}{2}} + \left[\varepsilon + \sqrt{|\alpha|} \right] \left[1 + \sup_{0 \le s \le t} \|v_2(s)\|_{H^1_{\mathrm{ul}}} \right] \right) \mathrm{d}s \right],$$

and therefore

$$\sup_{0 \le t \le T_1} \|v_2(t)\|_{H^1_{\mathrm{ul}}} \le K_{10} \left(\|v_2(0)\|_{H^1_{\mathrm{ul}}} + \left[\varepsilon + \sqrt{|\alpha|} \right]^{\frac{1}{2}} + \left[\varepsilon + \sqrt{|\alpha|} \right] \sup_{0 \le t \le T_1} \|v_2(t)\|_{H^1_{\mathrm{ul}}} \right)$$

for an appropriate constant K_{10} that does not depend on α, ε, or T_1 (as long as $T_1 \le 1$). We now choose positive bounds α_0 and ε_0 for α and ε, respectively, that are so small that $K_{10}[\varepsilon_0 + \sqrt{\alpha_0}] \le \frac{1}{2}$. For all (α, ε) with $|\alpha| < \alpha_0$ and $|\varepsilon| < \varepsilon_0$ we then have

$$\sup_{0 \le t \le T_1} \|v_2(t)\|_{H^1_{\mathrm{ul}}} \le 2K_{10} \left[\|v_2(0)\|_{H^1_{\mathrm{ul}}} + \left[\varepsilon + \sqrt{|\alpha|} \right]^{\frac{1}{2}} \right].$$

To obtain estimates for $\|v_2(t)\|_{H^1_{\mathrm{ul}}}$ for $t > 1$, we use the variation-of-constants formula

$$v_2(t) = \mathrm{e}^{\mathcal{A}_2(t-\tau)} v_2(\tau) + \int_\tau^t \mathrm{e}^{\mathcal{A}_2(t-s)} \left[\alpha v_2(s) - v_2^3(s) - \gamma_2 v_1(s) v_2(s) + \mathcal{R}_2(x - cs - q(s)) \rho_\beta(x - cs)^{-1} w_2(s) \right] \mathrm{d}s$$

on $[\tau, t]$ for each τ with $t - 1 \le \tau \le t$. As before, we obtain

$$\|v_2(t)\|_{H^1_{\mathrm{ul}}} \le K_0 (t-\tau)^{-\frac{1}{4}} \|v_2(\tau)\|_{L^2_{\mathrm{ul}}} + K_0 \int_\tau^t (t-s)^{-\frac{1}{4}} \left[|\alpha| \|v_2(s)\|_{L^2_{\mathrm{ul}}} + \|w_2(s)\|_{H^1_{\mathrm{ul}}} + \|v_2(s)\|^2_{L^2_{\mathrm{ul}}} \|v_2(s)\|_{H^1_{\mathrm{ul}}} + \|v_1(s)\|_{H^1_{\mathrm{ul}}} \|v_2(s)\|_{L^2_{\mathrm{ul}}} \right] \mathrm{d}s$$

and consequently

$$\|v_2(t)\|_{H^1_{\mathrm{ul}}} \le K_0 K_7 (t-\tau)^{-\frac{1}{4}} \left[\varepsilon + \sqrt{|\alpha|} \right]^{\frac{1}{2}} + K_0 \left[\varepsilon + \sqrt{|\alpha|} \right] \int_\tau^t (t-s)^{-\frac{1}{4}} \left[1 + \|v_2(s)\|_{H^1_{\mathrm{ul}}} \right] \mathrm{d}s. \tag{3.35}$$

Upon setting

$$\mathcal{J}_1(\tau) := \left(\int_\tau^{\tau+1} \|v_2(s)\|^3_{H^1_{\mathrm{ul}}} \, \mathrm{d}s \right)^{1/3},$$

Hölder's inequality gives

$$\int_\tau^t (t-s)^{-\frac{1}{4}} \|v_2(s)\|_{H^1_{\mathrm{ul}}} \, \mathrm{d}s \le \left(\int_\tau^t (t-s)^{-\frac{3}{8}} \, \mathrm{d}s \right)^{2/3} \left(\int_\tau^t \|v_2(s)\|^3_{H^1_{\mathrm{ul}}} \, \mathrm{d}s \right)^{1/3} \le K_0 \mathcal{J}_1(\tau),$$

and therefore

$$\left(\int_\tau^{\tau+1}\left[\int_\tau^t (t-s)^{-\frac{1}{4}}\|v_2(s)\|_{H^1_{\mathrm{ul}}}\,\mathrm{d}s\right]^3\mathrm{d}t\right)^{\frac{1}{3}} \le K_0\mathcal{J}_1(\tau). \tag{3.36}$$

Upon raising (3.35) to the power three, integrating both sides over $t \in [\tau, \tau+1]$, taking the third root, and using (3.36), we see that

$$\mathcal{J}_1(\tau) \le K_{11}\left(\left[\varepsilon+\sqrt{|\alpha|}\right]^{\frac{1}{2}} + \left[\varepsilon+\sqrt{|\alpha|}\right]\mathcal{J}_1(\tau)\right)$$

for an appropriate constant K_{11} that does not depend on t or τ. Making α_0 and ε_0 smaller if necessary, we conclude that

$$\mathcal{J}_1(\tau) \le 2K_{11}\left[\varepsilon+\sqrt{|\alpha|}\right]^{\frac{1}{2}},$$

and using this estimate in (3.35), we finally obtain the pointwise estimate

$$\|v_2(\tau+1)\|_{H^1_{\mathrm{ul}}} \le K_0K_{11}\left[\varepsilon+\sqrt{|\alpha|}\right]^{\frac{1}{2}},$$

which is valid for any $\tau > 0$. This completes the proof of the lemma. ■

We are now ready to complete the proof of Proposition 3.3.

Proof of Proposition 3.3. For sufficiently small $\varepsilon > 0$, Lemma 3.6 shows that $|q(t)| + \|V(t)\|_{H^1_{\mathrm{ul}}}\| \le \frac{1}{2}\eta_0$ for $0 \le t < T_{\max}$, which contradicts the maximality of $T_{\max}$ (see (3.13)) if $T_{\max}$ is finite. Thus, (3.13) holds for any t, which in turn implies that (3.17) and (3.20) are valid for all times. Therefore, (3.32) holds with $T_{\max} = \infty$, which completes the proof of Proposition 3.3. ■

4. Implications of nonlinear stability and comparison with simulations. Throughout this section, we consider the system

$$\begin{aligned}\partial_t u_1 &= \partial_\xi^2 u_1 + c\partial_\xi u_1 + \frac{1}{2}(u_1-c)(1-u_1^2) + \gamma_1 u_2^2,\\ \partial_t u_2 &= -(1+\partial_\xi^2)^2 u_2 + c\partial_\xi u_2 + \alpha u_2 - u_2^3 - \gamma_2 u_2(1+u_1)\end{aligned} \tag{4.1}$$

exclusively in the frame $\xi = x - ct$ that moves with the front. In this frame, the front solution is stationary and is given by

$$U_h(\xi) = \begin{pmatrix}\tanh(\xi/2)\\ 0\end{pmatrix}.$$

We shall also assume that the coefficients appearing in (4.1) satisfy the assumptions required in Theorem 2. This theorem then asserts that, for any function V_0 for which $\|V_0\|_{H^1_{\mathrm{ul}}}$ is sufficiently small, the solution $U(\xi,t)$ with initial data $U(\cdot,0) = U_h + V_0$ can be written as $U(\xi,t) = U_h(\xi - q(t)) + V(\xi,t)$ and that there is a number q_* so that

$$\|\rho_\beta(\cdot)V(\cdot,t)\|_{H^1_{\mathrm{ul}}} + |q(t)-q_*| \le K\mathrm{e}^{-\Lambda_* t}, \qquad \|V(\cdot,t)\|_{H^1_{\mathrm{ul}}} \le K\left[\|V_0\|_{H^1_{\mathrm{ul}}} + \sqrt{|\alpha|}\right]^{\frac{1}{2}} \tag{4.2}$$

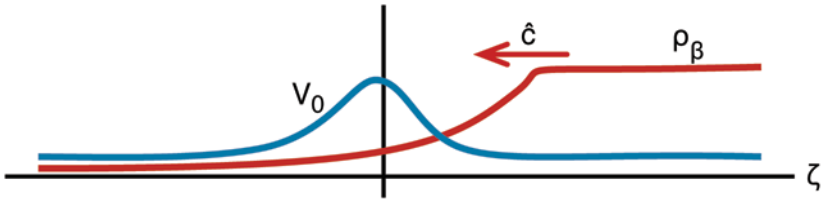


Figure 3. *A schematic illustration of the interplay between the perturbation $V_0(\zeta)$ and the weight $\rho_\beta(\zeta+\hat{c}t)$ in the frame $\zeta = \xi + \hat{c}t$ that moves with the perturbation. The speed $\hat{c}$ needs to be negative so that the weight travels to the right, since (4.3) implies that the product $\rho_\beta(\zeta+\hat{c}t)V_0(\zeta)$ tends to zero as $t \to \infty$.*

for $t \geq 0$.

To see what the implications of the nonlinear stability estimates (4.2) are, let us suppose that $V(\xi,t)$ assumes the form of a traveling wave of speed $\hat{c}$ so that $V(\xi,t) = V_0(\xi - \hat{c}t)$. Using the coordinate $\zeta = \xi - \hat{c}t$ that moves with the perturbation, we obtain from (4.2) that

$$K\mathrm{e}^{-\Lambda_* t} \geq \|\rho_{\beta_*}(\cdot)V(\cdot,t)\|_{H^1_{\mathrm{ul}}} = \|\rho_{\beta_*}(\xi)V_0(\xi-\hat{c}t)\|_{H^1_{\mathrm{ul}}} = \|\rho_{\beta_*}(\zeta+\hat{c}t)V_0(\zeta)\|_{H^1_{\mathrm{ul}}}, \tag{4.3}$$

which implies that the speed $\hat{c}$ needs to satisfy

$$\hat{c} \leq -\frac{\Lambda_*}{\beta_*}. \tag{4.4}$$

In particular, $\hat{c}$ is negative, meaning that any traveling wave that exists in the wake of the stationary front U_h moves toward $\xi = -\infty$, that is, away from the front U_h; see Figure 3 for an illustration.

One possible candidate for perturbations of traveling wave type in the wake of the front U_h are *Turing fronts* which, by definition, connect the spatially periodic Turing patterns U_{per} discussed in (1.4) at $\xi = -\infty$ to the unstable homogeneous rest state U_- in the wake of the front. We shall first derive an explicit estimate c^* for the maximal speed with which they can move.

The Turing patterns (1.4) have amplitude of order $\sqrt{\alpha}$, and we therefore set $\varepsilon = K\sqrt{\alpha}$ in Theorem 2 for a sufficiently large constant K. The proof of Proposition 2.2, and in particular (2.4), shows that the spectrum of the linearization $\mathcal{L}_\beta$ of (4.1) about U_h in the weighted space lies to the left of the line $\mathrm{Re}\,\lambda = \alpha - c\beta + 4\beta^2 + 8\beta^4$ with the exception of the translation eigenvalue at the origin. The decay rate Λ_* in Theorem 2 is chosen in Lemma 3.2: Choosing $\hat{\eta}_0 = K\sqrt{\alpha}$ allows us to set $\Lambda_* = \Lambda = K\sqrt{\alpha} - [\alpha - c\beta + 4\beta^2 + 8\beta^4]$ (we remark here that the constants K_2 and K_3 in Lemma 3.2 depend only on the value of the left-hand side of (3.14) but not on the values of Λ and $\hat{\eta}_0$). Substituting this expression for Λ_* into (4.4), we obtain

$$\hat{c} \leq \frac{K\sqrt{\alpha} + \alpha - c\beta + 4\beta^2 + 8\beta^4}{\beta} \leq -c + \frac{K\sqrt{\alpha} + \alpha + 4\beta^2 + 8\beta^4}{\beta}.$$

The minimum $4\sqrt{K}\alpha^{1/4} + \mathrm{O}(\alpha^{3/4})$ of the right-hand side over $\beta > 0$ is achieved at $\beta = \sqrt{K}\alpha^{1/4} + \mathrm{O}(\alpha^{3/4})$, which gives the upper bound

$$\hat{c} \leq c^* := -c + 4\sqrt{K}\alpha^{1/4} + \mathrm{O}(\alpha^{3/4}) \tag{4.5}$$

for the speed of traveling fronts with amplitude bounded by $K\sqrt{\alpha}$ to the left of the stationary front U_h. We remark that similar upper bounds for more general solutions of the Swift–Hohenberg equation, but without the presence of a front to the right, were obtained in [5].

Next, we complement the upper bound (4.5) for $\hat{c}$ by formal lower bounds using the results in [6, 20] by van Saarloos and his collaborators who derived lower bounds c_* for the propagation of Turing patterns into the unstable homogeneous rest state U_-. Applying the formulas in [20, section 2.11] to the u_2-component

$$\partial_t v_2 = -(1+\partial_\xi^2)^2 v_2 + c\partial_\xi v_2 + \alpha v_2$$

of the linearization of (4.1) about $U_- = (1,0)$, we obtain the lower bound

$$c_* = -c + 4\sqrt{\alpha} + \mathrm{O}(\alpha^{3/2}) \tag{4.6}$$

for Turing fronts. Thus, combining (4.5) and (4.6), we expect that Turing fronts in the wake of the stationary front U_h travel at a speed $\hat{c}$ with $c_* \le \hat{c} \le c^*$ to the left.

In summary, Theorem 2 shows that small perturbations to the front should move away from the front at a speed $\hat{c}$ that satisfies (4.4). If the perturbations are of order $\sqrt{\alpha}$, then the speed at which they have to move to the left satisfies the more explicit estimate (4.5). Furthermore, if we construct an initial condition that consists of the small-amplitude Turing patterns at $\xi = -\infty$ with the front U_h to their right, then we expect that the solution to (4.1) is the superposition of the stationary front U_h with a Turing front in its wake whose speed $\hat{c}$ lies between the lower and upper bounds provided by (4.6) and (4.5), respectively. This last claim is based only on formal arguments, though.

We now compare these predictions with numerical simulations of (4.1). Throughout, we set

$$\gamma_1 = 0.5, \quad \gamma_2 = 0.6, \quad c = 0.5 \tag{4.7}$$

and solve (4.1) on the interval $(0,\ell)$ for $\ell = 1000$ with the boundary conditions

$$u_1(0,t) = -1, \quad u_1(\ell,t) = 1, \quad u_2(0,t) = \partial_\xi u_2(0,t) = u_2(\ell,t) = \partial_\xi u_2(\ell,t) = 0.$$

We discretized (4.1) using centered finite differences with step size 0.05 and integrated the resulting ODE using the explicit Runge–Kutta–Chebyshev scheme developed in [27]. Throughout, we pick the initial condition

$$U_0(\xi) = \begin{pmatrix} \tanh[0.5(\xi - 950)] + 0.045\cos\xi \\ 0.045\cos\xi \end{pmatrix}, \tag{4.8}$$

which excites the most unstable linear mode over the entire domain, and comment in section 5 on other initial data.

First, we choose $\alpha = 0.001$ to be very small in order to test the speed predictions from (4.5) and (4.6). In Figure 4, we plot the difference $V(\xi,t) = (v_1,v_2)(\xi,t)$ between the solution $U(\xi,t)$ and the front $U_h(\xi - 950.024)$ at $t = 1000$ for $\alpha = 0.001$. The relative offset 0.024 to the front interface at $\xi = 950$ for $t = 0$ minimizes the difference between U and U_h near

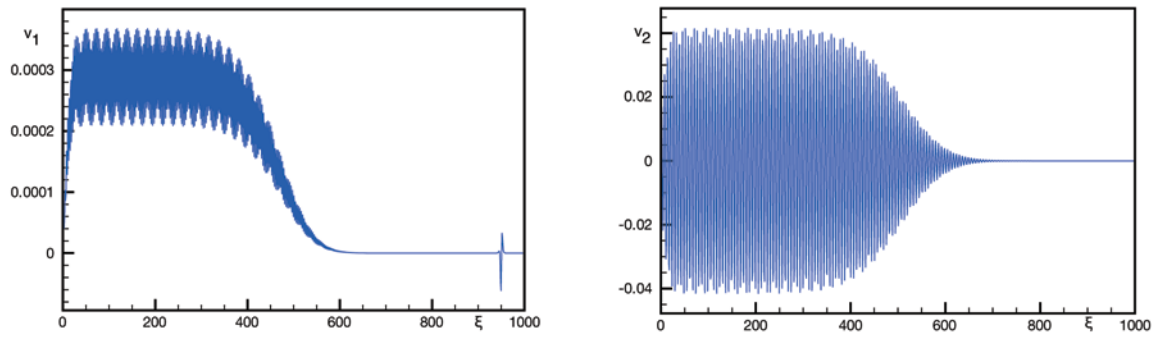


Figure 4. *Plotted are the values of* $v_1 = u_1 - h(\cdot - 950.024)$ *(left) and of* $v_2 = u_2$ *(right) for* $\alpha = 0.001$ *in the comoving frame as functions of* ξ *for* $t = 1000$*: The Turing patterns behind the front have been pushed to the left of the front interface which is located at* $\xi = 950.024$ *for* $t = 1000$*.*

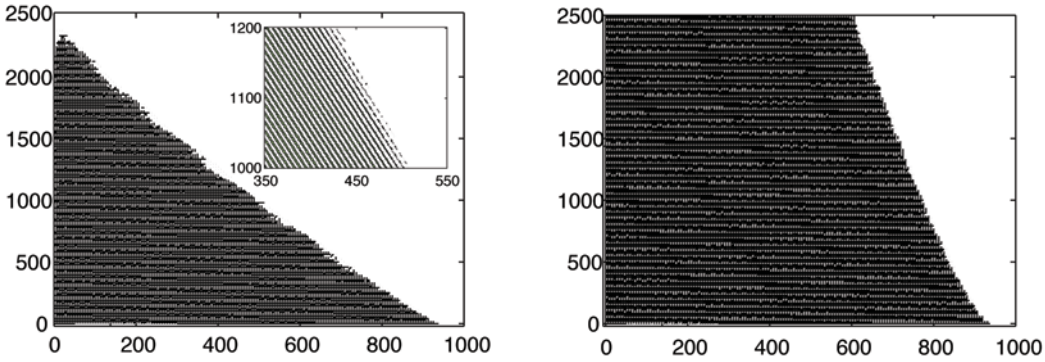


Figure 5. *Space-time contour plots of* $v_2 = u_2$ *are shown for* $\alpha = 0.001$ *(left) and* $\alpha = 0.01$ *(right) in the comoving frame (time* t *upward and space* ξ *horizontal). The inset illustrates that small individual Turing patterns behind the front travel to the left as expected. For* $\alpha = 0.001$*, the overall perturbation also travels to the left at an approximately constant speed* -0.384*. For* $\alpha = 0.01$*, the perturbation still travels to the left, but at a much smaller speed.*

$\xi = 950$ and accounts therefore for the shift $q(t)$ from Theorem 2. Figure 4 indicates that V becomes very small ahead of and near the front as expected, while it develops Turing patterns to the left of the front, that is, in the spatial regime where the background state is unstable. Upon measuring the slope of the Turing front interface in the contour plot shown in Figure 5 (left plot), we find that the Turing front travels at speed -0.384 to the left. This is in agreement with the formal lower bound (4.6) which gives the minimal speed $c_* = -0.373$ upon substituting $\alpha = 0.001$ into (4.6).

For larger values α, the perturbation will still evolve into Turing fronts which are pushed to the left of the front, but the relative speed between the Turing front and the large front h will decrease as α increases. This is illustrated in the simulation for $\alpha = 0.01$ shown in the right plot of Figure 5.

Eventually, for sufficiently large parameters α, the instability of the background state $u = 0$, which is initially convective in nature, will become an absolute instability: Perturbations of the background state will then no longer be convected away but will grow in amplitude

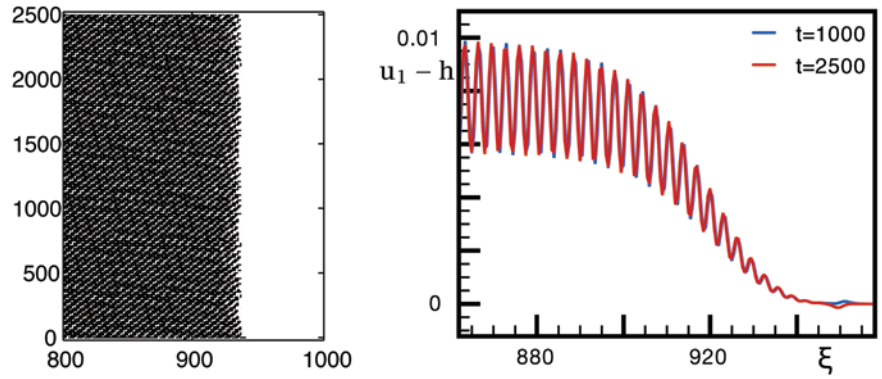


Figure 6. *The simulations in this figure are for $\alpha = 0.03$. On the left, a space-time contour plot of $v_2 = u_2$ is shown in the comoving frame (time t upward and space ξ horizontal). On the right, we plotted $v_1 = u_1 - h(\cdot - 950.02)$ at $t = 1000$ in blue and $v_1 = u_1 - h(\cdot - 950.0135)$ at $t = 2500$ in red: Note that the two graphs lie on top of each other which indicates that the Turing patterns have locked to the front h.*

as time increases at each fixed point in space. In this situation, we can no longer expect that perturbations will be pushed away by the front. Instead, the Turing patterns behind the front may lock to the front, yielding a time-periodic wave in an appropriate comoving frame. The convective instability changes to an absolute instability when the linear dispersion relation $\lambda = \lambda_-(k)$ from (2.3) has a double root $k_{\mathrm{bp}} \in \mathbb{C}$ with $\lambda \in \mathrm{i}\mathbb{R}$; see, for instance, [3] or [21] and the references therein. For our system with parameter values as in (4.7), the transition from convective to absolute instability occurs at $\alpha = 0.015$. Figure 6 shows simulations for $\alpha = 0.03$ which illustrate that the expected locking behavior between the front and the Turing pattern in its wake indeed occurs in our system.

In summary, our numerical simulations with initial data (4.8) show that the corresponding solution is indeed composed of two fronts—a Turing front which connects the Turing pattern at $-\infty$ to the unstable homogeneous rest state $U_- = (-1, 0)$, and the primary front h which connects $(-1, 0)$ to the stable homogeneous rest state $(1, 0)$ at $+\infty$. In the parameter regime where the equilibrium U_- is only convectively unstable, the relative speed between the two fronts is positive, and the front h therefore is asymptotically stable in the weighted norm as predicted by Theorem 2. As α increases, the relative speed between the two interfaces decreases. For sufficiently large α, the equilibrium U_- is absolutely unstable, and we then observe locking of the front h and the Turing pattern in its wake.

5. Discussion. In this paper, we discussed the nonlinear stability of convectively unstable fronts near supercritical Turing instabilities for the specific system (1.2). To prove nonlinear stability, we established a priori H^1_{ul}-estimates for solutions with initial data close to the front and used these estimates to show exponential temporal decay of solutions when measured in exponentially weighted H^1_{ul}-norms. While the second part of the proof generalizes easily to general partial differential equations, our proof of a priori estimates relies on the comparison principle and depends therefore on the special structure of our model system.

We expect nevertheless that our nonlinear stability result remains true for general partial differential equations, and there is indeed much numerical evidence that supports this belief.

For instance, the Gray–Scott system

$$\partial_t U_1 = d_1 \partial_x^2 U_1 - U_1 U_2^2 + F(1 - U_1),$$
$$\partial_t U_2 = d_2 \partial_x^2 U_2 + U_1 U_2^2 - (F + k) U_1$$

is known to have Turing bifurcations, and direct numerical computations show that these are supercritical in certain parameter regions [15]. The direct numerical partial differential equation simulations in [23, section 8] show that the Gray–Scott system exhibits fronts in parts of this parameter regime which become convectively unstable at the supercritical Turing bifurcation. In particular, [23, Figures 14–15] indicate that the convectively unstable fronts are nonlinearly stable in the weighted norm.

Our results should also remain true if the homogeneous equilibrium U_- behind the front undergoes a supercritical Hopf bifurcation, rather than a Turing bifurcation. In both cases, the dynamics near U_- is captured by the complex Ginzburg–Landau equation (CGL)

$$A_t = (1 + \mathrm{i}a)\partial_x^2 A + \alpha A - (1 + \mathrm{i}b)|A|^2 A, \tag{5.1}$$

but the coefficients a and b vanish for Turing bifurcations, while they are generically nonzero for Hopf bifurcations. Depending on the values of the coefficients a and b, the Ginzburg–Landau equation may exhibit stable oscillatory waves or spatio-temporally complex patterns which, beyond onset, appear behind the front. Again, we expect that the front should outrun these structures in its wake, while leaving a growing spatial region behind it where the solution converges to the unstable equilibrium U_-. Sherratt [25] confirmed this picture, through a formal analysis, for fronts near supercritical Hopf bifurcations in the case when these can be described by λ-ω systems, i.e., for $a = 0$. We also refer the reader to [14, 25, 26] for numerical simulations in this setup and for applications to predator-prey systems.

The numerical simulations presented in section 4 used the initial condition (4.8) which selected a single linearly unstable spatially periodic mode. For more general initial data close to the front h, the perturbation remains small and is still pushed to the left for α near zero. The dynamics in the wake of the front may, however, be more complex and may, in particular, involve amplitude-modulated Turing patterns whose spatial periods vary over space: The evolution behind the front is, on a formal level, captured by the Ginzburg–Landau approximation (5.1). As mentioned previously, the nonlinear stability result presented here is valid independently of the particular dynamics behind the front, as long as a priori bounds for solutions are available.

As illustrated in Figure 6 in section 4, we expect that the stability properties of the front change when the equilibrium behind the front becomes absolutely unstable, since perturbations will then grow pointwise in space, rather than being convected toward $-\infty$. In our specific model problem, the Turing patterns lock to the front, and a modulated front (time-periodic in an appropriate moving coordinate frame) emerges which converges to spatially periodic patterns in its wake. In particular, the original front is no longer stable in weighted norms.

Finally, we comment on subcritical bifurcations behind the front. In this case, we cannot expect to have nonlinear stability in weighted norms since perturbations will not necessarily stay small in the wake of the front but may grow to finite amplitude. In particular, there is no

reason to believe that the solution in the wake will not strongly influence the front ahead, thus possibly precluding nonlinear stability; note though that the front may still be nonlinearly stable if conditions are right. In the system (4.1), these different behaviors can be observed: The Turing bifurcation is subcritical for parameters as in (4.7) with $\gamma_1 = -4$ or $\gamma_1 = -8$. Numerical simulations of (4.1) for $\gamma_1 = -4$ show that the solution behind the front converges to a spatially periodic pattern of finite amplitude which is again pushed away by the front, and the front therefore seems to be nonlinearly stable in weighted norms. For $\gamma_1 = -8$, on the other hand, the periodic patterns in the wake have much larger amplitude and lock to the front, which is therefore no longer nonlinearly stable.

REFERENCES

[1] T. Brand, M. Kunze, G. Schneider, and T. Seelbach, *Hopf bifurcation and exchange of stability in diffusive media*, Arch. Ration. Mech. Anal., 171 (2004), pp. 263–296.

[2] J. Bricmont and A. Kupiainen, *Renormalization group and the Ginzburg-Landau equation*, Comm. Math. Phys., 150 (1992), pp. 193–208.

[3] R. J. Briggs, *Electron-Stream Interaction with Plasmas*, MIT Press, Cambridge, MA, 1964.

[4] P. Collet and J.-P. Eckmann, *Instabilities and Fronts in Extended Systems*, Princeton University Press, Princeton, NJ, 1990.

[5] P. Collet and J.-P. Eckmann, *A rigorous upper bound on the propagation speed for the Swift–Hohenberg and related equations*, J. Statist. Phys., 108 (2002), pp. 1107–1124.

[6] U. Ebert and W. van Saarloos, *Front propagation into unstable states: Universal algebraic convergence towards uniformly translating pulled fronts*, Phys. D, 146 (2000), pp. 1–99.

[7] J.-P. Eckmann and G. Schneider, *Nonlinear stability of bifurcating front solutions for the Taylor–Couette problem*, ZAMM Z. Angew. Math. Mech., 80 (2000), pp. 745–753.

[8] J.-P. Eckmann and G. Schneider, *Non-linear stability of modulated fronts for the Swift–Hohenberg equation*, Comm. Math. Phys., 225 (2002), pp. 361–397.

[9] T. Gallay, *Local stability of critical fronts in parabolic partial differential equations*, Nonlinearity, 7 (1994), pp. 741–764.

[10] T. Gallay, G. Schneider, and H. Uecker, *Stable transport of information near essentially unstable localized structures*, Discrete Contin. Dyn. Syst. Ser. B, 4 (2004), pp. 349–390.

[11] A. R. Ghazaryan, *Nonlinear Convective Instability of Fronts: A Case Study*, Ph.D. thesis, Ohio State University, http://www.ohiolink.edu/etd/view.cgi?osu1117552079.

[12] D. Henry, *Geometric Theory of Semilinear Parabolic Equations*, Lecture Notes in Math. 840, Springer-Verlag, New York, 1981.

[13] E. Hille, *Lectures on Ordinary Differential Equations*, Addison–Wesley, Reading, MA, 1969.

[14] A. L. Kay and J. A. Sherratt, *On the persistence of spatiotemporal oscillations generated by invasion*, IMA J. Appl. Math., 63 (1999), pp. 199–216.

[15] W. Mazin, K. E. Rasmussen, E. Mosekilde, P. Borckmans, and G. Dewel, *Pattern formation in the bistable Gray-Scott model*, Math. Comput. Simulation, 40 (1996), pp. 371–396.

[16] A. Mielke and G. Schneider, *Attractors for modulation equations on unbounded domains—existence and comparison*, Nonlinearity, 8 (1995), pp. 743–768.

[17] K. J. Palmer, *Exponential dichotomies and Fredholm operators*, Proc. Amer. Math. Soc., 104 (1988), pp. 149–156.

[18] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[19] R. L. Pego and M. I. Weinstein, *Asymptotic stability of solitary waves*, Comm. Math. Phys., 164 (1994), pp. 305–349.

[20] W. van Saarloos, *Front propagation into unstable states*, Phys. Rep., 386 (2003), pp. 29–222.

[21] B. Sandstede and A. Scheel, *Absolute and convective instabilities of waves on unbounded and large bounded domains*, Phys. D, 145 (2000), pp. 233–277.

[22] B. Sandstede and A. Scheel, *Spectral stability of modulated travelling waves bifurcating near essential instabilities*, Proc. Roy. Soc. Edinburgh Sect. A, 130 (2000), pp. 419–448.
[23] B. Sandstede and A. Scheel, *Essential instabilities of fronts: Bifurcation, and bifurcation failure*, Dyn. Syst., 16 (2001), pp. 1–28.
[24] D. H. Sattinger, *Weighted norms for the stability of travelling waves*, J. Differential Equations, 25 (1977), pp. 130–144.
[25] J. A. Sherratt, *Invading wave fronts and their oscillatory wakes linked by a modulated travelling phase resetting wave*, Phys. D, 117 (1998), pp. 145–166.
[26] J. A. Sherratt, B. T. Eagan, and M. A. Lewis, *Oscillations and chaos behind predator-prey invasion: Mathematical artifact or ecological reality?*, Phil. Trans. Roy. Soc. London B, 352 (1997), pp. 21–38.
[27] B. P. Sommeijer, L. F. Shampine, and J. G. Verwer, RKC*: An explicit solver for parabolic PDEs*, J. Comput. Appl. Math., 88 (1998), pp. 315–326.

SIAM J. APPLIED DYNAMICAL SYSTEMS
Vol. 6, No. 2, pp. 348–377

A Multilegged Modular Robot That Meanders: Investigation of Turning Maneuvers Using Its Inherent Dynamic Characteristics*

Shinya Aoi†, Hitoshi Sasaki†, and Kazuo Tsuchiya†

Abstract. This paper deals with the motion of a multilegged modular robot. The robot consists of a set of homogenous modules, each of which has a body and two legs and is connected to the others through a three-degree-of-freedom rotary joint. The leg joints are manipulated to follow periodic desired trajectories, and the joints between the modules act like a passive spring with a damper. This robot has characteristic dynamic properties. Specifically, a straight walk naturally turns into a meandering walk by changing the compliance of the joints between the modules without incorporation of any oscillatory inputs. We first show that this transition is excited due to a Hopf bifurcation, based on a numerical simulation and Floquet analysis. Following that, we examine whether the maneuverability and agility of the robot increase by utilizing the dynamic characteristics inherent in the robot. In particular, we conduct an experiment in which the robot pursues a target moving across the floor. We propose a simple controller to accomplish the task and achieve high maneuverability and agility by making the most of the robot's dynamic features.

Key words. multilegged modular robot, turning maneuvers, meandering walk, maneuverability, Hopf bifurcation, Floquet analysis

AMS subject classifications. 70E60, 70K50

DOI. 10.1137/060664756

1. Introduction.

Modular robots consist of a set of robotic modules that change the configuration and strength of their connection, which allows them to deal with a wide variety of tasks. In the literature, many modular robots have been developed that have capabilities such as self-reconfiguration, fault tolerance, and locomotion [13, 21, 40, 47, 58, 80, 81, 82]. In particular, legged-type modular robots, which are specialized for locomotion, have a high possibility of moving across uneven terrain and high adaptability to various environments [16, 36, 37, 39, 74]. They are expected to display their great ability in a lot of places such as in space [66, 81].

However, it is still difficult to create sophisticated legged robots and their control systems. In particular, (1) the robot is a mechanical system with many degrees of freedom, composed of many links that are connected with others by joints, some of which are redundant in achieving walking. As Bernstein [11] pointed out, the essential problem is how to coordinate motion.

*Received by the editors July 11, 2006; accepted for publication (in revised form) by R. Murray January 2, 2007; published electronically May 21, 2007. This research was supported in part by the Center of Excellence for Research and Education on Complex Functional Mechanical Systems (COE program of the Ministry of Education, Culture, Sports, Science and Technology, Japan) and by a Grant-in-Aid for Scientific Research on Priority Areas "Emergence of Adaptive Motor Function through Interaction between Body, Brain and Environment" from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

http://www.siam.org/journals/siads/6-2/66475.html

†Department of Aeronautics and Astronautics, Graduate School of Engineering, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan (shinya_aoi@kuaero.kyoto-u.ac.jp, hitoshi_sasaki@kuaero.kyoto-u.ac.jp, tsuchiya@kuaero.kyoto-u.ac.jp).

(2) The leg motion consists of the swing and stance phases. The swing leg lands on the ground and in turn becomes the stance leg. Therefore, periodically and intermittently, the legs receive reaction forces from the ground. In other words, the condition of foot-to-ground contact is changeable, resulting in changes of the dynamics that govern the walking motion and influence the walking stability. To overcome these difficulties, studies have been widely carried out based on the model-based approach [12, 27, 53]. In this approach, the robot motion is basically generated by the inverse kinematics and kinetics, for example, by calculating the foot landing positions to keep the walk stable and then by computing the joint motions. However, complicated computations are required, as is precise modeling of the robot and environment, which restricts the possibility of attaining adaptability and robustness. In addition to these difficulties, a robot that possesses many legs has the following characteristic problem: since its many legs are in contact with the ground and support the robot, they keep the robot from falling over. However, all those contact legs also keep the robot from accomplishing maneuverable and agile motions such as a quick turn. Therefore, it is difficult to simply design a locomotion control system and achieve high adaptability and maneuverability of the robot motion.

This is in contrast to the millions of animal species that adapt themselves to various environments by manipulating their complicated and redundant musculoskeletal system, giving them marvelous maneuverability and agility. Recently, many researchers have developed biologically inspired robots and aimed to clarify the control mechanisms of animals based on the constructive approach. In particular, neurophysiological studies have revealed that animal walking is generated by central pattern generators (CPGs) [32, 33, 52]. CPGs comprise a set of neural oscillators present in the spinal cord, generating rhythmic signals that activate their limbs. The CPGs modulate signal generation in response to the sensory signals, resulting in adaptive motions. The CPGs are widely modeled using nonlinear oscillators [68, 69, 70], and based on such CPG models many locomotion robots and their control systems have been developed. For example, Fukuoka, Kimura, and Cohen [22] and Kimura et al. [43, 44] created quadruped robots and achieved adaptive and agile walking on an irregular terrain by employing CPG and reflex models. Tsujita, Tsuchiya, and Onat [75] proposed a locomotion control system for a quadruped robot using nonlinear oscillators and built a quadruped robot that obtained adaptive walking by changing the gait pattern depending on the walking speed and environmental variations. Inagaki et al. [36, 37] developed a six-legged modular robot and generated its gait pattern through their CPG model. Ijspeert et al. [35], Crespi et al. [18], Ijspeert, Crespi, and Cabelguen [34], and Inoue, Ma, and Jin [38] created salamander and snake-like modular robots that accomplished a serpentine meandering locomotion through a neural oscillator network. Lewis et al. [48], Nakanishi et al. [51], and Aoi and Tsuchiya [5, 6, 8] realized adaptive walking of biped robots using their CPG models and nonlinear oscillators.

Also, from the field of neurophysiology, it has been revealed that, as well as rhythm control, muscle tone control has an important role in generating adaptive motions [50, 57, 71, 72], suggesting the importance of compliance in locomotion. Actually, many studies on robotics demonstrated the essential roles of the compliance. Specifically, by appropriately employing the mechanical compliance of the robots, simple control systems attained highly adaptive, robust, and agile motions, especially in hexapod robots [2, 14, 15, 19, 56, 59, 60], quadruped robots [22, 43, 54, 55], and biped robots [73, 79].

Animals generate their motions by skillfully applying the intrinsic characteristics of their musculoskeletal system. In particular, many researchers have used simple models and analyzed self-stabilizing properties embedded in animals' musculoskeletal systems that accomplish stable motions without depending on external sensory information. Wagner and Blickhan [78] exhibited such self-stabilizing oscillatory leg movement by using such muscle properties as the force-length relationship, the force-velocity relationship, and muscle geometry. For a running model, Geyer, Seyfarth, and Blickhan [25] employed a simple spring-mass model and Ghigliazza et al. [26] used a spring-loaded inverted pendulum (SLIP). They demonstrated that by appropriately controlling the leg angle at touchdown their models could obtain an asymptotically stable running motion. Schmitt et al. [61, 62, 63, 64], meanwhile, analyzed the hexapod walking of a cockroach by using a simple model composed of a rigid body and a pair of massless elastic legs.

As well as with animals, many studies have been carried out to elucidate such self-stabilization inherent in locomotion robots also by employing simple models. Garcia et al. [24] investigated the asymptotic stability of passive dynamic walking, originally performed by McGeer [49] using hardware experiments and numerical simulations. Altendorfer, Koditschek, and Holmes [3, 4] analyzed the running motion of a hexapod robot, RHex [2, 59, 60], which has passively compliant legs, and Aoi and Tsuchiya [7, 9] examined biped robot walking driven by a rhythmic signal from an oscillator.

Dynamic characteristics such as stability must strongly affect the maneuverability and agility of locomotion. For example, cockroaches are highly agile and have a great range of maneuverability [23, 41, 42, 45]. Schmitt et al. [61, 62, 63, 64] simply modeled the hexapod walking of a cockroach, and then analytically demonstrated that it successfully achieves a quick turn by virtue of destabilizing its straight walking motion by changing its dynamic features. It would be very interesting and helpful if we attained such maneuverability and agility of locomotion robots by using the dynamic properties inherent in the robots.

In this paper, we deal with a multilegged modular robot whose model is introduced in section 2 and in particular we study its rudimentary locomotion. The robot consists of a set of homogenous modules, each of which has a body and two legs and is connected to the others through a three-degree-of-freedom (-DOF) rotary joint. The robot has a simple controller that generates periodic leg trajectories. The leg joints are manipulated to follow periodic desired trajectories and each joint between modules acts like a passive spring with a damper. This robot features characteristic dynamic properties. Specifically, a straight walk by the robot naturally turns into a meandering walk by changing the strength of the connection between the modules without actually incorporating any oscillatory inputs. That is, the dynamic stability of a straight walk varies depending on the compliance of the joints between the modules. In section 3, we show that the transition from a straight to a meandering walk is excited due to a Hopf bifurcation, based on a numerical simulation and Floquet analysis.

As described above, it is difficult for a locomotion robot that has many legs to achieve agile motions such as a quick turn because of the contact legs and motion planning. Since such a motion is a fundamental behavior for a locomotion robot, its dynamic characteristics should be thoroughly analyzed and the problem should be solved. Although the model-based approach is generally used, it requires precise modeling and complicated computations, preventing one from achieving adaptive locomotion and simple control system. In this paper,

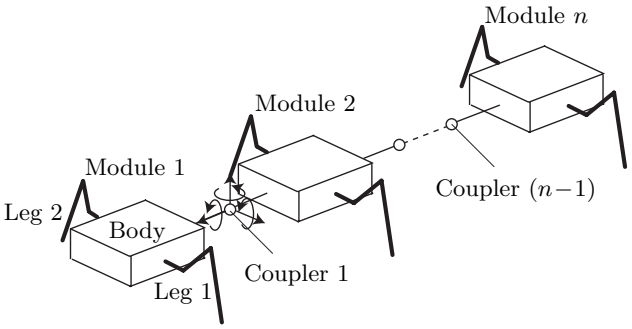


Figure 1. *Schematic model of a multilegged modular robot.*

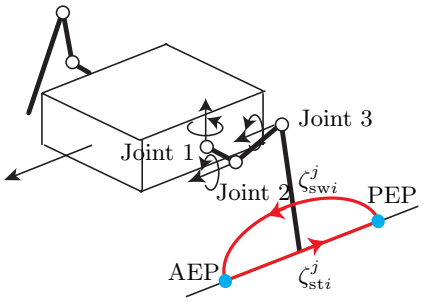


Figure 2. *Nominal trajectory of the leg.*

we especially focus on the dynamic characteristics embedded in the multilegged modular robot as one of the solutions to the problem. In particular, in section 4, we investigate whether the maneuverability and agility of the robot increase by using the dynamic characteristics. We conduct an experiment in which the robot pursues a target moving across the floor. We then propose a simple controller to accomplish the task and achieve high maneuverability and agility by making the most of the robot's dynamic features. Sections 5 and 6 present the discussion and conclusion, respectively.

2. A multilegged modular robot.

2.1. Robot model. Figure 1 shows a schematic diagram of the multilegged modular robot considered in this paper. The robot has n homogenous modules, each with one body and two legs. Each leg consists of three links that are connected to each other through a one-DOF rotational joint (see Figure 2). The legs are articulated to the body also by a one-DOF rotational joint. Each module is connected to the next through a coupler composed of roll, pitch, and yaw joints, with each joint manipulated by a motor. The modules are enumerated from Module 1 to Module n, and the coupler between Module i and Module $(i+1)$ is numbered Coupler i $(i = 1, \ldots, n-1)$. The left and right legs are numbered Legs 1 and 2, respectively. The joints and links of each leg are numbered from the side of the body as Joints 1, 2, and 3 and Links 1, 2, and 3, respectively. The position vector of the body of Module 1 is given by vector $[\, x_1 \, x_2 \, x_3 \,]$ expressed on the ground, where x_1 and x_3 are toward the nominal direction of locomotion and the vertical direction, respectively. The posture of the body of Module 1 is

given by Euler angles $[\,\theta_{11}\,\theta_{12}\,\theta_{13}\,]$, where θ_{11}, θ_{12}, and θ_{13} correspond to roll, pitch, and yaw angles, respectively. Similarly, angles θ_{im} $(i = 2, \ldots, n,\ m = 1, 2, 3)$ are the components of relative angles of Module i with respect to Module $(i-1)$, which correspond to the angles of Coupler $(i-1)$. Angles θ_{ik}^{j} $(i = 1, \ldots, n,\ j = 1, 2,\ k = 1, 2, 3)$ are the relative angles of Joint k of Leg j of Module i.

State variable $q \in \mathbb{R}^{3+9n}$ is defined as $q^{\mathrm{T}} = [\,x_m\,\theta_{im}\,\theta_{ik}^{j}\,]$ $(i = 1, \ldots, n,\ j = 1, 2,\ k = 1, 2, 3,\ m = 1, 2, 3)$. An equation of motion for the state variable is derived using the Lagrangian formulation by

$$K(q)\ddot{q} + h(q, \dot{q}) = g(q) + u + \lambda, \tag{2.1}$$

where $K(q) \in \mathbb{R}^{(3+9n)\times(3+9n)}$ is the inertia matrix, $h(q, \dot{q}) \in \mathbb{R}^{3+9n}$ is the nonlinear term that includes Coriolis and centrifugal forces, $g(q) \in \mathbb{R}^{3+9n}$ is the gravity term, $u \in \mathbb{R}^{3+9n}$ is the input torque term, and $\lambda \in \mathbb{R}^{3+9n}$ is the reaction force from the ground. The ground is modeled as a spring with a damper in the vertical direction and a viscous damper in the horizontal direction (see Appendix A). In this paper, numerical simulations are carried out based on this equation of motion.

2.2. Clock-driven leg controller. The robot's walking motion is generated by the motions of the leg joints and coupler joints. The main role of the legs in walking is to support the robot's weight and provide propulsive forces to move forward. We have designed the physical kinematics of the legs, whereby the leg joints are manipulated by motors using a proportional-derivative (PD) controller. Specifically, we employ a simple clock-driven, open-loop gait, and the leg joints are controlled by incorporating periodic desired angles. Therefore, input torque u_{ik}^{j} $(i = 1, \ldots, n,\ j = 1, 2,\ k = 1, 2, 3)$ at Joint k of Leg j of Module i is given by

$$\begin{aligned} u_{ik}^{j} &= -\kappa_{ik}^{j}(\theta_{ik}^{j} - \hat{\theta}_{ik}^{j}) - \sigma_{ik}^{j}(\dot{\theta}_{ik}^{j} - \dot{\hat{\theta}}_{ik}^{j}), \\ &\qquad i = 1, \ldots, n,\ j = 1, 2,\ k = 1, 2, 3, \end{aligned} \tag{2.2}$$

where $\hat{\theta}_{ik}^{j}$ is the periodic desired angle and κ_{ik}^{j} and σ_{ik}^{j} are feedback gains $(i = 1, \ldots, n,\ j = 1, 2,\ k = 1, 2, 3)$. On the other hand, the coupler joints are controlled with a desired angle maintained at zero to generate a straight walk. Thus, input torque u_{ij} $(i = 2, \ldots, n,\ j = 1, 2, 3)$ at angles of Coupler $(i-1)$ is given by

$$u_{ij} = -\kappa_{ij}\theta_{ij} - \sigma_{ij}\dot{\theta}_{ij}, \qquad i = 2, \ldots, n,\ j = 1, 2, 3, \tag{2.3}$$

where κ_{ij} and σ_{ij} are feedback gains $(i = 2, \ldots, n,\ j = 1, 2, 3)$. Therefore, the coupler joints act like a passive spring with a damper.

To design the periodic desired angles of the leg joints, we first introduce an oscillator for each leg. In particular, we introduce Oscillator i, j $(i = 1, \ldots, n,\ j = 1, 2)$ for Leg j of Module i. Oscillator i, j has a phase ϕ_i^j whose angular velocity is constant. Second, we design nominal trajectory ζ_i^j $(i = 1, \ldots, n,\ j = 1, 2)$ of the tip of Leg j of Module i in the sagittal plane as a function of phase ϕ_i^j, that is, $\zeta_i^j = \zeta_i^j(\phi_i^j)$. Trajectory ζ_i^j is expressed in the body of Module i and consists of trajectories ζ_{swi}^j for the swing phase and ζ_{sti}^j for the stance phase $(i = 1, \ldots, n,\ j = 1, 2)$ (see Figure 2). Trajectory ζ_{swi}^j is composed of half of an elliptic curve

that includes the anterior extreme position (AEP) and the posterior extreme position (PEP). Note that the distance between points AEP and PEP implies nominal stride s. Trajectory $\zeta^j_{\mathrm{st}i}$ is comprised of a straight line that also involves points AEP and PEP. During the stance phase, the tip of the leg moves at constant speed v with respect to the body in the opposite walking direction, which is given by

$$v = s/(\beta\tau), \tag{2.4}$$

where τ is the nominal step cycle and β indicates the nominal duty ratio that expresses the ratio between the nominal stance phase duration and the nominal step cycle. Since we consider that the tip of the leg is constrained on the ground and hardly slips relative to the ground during the stance phase, the body moves in the walking direction at nominal locomotion speed v with respect to the ground. In light of the above description, trajectory ζ^j_i is given by (see details in [5])

$$\zeta^j_i(\phi^j_i) = \begin{cases} \zeta^j_{\mathrm{sw}i}(\phi^j_i), & 0 \le \phi^j_i < \phi_{\mathrm{a}}, \\ \zeta^j_{\mathrm{st}i}(\phi^j_i), & \phi_{\mathrm{a}} \le \phi^j_i < 2\pi, \end{cases} \qquad i = 1, \ldots, n,\ j = 1, 2, \tag{2.5}$$

where $\phi_{\mathrm{a}} = 2\pi(1-\beta)$, which indicates the nominal phase value at point AEP (0 at point PEP). When trajectory ζ^j_i is on point PEP at $t = 0$, it arrives at point AEP at $t = (1-\beta)\tau$ through the swing phase and turns into the stance phase. It then reaches point PEP at $t = \tau$ and returns to the swing phase. Finally, based on the inverse kinematics, we obtain the desired angles $\hat{\theta}^j_{ik}$ ($i = 1, \ldots, n$, $j = 1, 2$, $k = 1, 2, 3$) of Joint k of Leg j of Module i as the function of phase ϕ^j_i.

In numerical simulations, regarding the gait pattern of the robot, the unilateral legs on adjacent modules move out of phase with each other. That is, the phases of the oscillators have relationships such that $\phi^j_i - \phi^j_{i+1} = \pi$ ($i = 1, \ldots, n-1$, $j = 1, 2$). On the other hand, the relationship between the contralateral legs on each module is determined from the phase relationship $\phi^1_i - \phi^2_i = \Delta\phi_{\mathrm{cntrl}}$ ($i = 1, \ldots, n$). For example, when $\Delta\phi_{\mathrm{cntrl}} = \pi$, the contralateral legs on each module move out of phase with each other (which we call an "anti-phase gait pattern"), and when $\Delta\phi_{\mathrm{cntrl}} = 0$, the contralateral legs on each module move in phase (we call this an "in-phase gait pattern"). Nominal stride s, duty ratio β, and step cycle τ of each leg are set to 5 cm, 0.5, and 0.5 s, respectively. In this case, nominal locomotion speed v becomes equivalent to 0.2 m/s from (2.4). Point AEP of both legs of each module is located 4.5 cm ahead of and 7.5 cm outside of the center of the module in the nominal direction of locomotion.

3. Dynamic properties of the multilegged modular robot.

3.1. Transition from a straight to a meandering walk. This robot has interesting and essential characteristics in locomotion. Since the couplers act like a passive spring with a damper, the robot is able to achieve a straight walk as long as it moves its legs in a similar manner. However, it is clear that when we decrease the feedback gains of the couplers' yaw joints, a meandering walk appears beyond a critical point without incorporation of any

Table 1

Physical parameters of the multilegged modular robot. The same values are used in each module.

Link	Parameter	Value
Body	Mass [kg]	0.6
	Length [m]	0.2
	Width [m]	0.15
	Inertia [$\times 10^{-3}$kgm^2]	3.1
Link 1	Mass [kg]	0.02
	Length [m]	0.015
Link 2	Mass [kg]	0.03
	Length [m]	0.045
Link 3	Mass [kg]	0.05
	Length [m]	0.15
Motor	Gear ratio	100
	Rotor inertia [$\times 10^{-7}$kgm^2]	9.7

oscillatory inputs into the couplers. This means that the robot's walking motion naturally varies according to changes in its mechanical properties.

In this section, we present detailed results obtained through numerical simulations using six modules ($n = 6$) and an anti-phase gait pattern ($\Delta\phi_{\text{cntrl}} = \pi$). Table 1 displays the physical parameters of the multilegged modular robot used in the numerical simulations. The center of mass of each module is located 0.9 cm ahead of the center of the module in the nominal direction of locomotion. The damping coefficient of the ground in the horizontal direction is set to 19.6 Ns/m. To change the mechanical features of the robot, we parameterize proportional feedback gain κ_{i3} and derivative feedback gain σ_{i3} ($i = 2, \ldots, n$) of the couplers' yaw joints by using parameter f,

$$\kappa_{i3} = \kappa_0(2\pi f)^2, \ \sigma_{i3} = 2\kappa_0\zeta_0(2\pi f), \quad i = 2, \ldots, n, \tag{3.1}$$

where κ_0 and ζ_0 are set to 0.0097 and 0.8, respectively, to provide adequate damping. Note that for other joints high feedback gains are used and parameter f is set to 10 (e.g., $\kappa_{i1,2}^{-1} = 0.0261$, $i = 2, \ldots, n$). Also note that this change in the feedback gains indicates a change in the joints' compliance. Figure 3(a) shows yaw angle θ_{43} of Coupler 3 at $\phi_1^1 = 0$ with respect to the reciprocal of gain parameter κ_{i3}^{-1}. Figure 3(b) shows the power spectrum of yaw angle θ_{43} of Coupler 3 using $\kappa_{i3}^{-1} = 10$ and $\kappa_{i3}^{-1} = 17$. These figures reveal that undulatory motion is excited over a bifurcation point. Figure 3(a) also implies that since the angles are plotted with respect to each step cycle of the leg motion, this meandering motion is not synchronized with the leg motion. Therefore, it has an independent frequency (see details in Figure 7(c) shown below). Figures 4(a) and (b) are snapshots of the front and top views, respectively, of the walking motion using $\kappa_{i3}^{-1} = 21$, showing that wavy motion appears. A video of this behavior is available as 66475_01.avi [10.3MB].

3.2. Investigation of the transition mechanism. As shown above, a straight walk naturally turns into a meandering walk through changes in the robot's mechanical characteristics. This robot receives dynamical interactions between the contact legs and the ground. Since the robot moves the legs parallel to their modules, it receives reaction forces from the ground also parallel to their modules. Therefore, the robot is supposed to walk straight. However,

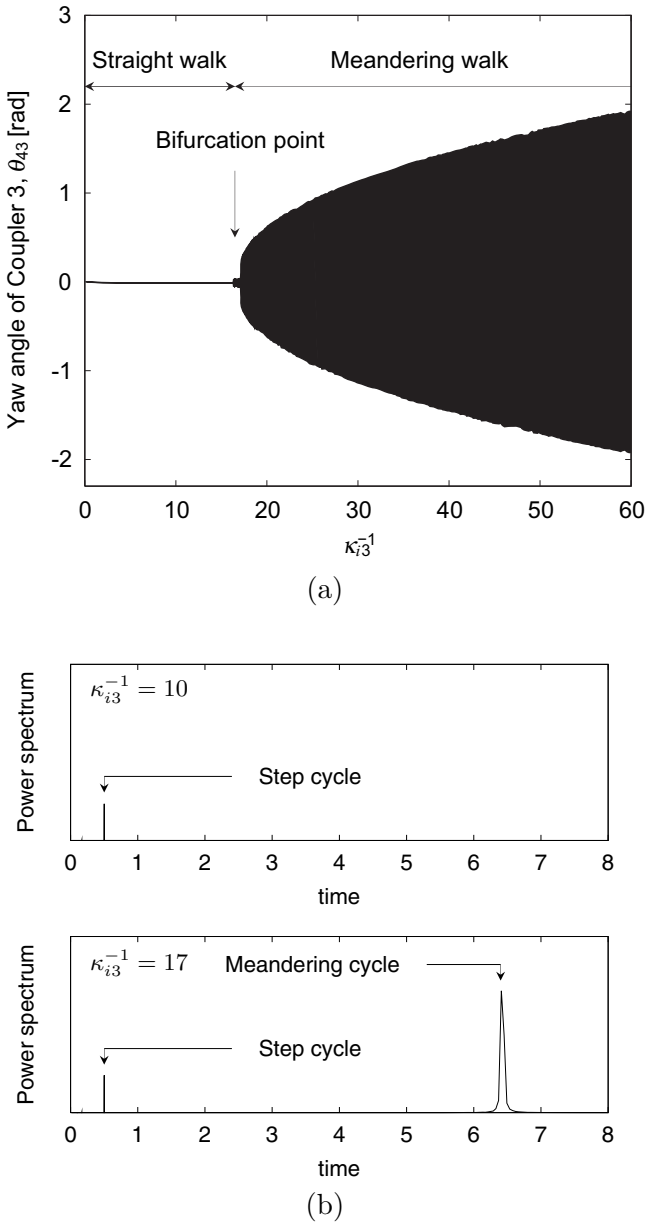
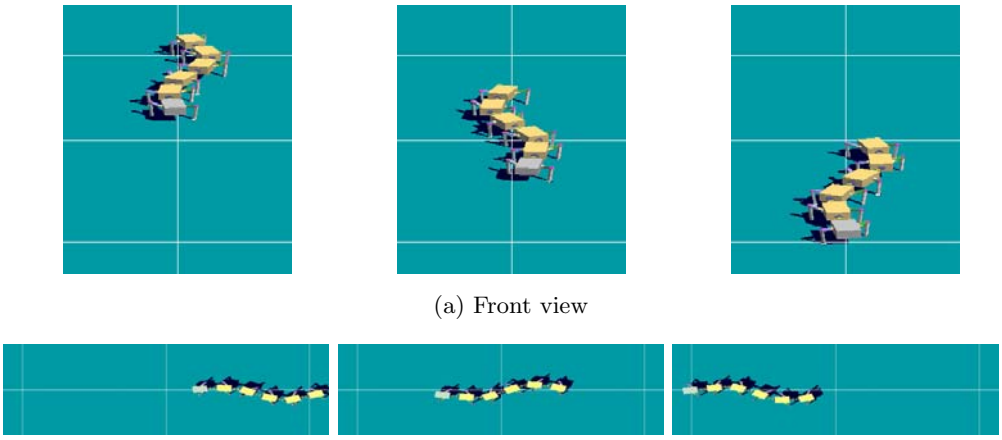


Figure 3. *Bifurcation of the walking motion.* (a) *Yaw angle of Coupler* 3, θ_{43}, *at* $\phi_1^1 = 0$ *versus the reciprocal of gain parameter* κ_{i3}^{-1}. (b) *Power spectrum of angle* θ_{43}.

when the feedback gains of the couplers' yaw joints decrease beyond the critical threshold, lateral motions appear and wavy motion is generated. Why does this happen? In this section we investigate its mechanism in detail.

In a straight walk, the couplers' yaw joints hardly move and the modules keep straight, suggesting that zero is stable for the joint motions. On the contrary, the joints achieve periodic oscillation in a meandering walk. Depending on the feedback gains, a straight walk



(a) Front view

(b) Top view

Figure 4. *Meandering motion. See also the accompanying animation* (66475_01.avi [10.3MB]).

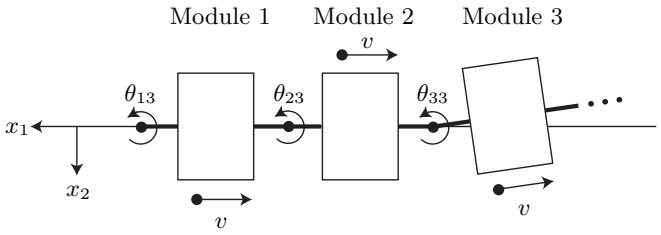


Figure 5. *Overhead view of the simplified model. Black dots beside the bodies indicate the positions of the leg tips during the stance phase.*

turns into a meandering walk. These facts imply that when the gains are large enough a straight walk is stable. On the other hand, when the gains decrease beyond a critical value, a Hopf bifurcation occurs and the straight walk is destabilized, exciting undulatory motion. As a result, the transition from a straight to a meandering walk takes place as obtained in numerical simulations. Here, we examine whether this suggestion actually explains the transition mechanism.

Although the above suggestion comes from simulation results based on the robot model described in section 2.1, we simplify the robot model by assuming the following for its straight walk to extract the essence of the transition mechanism:

1. The leg mass is too small in comparison to the body mass and the leg joints completely follow the kinematically designed trajectories as shown in Figure 2.
2. Many legs are in contact with the ground and support the robot. Since the legs' nominal motions do not cause the up-and-down, roll, and pitch motions of the robot, those motions are sufficiently small with respect to other motions and can be ignored.
3. The robot walks at constant speed v in the nominal direction of locomotion.

Figure 5 shows an overhead view of the simplified model, where black dots beside the bodies indicate the positions of the leg tips during the stance phase. Note that the legs receive the forces from the ground only when they are in the stance phase.

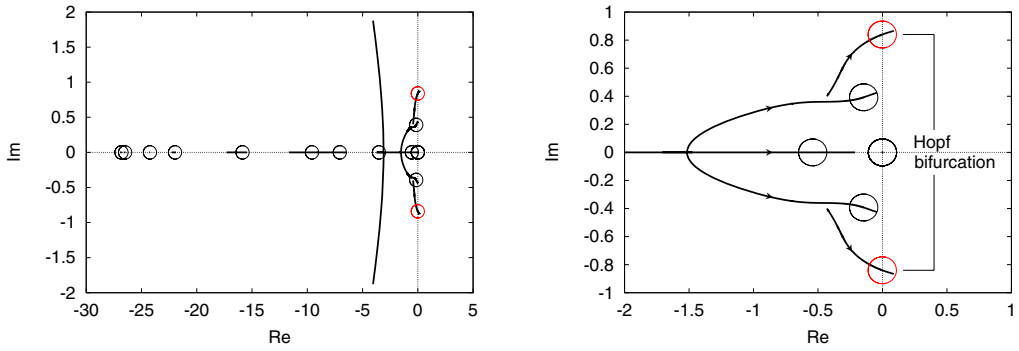


Figure 6. *Trajectories of Floquet exponents.* (a) *Trajectories of all the Floquet exponents.* (b) *Enlarged detail.*

Under these assumptions, state variable $q \in \mathbb{R}^{n+2}$ is redefined as $q^{\mathrm{T}} = [\, x_1\, x_2\, \theta_{13} \cdots \theta_{n3} \,]$. Then, we define state $\xi \in \mathbb{R}^{2n+4}$ as $\xi^{\mathrm{T}} = [\, \dot{q}^{\mathrm{T}}\, q^{\mathrm{T}} \,]$. In a straight walk, state ξ can be written as $\xi_{\mathrm{str}}^{\mathrm{T}} = [\, v\ 0 \cdots 0\ vt + x_{10}\ 0 \cdots 0 \,]$, where x_{10} is the state of x_1 at $t = 0$. Perturbed state $\xi \in \mathbb{R}^{2n+4}$ from a straight walk is defined as $\xi = \xi_{\mathrm{str}} + \delta\xi$, where $\delta\xi \in \mathbb{R}^{2n+4}$ is the perturbation. By contracting the equation of motion (2.1) and then linearizing it around state ξ_{str}, we obtain (see Appendices B and C)

$$\delta\dot{\xi} = A(t)\delta\xi, \tag{3.2}$$

where matrix $A(t) \in \mathbb{R}^{(2n+4)\times(2n+4)}$ is periodic and $A(t+\tau) = A(t)$ for step cycle τ of the leg motion. Note that matrix $A(t)$ has no element that contains x_{10}.

Based on the Floquet theory [20], we analyze the stability of a straight walk, where for the simplified model the mass and inertia of each module are set to 0.8 kg and 4.2×10^{-3} kgm^2, respectively. Figures 6(a) and (b) show the trajectories of the Floquet (characteristic) exponents while the feedback gains of the yaw joints of the couplers change. Specifically, Figure 6(a) shows all the trajectories of the Floquet exponents, revealing that a Hopf bifurcation occurs by crossing the imaginary axis. Displayed circles indicate all the $(2n + 4)$ Floquet exponents when the Hopf bifurcation takes place, where red circles correspond to the Hopf bifurcation. Figure 6(b) shows trajectories in detail by focusing on the vicinity of the imaginary axis.

The above result clarifies that a Hopf bifurcation occurs and a straight walk becomes unstable according to the feedback gains. However, this result comes from the simplified model. Next, we verify whether this Hopf bifurcation actually explains the transition from a straight to a meandering walk obtained in numerical simulations. Figure 7 shows a comparison between the results of the numerical simulation and the Floquet analysis, where the obtained meandering motion in the numerical simulation is analyzed using fast Fourier transform and the eigenvector corresponding to the destabilized Floquet exponent is used for the Floquet analysis. In particular, Figures 7(a), (b), and (c) illustrate the phase difference with respect to angle θ_{13}, the amplitude ratio between the angles, and the period of the meandering motion, respectively. Although the simplification causes some discrepancies, these results are almost the same in quality and quantity. Therefore, we conclude that the transition from a straight

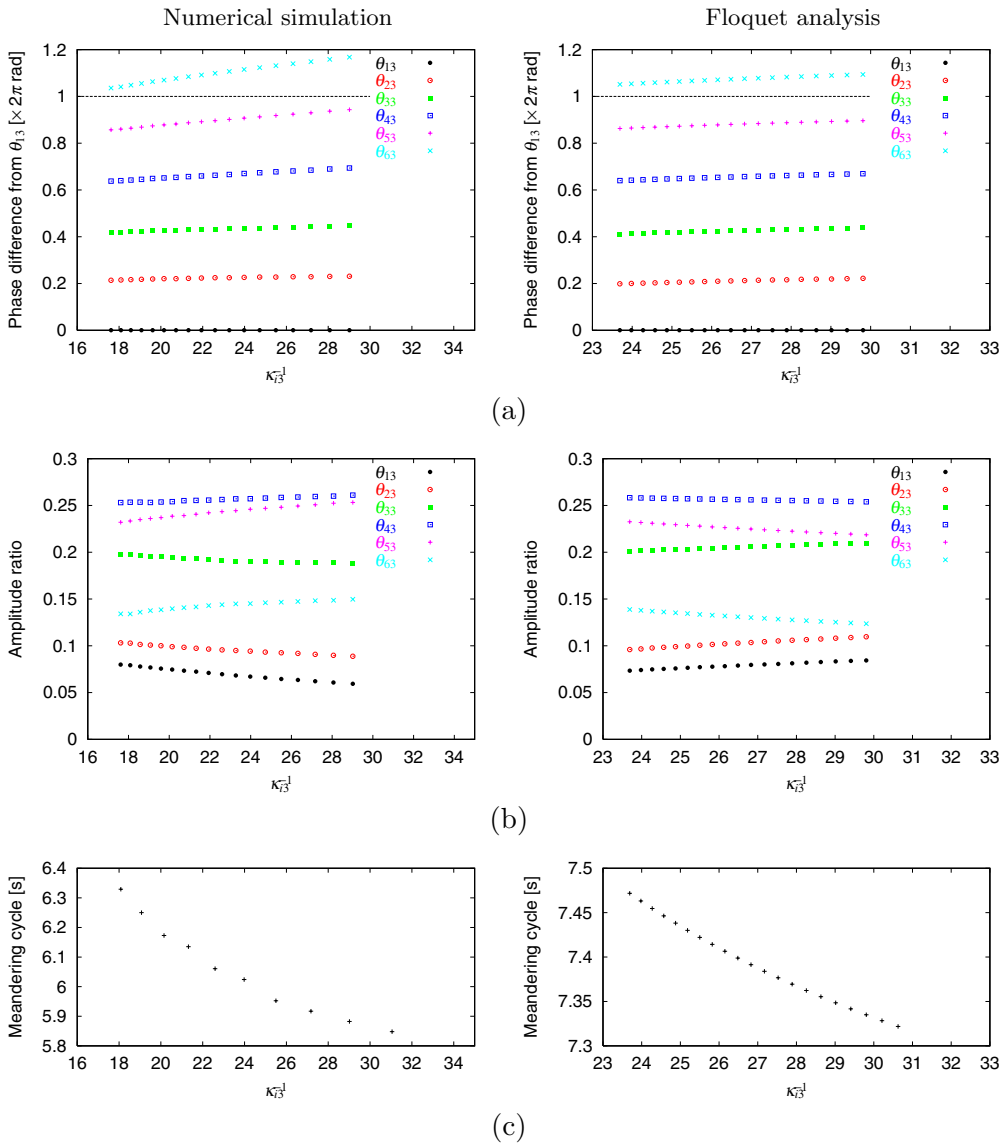


Figure 7. *Comparison between numerical and analytical results.* (a) *Phase difference with respect to* θ_{13}. (b) *Amplitude ratio between angles.* (c) *Period of meandering motion.*

to a meandering walk is caused by a Hopf bifurcation due to the change of the feedback gains of the couplers' yaw joints.

3.3. Investigation of the model dependence of the transition. In the previous sections, we examined the transition from a straight to a meandering walk with respect to a specific robot model using six modules ($n = 6$), an anti-phase gait pattern ($\Delta\phi_{\mathrm{cntrl}} = \pi$), and a viscous damper model for the ground reaction force in the horizontal direction. In this section, we show that this transition is not a particular property caused by the specific robot model but a qualitatively universal characteristic inherent in the robot by similarly using numerical

simulations and Floquet analysis.

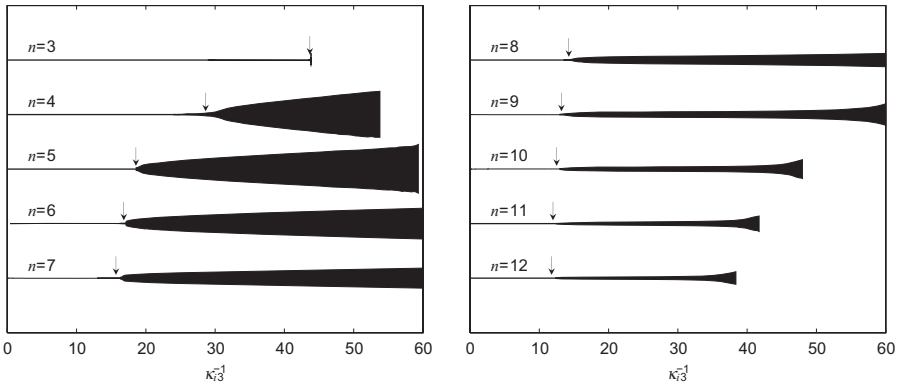
3.3.1. Dependence on the number of modules. Although up to this section we used six modules for the robot model ($n = 6$), we here change the number of modules n and investigate the effect on the characteristics of walking motion. Note that we use an anti-phase gait pattern ($\Delta\phi_{\mathrm{cntrl}} = \pi$) and the viscous damper model as the ground reaction force model in the horizontal direction.

Figure 8 depicts the results. Specifically, Figure 8(a) shows yaw angle θ_{33} of Coupler 2 at $\phi_1^1 = 0$ with respect to the reciprocal of gain parameter κ_{i3}^{-1} for number of modules n (≥ 3) obtained through numerical simulations, where arrows indicate the bifurcation points. This figure reveals that the multilegged modular robot, composed of more than two modules, has the dynamic properties to change the walking motion beyond a critical point of the compliance of the couplers' yaw joints. Figure 8(b) shows all the trajectories of the Floquet exponents with respect to number of modules n achieved by the Floquet analysis. Figure 8(c) illustrates a comparison between the numerical and analytical results, showing the phase difference with respect to angle θ_{13}, the amplitude ratio between the angles, the period of the meandering motion, and the bifurcation point of the reciprocal of gain parameter κ_{i3}^{-1} (cf. Figure 7). These figures indicate that a Hopf bifurcation occurs, resulting in the transition from a straight to a meandering walk. Although such dynamic features as the amplitude of the angles and the wavelength of the meandering motion quantitatively depend on the number of modules, it is clarified that this transition is intrinsic in the robot when $n \geq 3$. Figures 9(a) and (b) are snapshots of the front and top views, respectively, of the walking motion using $n = 12$ and $\kappa_{i3}^{-1} = 29$.

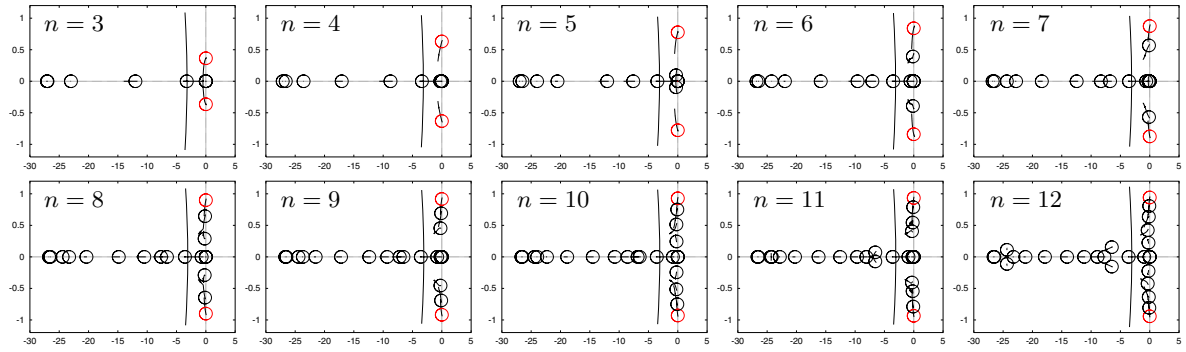
Although this analysis demonstrates that when $n \geq 3$ a straight walk is destabilized and undulatory motion is excited through a Hopf bifurcation, when $n \leq 2$ it is clear that the robot has no such characteristic that changes the walking motion due to a Hopf bifurcation. Figure 10 presents all the trajectories of the eight Floquet exponents using $n = 2$ and verifies that the changes in the couplers' yaw joints cause no bifurcation and that a meandering motion does not take place. In fact, no matter how small the feedback gains of the couplers' yaw joints are in numerical simulations, wavy motion does not appear.

3.3.2. Dependence on the gait pattern. In the previous sections, we used an anti-phase gait pattern ($\Delta\phi_{\mathrm{cntrl}} = \pi$) for the leg motions. In this section, we employ another gait pattern and examine its influence on the walking motion. As described above, when we used an anti-phase gait pattern, a meandering motion was generated without synchronization with the leg motions. Therefore, we expect that the transition from a straight to a meandering walk will appear irrespective of the gait pattern. In particular, this section employs an in-phase gait pattern ($\Delta\phi_{\mathrm{cntrl}} = 0$), in which the contralateral legs on each module move in phase. In contrast to an anti-phase gait pattern, an in-phase gait pattern has completely symmetric leg motions in each module and thus has no direct influence on the motions of the couplers' yaw joints. In that sense, this gait pattern is very useful for examining the mechanism of this transition. This section also uses six modules ($n = 6$) and the viscous damper model for the ground reaction force in the horizontal direction.

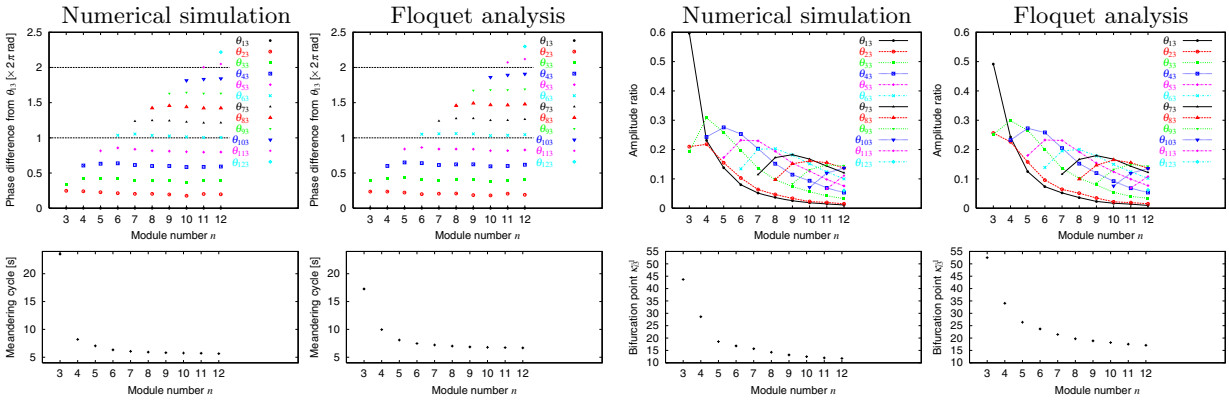
Figure 11(a) shows yaw angle θ_{43} of Coupler 3 at $\phi_1^1 = 0$ with respect to the reciprocal of gain parameter κ_{i3}^{-1} obtained by numerical simulations, illustrating that undulatory mo-



(a)



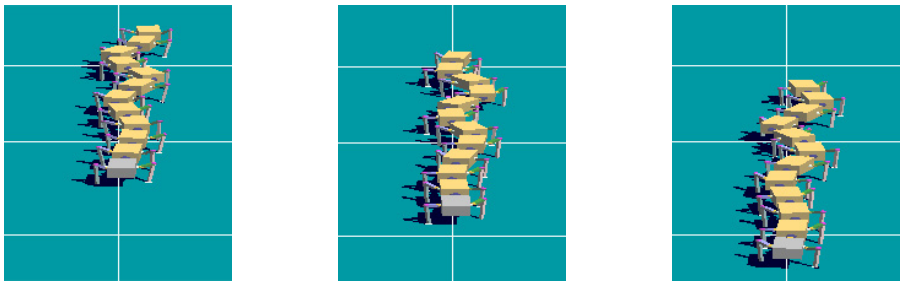
(b)



(c)

Figure 8. *Numerical and analytical results with respect to number of modules n.* (a) *Yaw angle of Coupler* 2, θ_{33}, *at* $\phi_1^1 = 0$. (b) *Trajectories of Floquet exponents.* (c) *Comparison between numerical and analytical results.*

tion appears over a bifurcation point. Figure 11(b) shows all the trajectories of the Floquet exponents calculated in the Floquet analysis. Figure 11(c) presents a comparison between the results of the numerical simulation and the Floquet analysis. Specifically, it displays the phase difference with respect to angle θ_{13}, the amplitude ratio between the angles, and the meandering cycle versus the reciprocal of gain parameter κ_{i3}^{-1} (cf. Figures 7 and 8(c)), verify-



(a) Front view

(b) Top view

Figure 9. *Meandering motion when* $n = 12$.

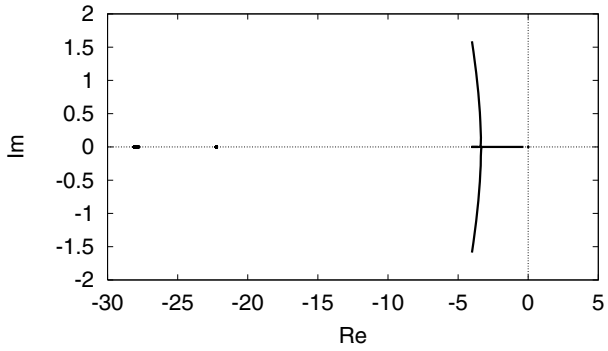


Figure 10. *Trajectories of Floquet exponents when* $n = 2$.

ing that the transition is caused by a Hopf bifurcation. These figures demonstrate that this transition is a universal feature embedded in the robot independent of the gait pattern.

3.3.3. Dependence on the ground reaction force model. Reaction forces from the ground have important roles in locomotion. In particular, for a locomotion robot that has many legs, vertical forces help it to achieve stable walking. With the support of those legs, horizontal forces affect the generation of a planar walking motion of the robot. In the above sections, we used the viscous damper model for the ground reaction force in the horizontal direction, and we demonstrated that the transition from a straight to a meandering walk takes place depending on the mechanical characteristics. The purpose of this section is to explain that this transition is not specific to such a ground reaction force model. To do this, we specifically employ a spring with a damper model and also a Coulomb friction model for the horizontal reaction force, which are often used in numerical simulations [1, 5, 77] (see Appendix A). Here, we use six modules ($n = 6$) and an anti-phase gait pattern ($\Delta\phi_{\mathrm{cntrl}} = \pi$).

Figure 12 shows the numerical results obtained using a spring with a damper model and a Coulomb friction model. Specifically, it exhibits yaw angle θ_{43} of Coupler 3 at $\phi_1^1 = 0$, the

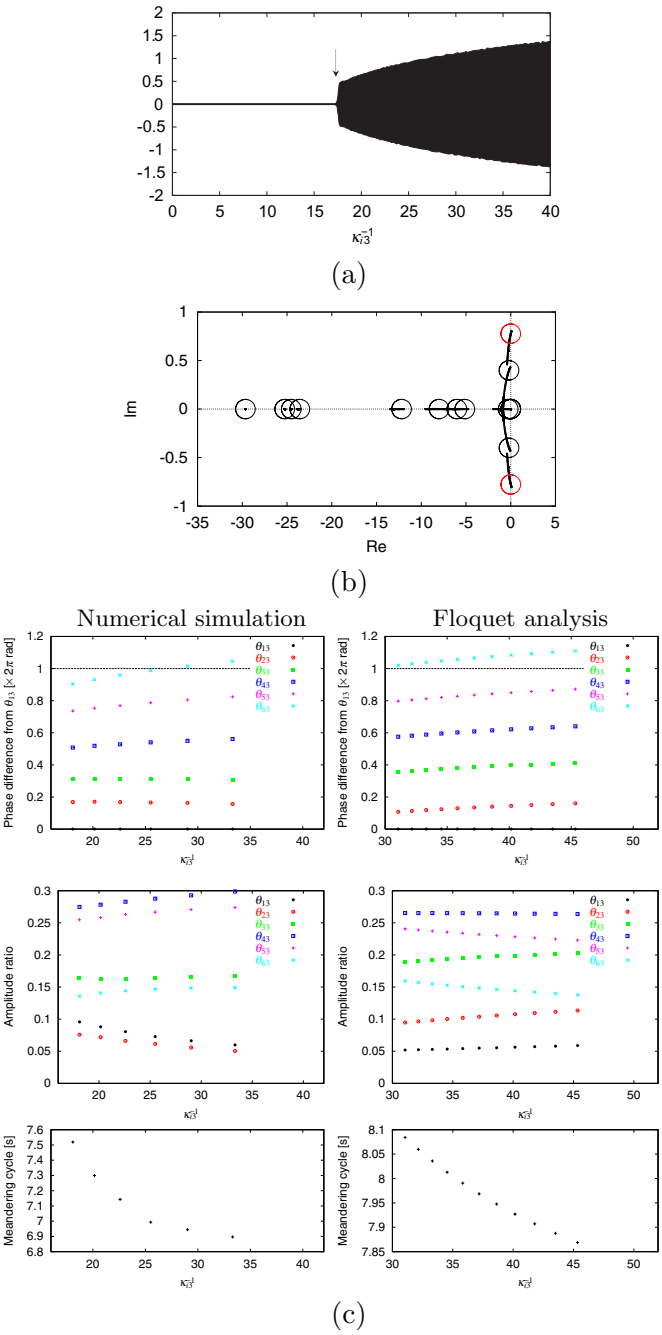


Figure 11. *Numerical and analytical results using an in-phase gait pattern ($\Delta\phi_{\mathrm{cntrl}} = 0$).* (a) *Yaw angle of Coupler* 3, θ_{43}, *at* $\phi_1^1 = 0$. (b) *Trajectories of Floquet exponents.* (c) *Comparison between numerical and analytical results.*

phase difference with respect to angle θ_{13}, and the amplitude ratio between the angles versus the reciprocal of gain parameter κ_{i3}^{-1}. These numerical results reveal that wavy motion arises beyond a bifurcation point and that the transition from a straight to a meandering walk is

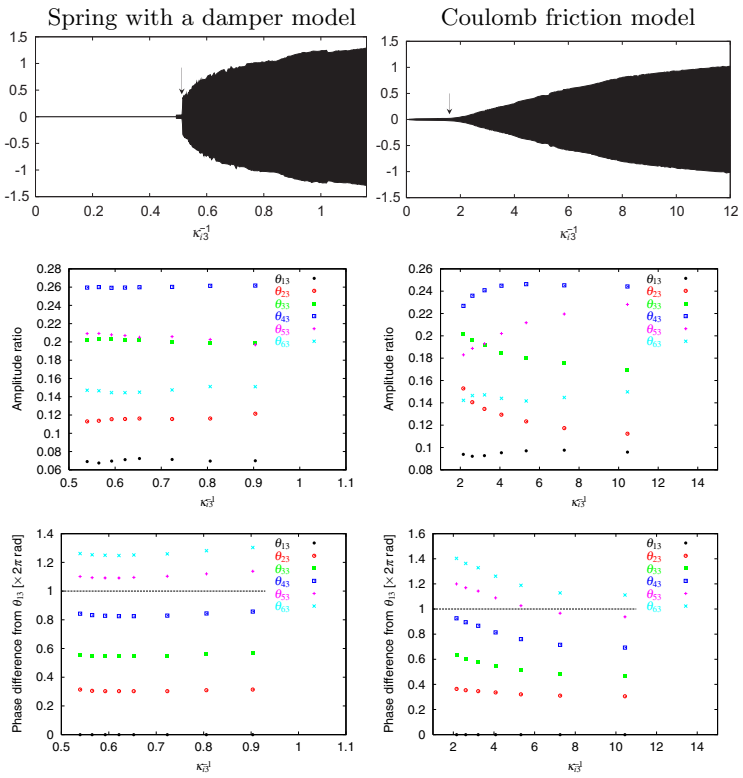


Figure 12. *Numerical results obtained using a spring with a damper and a Coulomb friction model.*

not specific to the ground reaction force model.

These results obtained from numerical simulations and Floquet analysis clarified the dynamic characteristics inherent in the multilegged modular robot. In the next section, we focus on the dynamic characteristics to achieve one solution to solve the problem in multilegged robots.

4. Turning maneuvers utilizing dynamic properties. Dynamic characteristics such as stability must greatly affect the maneuverability and agility of locomotion. Schmitt et al. [61, 62, 63, 64] simply modeled the hexapod walking of a cockroach, which has marvelous agility, and analytically demonstrated that it successfully achieves a quick turn by virtue of destabilizing its straight walk through changing its dynamic features.

As described above, there is a characteristic problem for a locomotion robot that has many legs: although many contact legs help the robot to achieve a stable walk, all those contact legs also keep the robot from accomplishing maneuverable and agile motions such as a quick turn. It would be interesting and helpful if we attained one solution to the problem by using the dynamic properties inherent in the robot as cockroaches use the dynamic features embedded in themselves. Therefore, in this section, we especially focus on the dynamic characteristics analyzed in the previous sections.

In particular, we investigate the relationship between the dynamic features and maneuverability of the robot. To clarify this relationship, we have the robot pursue a target moving

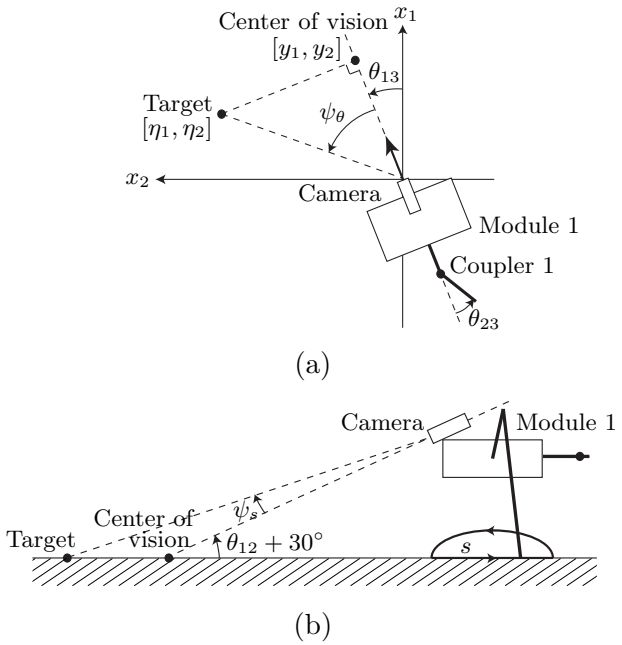


Figure 13. *Target pursuit.* (a) *Direction* ψ_θ. (b) *Distant angle* ψ_s.

across the floor. A camera is attached to the head of Module 1, whose elevation angle is $-30°$. From the visual image taken by the camera, the robot can obtain direction angle ψ_θ and distance angle ψ_s (see Figures 13(a) and (b)). The center of vision is the intersection point between the ground and the visual line of the camera, whose position is expressed by $[\,\eta_1\ \eta_2\,]$ on the floor. The position of the target is $[\,y_1\ y_2\,]$. The sampling frequency of the visual information is set to be 20 Hz. By using information ψ_θ and ψ_s, the robot attempts to follow the moving target. In this section, we use six modules ($n = 6$), an anti-phase gait pattern ($\Delta\phi_{\text{cntrl}} = \pi$), and the viscous damper model for the horizontal reaction force.

To manipulate the walking direction, input torque u_{13} at the yaw joint of Coupler 1 is activated by incorporating the desired angle regarding visual information ψ_θ, given by

$$u_{23} = -\kappa_{23}(\theta_{23} + \psi_\theta) - \sigma_{23}\dot{\theta}_{23}, \tag{4.1}$$

where feedback gains κ_{23} and σ_{23} are fixed and parameter f in (3.1) is set to 1.0 for them. The aim of this control is to point the first module in the direction of the target and then make the other modules follow the first module through their passive connections. To approach the target, by using visual information ψ_s, stride s is determined by

$$s = \kappa_s \psi_s, \tag{4.2}$$

where κ_s is set to 0.191 m/rad and stride s is limited with a saturation at ± 5 cm.

In the task of pursuing the target, the target moves straight at a constant speed of 0.18 m/s. The target then makes a right-angled turn to the left and continues to move straight at the same constant speed. First, the robot walks straight following the target, where feedback

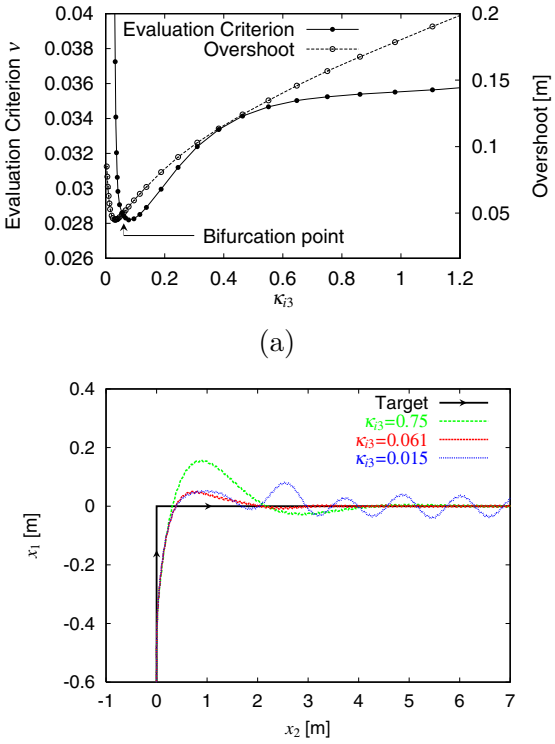


Figure 14. *Simulation results of the target pursuit.* (a) *Evaluation criterion* ν *and overshoot of the trajectory of Module* 1. (c) *Trajectories of the target and Module* 1.

gains κ_{i3} and σ_{i3} of the couplers' yaw joints ($i = 3, \ldots, 6$) are also fixed by setting $f = 1.0$, and wavy motion does not appear. Immediately after the target makes a turn, we change feedback gains κ_{i3} and σ_{i3} except for Coupler 1 ($i = 3, \ldots, 6$) and investigate the relationship between the maneuverability and dynamic characteristics dependent on the feedback gains. To examine the agility of locomotion, we employ evaluation criterion ν, which represents the mean square error between the target and center of vision, given by

$$\nu = \frac{1}{\tau_{\text{task}}} \int_0^{\tau_{\text{task}}} \sqrt{(\eta_1 - y_1)^2 + (\eta_2 - y_2)^2}\, dt, \tag{4.3}$$

where τ_{task} is the time interval to execute this task (set to 50 s).

14(a)

Figure 14(a) shows evaluation criterion ν and the overshoot of the trajectory of Module 1 with respect to gain parameter κ_{i3}. Figure 14(b) shows the trajectories of the target and Module 1 during the target pursuit, especially with respect to $\kappa_{i3} = 0.75$, 0.061 (close to the bifurcation point), and 0.015. Figure 15 is a snapshot of the turning behavior during the target pursuit ($\kappa_{i3} = 0.061$). A video of this turning behavior and the comparison between three cases is available as 66475_02.avi [2.1MB]. These figures show that the robot achieves high maneuverability and agility by using the feedback gains around the bifurcation point

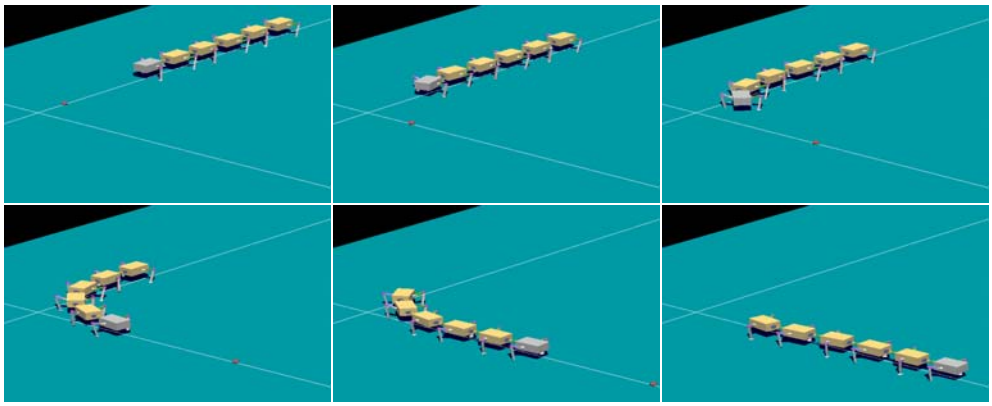


Figure 15. *Turning behavior during the target pursuit.*

where a straight walk becomes unstable. Although they do not necessarily mean that the bifurcation point is optimal, they do illustrate that there is an adequate region for feedback gains around the bifurcation point to achieve high maneuverability and agility. When the feedback gains are larger than the bifurcation point (i.e., when a straight walk is stable), evaluation criterion ν is also larger. This reflects that the robot is unable to obtain sufficient maneuverability, resulting in a large overshoot of the trajectory of Module 1, preventing the modules behind Module 1 from smoothly following Module 1. When the feedback gains are smaller than the bifurcation point (i.e., when a straight walk is unstable), evaluation criterion ν is also larger, partly because the undulatory motion is excited during target pursuit. Note that when the feedback gains are excessively small, the turning behavior is easily disintegrated. Therefore, the robot needs to use appropriate feedback gains. These results imply that the decrease in stability during a straight walk, due to a reduction in the strength of the connection between the modules, helps the robot to efficiently accomplish this task. In other words, the robot appears to achieve its maneuverability and agility by virtue of changes in the dynamic characteristics.

5. Discussion.

The contribution of this paper consists mainly of two parts. One is the elucidation of the dynamic characteristics and the mechanism of them embedded in a multilegged modular robot. The other is an investigation into improving the maneuverability and mobility of the robot by using the dynamic properties.

5.1. Dynamic characteristics inherent in the robot.

In this paper, we revealed that a straight walk by a multilegged modular robot that has more than two modules, each with one body and two legs, naturally turns into a meandering walk depending on the mechanical characteristics. This transition reflects the fact that undulatory motion appears according to the strength of the connection between the modules. Specifically, a straight walk becomes unstable and wavy motion is excited through a Hopf bifurcation. Furthermore, this dynamic property is qualitatively independent of the robot models such as the number of modules, gait pattern, and ground reaction force model.

The robot receives external forces from the ground when the tips of the legs are in contact with it. Since the robot manipulates the legs parallel to the body of the module as designed

in section 2.2, it obtains propulsive forces in that direction and moves forward. However, motions in the lateral direction appear and undulatory motions are generated due to changes in the mechanical characteristics. We clarified this mechanism by linearizing the equations of motion of the robot around the state of a straight walk under certain assumptions and by investigating the stability of the straight walk based on the Floquet theory.

In this Floquet analysis, the following should be noted. It is true that this analysis verifies that a straight walk is destabilized and turns into another walk, but it does not necessarily indicate formation of undulatory motion. Although this analysis surely suggests that when a straight walk becomes unstable the states diverge in accordance with the destabilized eigenvector and generate wavy motion, the nonlinearity in the dynamics is essential to saturate the divergence of the states and to complete a meandering motion.

Since this modular robot consists of many modules, all of which move independently and interact with each other through the couplers, we can consider this robot an autonomous, decentralized system. This system generates its motion pattern through the interaction between the dynamics of the mechanical system, the environment, and module connections. Furthermore, this system naturally changes the motion pattern according to the interaction between the modules. Although we often observe such a system that creates different motion patterns by changing control inputs depending on circumstances, not much work has been done on such a system that generates motion patterns by changing the mechanical and dynamic features and through a bifurcation like our results.

Golubitsky et al. [28, 29] analytically investigated the CPG network of many-legged animals such as Myriapods that display meandering walks, suggesting that there is an important type of neural system structure. On the other hand, we generate the meandering walk of a multilegged robot through the mechanical dynamics and the environment without incorporating the dynamical neural system. Animals must generate their walk by making the most of their inherent physical characteristics and neural structure. It is both mathematically and biologically important to elucidate their roles in locomotion dynamics.

5.2. Maneuverability dependent on the dynamic characteristics. The key issue in controlling a locomotion robot is to improve the stability and maneuverability of the robot's motion, where stability means that the robot continues locomotion without falling over. However, stability and maneuverability generally conflict with each other. Therefore, it is important to determine how to reach an acceptable trade-off between them in designing the controller. Regarding a multilegged robot, although an increase in the number of legs results in an improvement of stability, it makes the robot less maneuverable since the legs are constrained on the ground. Thus, it is essential to design a controller that enhances maneuverability and agility.

In designing the controller for a locomotion robot that has many legs and easily accomplishes statically stable walking, the model-based approach is widely used [12, 27, 53], in which the foot landing positions are calculated and then the joint angles are computed based on the inverse kinematics. However, the controller needs the calculations to be performed in real time to achieve desired motions such as target pursuit. It also requires more complicated and heavy computations as the numbers of modules and legs increase. Furthermore, this approach makes it difficult to attain adaptive and robust motions against environmental changes, errors in modeling, and disturbances.

To design the controller of a locomotion robot, it is essential to first undertake detailed research to better comprehend the intrinsic mechanical and dynamic characteristics of the robot. After that, it is important to design a simple controller to achieve the desired performance by making the most of the inherent characteristics. For example, the dynamic features of passive dynamic walkers have been examined in detail, revealing that they have a self-stabilizing property, a bifurcation characteristic, and chaotic behavior [24, 31, 49]. In particular, by using the intrinsic self-stabilizing property, many controllers have been successfully developed [10, 17, 30, 46, 65, 67, 73, 79].

This paper has so far clarified the inherent dynamic characteristics of a multilegged modular robot and revealed that a straight walk by the robot is destabilized due to its mechanical features and turns into a meandering walk. Such destabilization of a straight walk indicates that it becomes difficult for the robot to maintain its straight walk, and it does help the robot to increase the maneuverability. In section 4, we proposed simple turning maneuvers for the robot and investigated whether the robot achieves maneuverable and agile motions. The robot moves its legs periodically and manipulates the heading module according to the vision information. Despite having such a simple controller and strategy, the robot accomplishes efficient turning and walking by virtue of the dynamic characteristics. Making it harder for the robot to walk straight may be a very straightforward and acceptable approach to improving mobility. In this paper, the significance of the investigation of the dynamics properties embedded in the robot is to demonstrate that the robot can manipulate the ease and difficulty of a straight walk, and the crucial conclusion lies in the connection of the intrinsic dynamic characteristics with functions in locomotion, such as maneuverability.

5.3. Advanced controller and turning maneuvers. In the multilegged modular robot, the modules are connected with each other through the couplers. As the connections are weakened, the stability of its straight walk decreases and the mobility of each module increases. However, when the connections are excessively weakened, the walking motion will easily disintegrate. Therefore, a locomotion controller is required for determining adequate connection strength. In contrast to locomotion robots, animals modulate their muscle tone depending on circumstances and achieve adaptive walking. Although in this paper we examine the turning of the robot without changing the strength of the connections while performing an experiment, it is important to design a controller that modulates the strength according to the purpose and situation. Our results must provide guiding principles for designing such a controller and strategy.

As demonstrated in section 3.3.1, quadruped robots that consist of two modules and four legs, as well as biped robots comprising one module and two legs, do not possess dynamic characteristics in which a straight walk is destabilized and turns into a meandering walk. To be sure, the stability of a straight walk by a quadruped robot will deteriorate by decreasing the strength of the coupler's connection, but it is not sufficient to merely destabilize the straight walk; thus it is difficult to design a controller that makes the most of the intrinsic dynamic characteristics by using the maneuvers described in this paper.

Generally, in slow walking by a quadruped robot, the robot's motion is generated so that the center of gravity lies within the support leg polygon to maintain stable walking, making it relatively easy for a quadruped robot to walk slowly. However, since it is difficult to attain high-speed, dynamic, and agile walking, a controller is needed that stabilizes the

robot while retaining high mobility. Tsujita, Toui, and Tsuchiya [76] proposed a turning controller using nonlinear oscillators for a quadruped robot and accomplished dynamic turning by incorporating the kinematic control design and modulating the walking motion through sensory information and oscillators.

In this paper, we have analyzed the walking motion of the robot on flat terrain. However, terrain in the real world is not necessarily even and is often irregular and rough. One of the advantages of legged robots is that they are more adaptive and robust against such terrains than wheeled robots. Although irregular terrain produces up-and-down, roll, and pitch motions in a robot, the multilegged modular robot considered in this paper may deal with them by using the couplers' roll and pitch joints. In that case, the joint compliance and the strength of the connections will play important roles similar to the couplers' yaw joints.

6. Conclusion. In this paper, we dealt with locomotion of a multilegged modular robot. In particular, we showed that a straight walk by this robot naturally changes into a meandering walk by changes in the compliance of the yaw joints between the modules without incorporation of any oscillatory inputs. Based on the Floquet theory, we first clarified that this transition reflects that a straight walk is destabilized and undulatory motion is excited due to a Hopf bifurcation. Then, we investigated the role of this dynamic property in achieving maneuverability and agility, conducting an experiment in which the robot pursued a target moving across the floor. By using the proposed simple controller, the robot accomplished the task and performed high maneuverability and agile motions by making the most of the dynamic characteristics inherent in it.

Videos of these numerical simulations are available at http://control.kuaero.kyoto-u.ac.jp/member/aoi/.

Appendix A. Ground reaction force model. When the tip of the leg is in contact with the ground, the robot receives a reaction force from the ground. Position vector z of the leg tip and ground reaction force γ are expressed on the ground by $z = [\, z_1 \ z_2 \ z_3 \,]^{\mathrm{T}}$ and $\gamma = [\, \gamma_1 \ \gamma_2 \ \gamma_3 \,]^{\mathrm{T}}$, respectively. Ground reaction force γ_3 in the vertical direction is modeled using a spring with a damper [1, 5, 77], given by $\gamma_3 = -\kappa_{\mathrm{vrt}} z_3 - \sigma_{\mathrm{vrt}} \dot{z}_3$ during contact ($z_3 < 0$); otherwise, $\gamma_3 = 0$, where κ_{vrt} and σ_{vrt} are constants.

A multilegged modular robot generates its steady walking motion essentially in a horizontal plane by virtue of the support provided by its many legs. Therefore, ground reaction forces in the horizontal direction are important in generating such a horizontal motion. Throughout this paper, we employ a viscous damper model for the horizontal reaction force, expressed by $\gamma_i = -\sigma_{\mathrm{hrz}} \dot{z}_i$ when $\gamma_3 > 0$; otherwise, $\gamma_i = 0$ $(i = 1, 2)$, where σ_{hrz} is a constant.

In addition to the viscous damper model, we use another two models to investigate whether the dynamic characteristics of the robot considered in this paper are specific to such a ground reaction force model (see section 3.3.3). One is a spring with a damper model similar to the vertical direction [1, 5], expressed by $\gamma_i = -\kappa_{\mathrm{hrz}}(z_i - \overline{z}_i) - \sigma_{\mathrm{hrz}} \dot{z}_i$ during contact $(i = 1, 2)$, where κ_{hrz} is a constant and $\overline{z}_i$ is the position where the leg tip is constrained. The other is a Coulomb friction model, where the robot receives horizontal force γ_{v_z} in the direction of the derivative of z, given by $\gamma_{v_z} = -\mu\gamma_3$ during contact, where μ is the coefficient of friction. This force is approximated by using a pseudo-Coulomb friction force [77], given by $\gamma_{v_z} = -2\mu\gamma_3 \arctan(\sigma_{\mathrm{hrz}} \pi v_z / (2\mu\gamma_3))/\pi$ during contact, where $v_z = \sqrt{\dot{z}_1^2 + \dot{z}_2^2}$. When

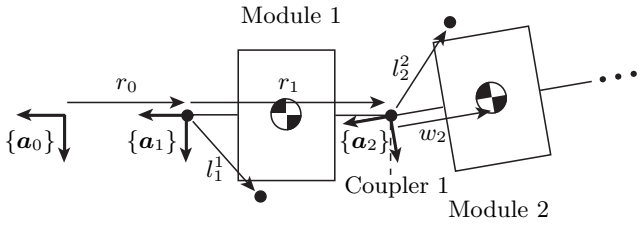


Figure 16. *Coordinate axes and vectors of the simplified model.*

$v_z \gg \mu\gamma_3/\sigma_{\mathrm{hrz}}$, γ_{v_z} approaches $-\mu\gamma_3$, which corresponds to a Coulomb friction model. On the other hand, when the tip of the leg slips slightly on the ground and speed v_z is slow, γ_{v_z} comes close to $-\sigma_{\mathrm{hrz}}v_z$, which corresponds to a viscous damper model. Since we consider the walking motion of the robot where the tips of the legs hardly slip with respect to the ground and speed v_z is slow compared to the locomotion speed, we employ the viscous damper model throughout this paper.

Appendix B. Equation of motion of the simplified model. In this appendix, we show the equation of motion for the simplified model of the multilegged modular robot described in section 3.2. First, we introduce the following coordinate axes and vectors (see Figure 16). Coordinate axes $\{\boldsymbol{a}_0\} = \{\boldsymbol{a}_{01}\ \boldsymbol{a}_{02}\}$ are fixed to the ground, where axes $\boldsymbol{a}_{01}$ and $\boldsymbol{a}_{02}$ are in the nominal walking direction and the lateral direction, respectively. Coordinate axes $\{\boldsymbol{a}_i\} = \{\boldsymbol{a}_{i1}\ \boldsymbol{a}_{i2}\}$ $(i = 1, \ldots, n)$ are fixed in the body of Module i, whose origin is located at Coupler $(i-1)$. Note that the origin of axes $\{\boldsymbol{a}_1\}$ is at the same point on Module 1 as the coordinate axes fixed to the other modules. Vector r_i $(i = 0, \ldots, n-1)$ is the distance vector from the origin of axes $\{\boldsymbol{a}_i\}$ to the origin of axes $\{\boldsymbol{a}_{i+1}\}$, expressed in axes $\{\boldsymbol{a}_i\}$, where $r_0 = [\,x_1\ x_2\,]^{\mathrm{T}}$. Vector w_i $(i = 1, \ldots, n)$ is the distance vector from the origin of axes $\{\boldsymbol{a}_i\}$ to the center of mass of Module i, expressed in axes $\{\boldsymbol{a}_i\}$. Vector l_i^j $(i = 1, \ldots, n,\ j = 1, 2)$ is the distance vector from the origin of axes $\{\boldsymbol{a}_i\}$ to the tip of Leg j of Module i, expressed in axes $\{\boldsymbol{a}_i\}$. Note that as assumed in section 3.2, the tips of the legs move along the trajectories kinematically designed in axes $\{\boldsymbol{a}_i\}$ and that vector l_i^j becomes a periodic function of time t, i.e., $l_i^j = l_i^j(t)$, whose cycle is equal to step cycle τ.

Using Lagrangian equations, the equation of motion for state variable $q^{\mathrm{T}} = [\,x_1\ x_2\ \theta_{13}\ \cdots\ \theta_{n3}\,] \in \mathbb{R}^{n+2}$ is derived by

$$K(q)\ddot{q} + h(q, \dot{q}) = u(q, \dot{q}) + \lambda(q, \dot{q}, t), \tag{B.1}$$

where $K(q) \in \mathbb{R}^{(n+2)\times(n+2)}$ is the inertia matrix, $h(q, \dot{q}) \in \mathbb{R}^{n+2}$ is the nonlinear term, $u(q, \dot{q}) \in \mathbb{R}^{n+2}$ is the input torque term, and $\lambda(q, \dot{q}, t) \in \mathbb{R}^{n+2}$ is the ground reaction force. Note that since the tips of the legs move periodically with respect to the body of each module, ground reaction force λ becomes a function of time t. Specifically, inertia matrix $K(q)$ is given by

$$K(q) = H^{\mathrm{T}}\{L(q)^{\mathrm{T}} M L(q) + J\}H + J_{\mathrm{g}}, \tag{B.2}$$

where

$$H = \left[\begin{array}{c|ccc} I_2 & & & \\ \hline & 1 & & \\ & \vdots & \ddots & \\ & 1 & \cdots & 1 \end{array}\right] \in \mathbb{R}^{(n+2)\times(n+2)},$$

$$L(q) = L_1(q) + L_2 \in \mathbb{R}^{2n\times(n+2)},$$

$$L_1(q) = \left[\begin{array}{c|cccc} R_{1,0}(q) & o_2 & & & \\ R_{2,0}(q) & R_{2,1}(q)\tilde{r}_1 & o_2 & & \\ \vdots & \vdots & \ddots & \ddots & \\ R_{n,0}(q) & R_{n,1}(q)\tilde{r}_1 & \cdots & R_{n,n-1}(q)\tilde{r}_{n-1} & o_2 \end{array}\right], \quad L_2 = \left[\begin{array}{c|ccc} \mathcal{O} & \tilde{w}_1 & & \\ & & \ddots & \\ & & & \tilde{w}_n \end{array}\right],$$

$$M = \mathrm{diag}[\, m_1\ m_1\ m_2\ m_2\ \cdots\ m_n\ m_n\,] \in \mathbb{R}^{2n\times 2n},$$

$$J = \mathrm{diag}[\, 0\ 0\ j_1\ j_2\ \cdots\ j_n\,] \in \mathbb{R}^{(n+2)\times(n+2)},$$

$$J_{\mathrm{g}} = \mathrm{diag}[\, 0\ 0\ 0\ n_{\mathrm{g}}^2 j_{\mathrm{g}}\ \cdots\ n_{\mathrm{g}}^2 j_{\mathrm{g}}\,] \in \mathbb{R}^{(n+2)\times(n+2)};$$

m_i and j_i $(i = 1, \ldots, n)$ are, respectively, the mass and inertia of Module i, j_{g} and n_{g} are, respectively, the rotor inertia and gear ratio of a motor, $o_2 \in \mathbb{R}^2$ is a zero vector, $I_2 \in \mathbb{R}^{2\times 2}$ is a unit matrix, $\mathcal{O}$ is an appropriately sized zero matrix, $R_{i,j}(q) \in \mathbb{R}^{2\times 2}$ $(i, j = 1, \ldots, n,\ i > j)$ is the coordinate transform matrix of axes $\{\boldsymbol{a}_i\}$ with respect to axes $\{\boldsymbol{a}_j\}$ given by

$$R_{i,j}(q) = \begin{bmatrix} \cos\theta_{i,j}(q) & \sin\theta_{i,j}(q) \\ -\sin\theta_{i,j}(q) & \cos\theta_{i,j}(q) \end{bmatrix} \quad i, j = 1, \ldots, n,\ i > j, \tag{B.3}$$

where

$$\theta_{i,j}(q) = \sum_{k=j+1}^{i} \theta_{k3},$$

and for vector $b = [\, b_1\ b_2\,]^{\mathrm{T}}$, $\tilde{b}$ is expressed as $\tilde{b} = [\, -b_2\ b_1\,]^{\mathrm{T}}$.

Nonlinear term $h(q, \dot{q})$ becomes

$$h(q, \dot{q}) = K_t(q, \dot{q})\dot{q} + V(q, \dot{q})p(q, \dot{q}), \tag{B.4}$$

where

$$K_t(q, \dot{q}) = \dot{K}(q),$$

$$V(q, \dot{q}) = \left[\begin{array}{c|ccc} & \mathcal{O} & & \\ \hline \tilde{v}_1^{\mathrm{T}}(q, \dot{q}) & & & \\ & & \ddots & \\ & & & \tilde{v}_n^{\mathrm{T}}(q, \dot{q}) \end{array}\right] \in \mathbb{R}^{(n+2)\times 2n},$$

$$[\, v_1^{\mathrm{T}}(q, \dot{q})\ \cdots\ v_n^{\mathrm{T}}(q, \dot{q})\,]^{\mathrm{T}} = L_1(q)H\dot{q} \in \mathbb{R}^{2n},$$

$$p(q, \dot{q}) = B(q)^{\mathrm{T}} M L(q) H\dot{q} \in \mathbb{R}^{2n},$$

$$B(q) = \left[\begin{array}{cccc} I_2 & & & \\ R_{2,1}(q) & I_2 & & \\ \vdots & \ddots & \ddots & \\ R_{n,1}(q) & \cdots & R_{n,n-1}(q) & I_2 \end{array}\right] \in \mathbb{R}^{2n\times 2n}.$$

Input torque term $u(q,\dot{q})$ is expressed by

$$u(q,\dot{q}) = -Pq - D\dot{q}, \tag{B.5}$$

where

$$P = \mathrm{diag}[\,0\ 0\ 0\ \kappa_{23}\ \cdots\ \kappa_{n3}\,] \in \mathbb{R}^{(n+2)\times(n+2)},$$
$$D = \mathrm{diag}[\,0\ 0\ 0\ \sigma_{23}\ \cdots\ \sigma_{n3}\,] \in \mathbb{R}^{(n+2)\times(n+2)}.$$

Reaction force term $\lambda(q,\dot{q},t)$ becomes equivalent to

$$\lambda(q,\dot{q},t) = -H^{\mathrm{T}} L_{\mathrm{leg}}(q,t)^{\mathrm{T}} \Lambda_{\mathrm{leg}}(t)\{L_{\mathrm{leg}}(q,t)H\dot{q} + v_{\mathrm{leg}}(t)\}, \tag{B.6}$$

where

$$L_{\mathrm{leg}}(q,t) = \left[\begin{array}{c|cccc} R_{1,0}(q) & \tilde{l}_1^1(t) & & & \\ R_{1,0}(q) & \tilde{l}_1^2(t) & & & \\ R_{2,0}(q) & R_{2,1}(q)\tilde{r}_1 & \tilde{l}_2^1(t) & & \\ R_{2,0}(q) & R_{2,1}(q)\tilde{r}_1 & \tilde{l}_2^2(t) & & \\ \vdots & \vdots & \ddots & \ddots & \\ R_{n,0}(q) & R_{n,1}(q)\tilde{r}_1 & \cdots & R_{n,n-1}(q)\tilde{r}_{n-1} & \tilde{l}_n^1(t) \\ R_{n,0}(q) & R_{n,1}(q)\tilde{r}_1 & \cdots & R_{n,n-1}(q)\tilde{r}_{n-1} & \tilde{l}_n^2(t) \end{array}\right] \in \mathbb{R}^{4n\times(n+2)},$$

$$v_{\mathrm{leg}}(t) = \left[\, l_1^1(t)^{\mathrm{T}}\ l_1^2(t)^{\mathrm{T}}\ \cdots\ l_n^1(t)^{\mathrm{T}}\ l_n^2(t)^{\mathrm{T}} \,\right]^{\mathrm{T}} \in \mathbb{R}^{4n},$$

$$\Lambda_{\mathrm{leg}}(t) = \sigma_{\mathrm{hrz}}\,\mathrm{diag}[\,\varepsilon_1^1(t)\ \varepsilon_1^1(t)\ \varepsilon_1^2(t)\ \varepsilon_1^2(t)\ \cdots\ \varepsilon_n^1(t)\ \varepsilon_n^1(t)\ \varepsilon_n^2(t)\ \varepsilon_n^2(t)\,] \in \mathbb{R}^{4n\times 4n},$$

$$\varepsilon_i^j(t) = \begin{cases} 1 & \text{if Leg } j \text{ of Module } i \text{ is in the stance phase,} \\ 0 & \text{otherwise,} \end{cases}$$
$$i = 1,\ldots,n,\ \ j = 1,2,$$

and σ_{hrz} is the damping coefficient of the ground. For example, in the anti-phase gait pattern ($\Delta\phi_{\mathrm{cntrl}} = \pi$), Leg j of Module i ($(i,j) = (1,1),(2,2),(3,1),(4,2),\ldots$) is in the stance phase from $t = 0$ to $t = \tau/2$ and vector $l_i^j(t)$ is given by

$$l_i^j(t) = \left[\, a_i^j - vt\ b_i^j \,\right]^{\mathrm{T}}, \quad 0 \le t \le \tau/2, \tag{B.7}$$

where $v = s/(\beta\tau)$, and a_i^j and b_i^j ($i = 1,\ldots,n$, $j = 1,2$) are the components of the distance vector from the origin of axes $\{\boldsymbol{a}_i\}$ to point AEP of the trajectory for Leg j of Module i. On the other hand, Leg j of Module i ($(i,j) = (1,2),(2,1),(3,2),(4,1),\ldots$) is in turn in the stance phase from $t = \tau/2$ to $t = \tau$, and vector $l_i^j(t)$ is expressed by

$$l_i^j(t) = \left[\, a_i^j - v(t-\tau/2)\ b_i^j \,\right]^{\mathrm{T}}, \quad \tau/2 \le t \le \tau. \tag{B.8}$$

In contrast to the anti-phase gait pattern, in the in-phase gait pattern ($\Delta\phi_{\mathrm{cntrl}} = 0$), Leg j of Module i ($(i,j) = (1,1),(1,2),(3,1),(3,2),\ldots$) is in the stance phase from $t = 0$ to $t = \tau/2$

and Leg j of Module i ($(i,j) = (2,1),(2,2),(4,1),(4,2),\dots$) is in the stance phase from $t = \tau/2$ to $t = \tau$.

Appendix C. Linearization of the equation of motion of the simplified model. In this section, we show the linearized equation of the equation of motion (B.1) for the simplified model of the multilegged modular robot around the states of a straight walk. This is done to investigate the transition mechanism from a straight to a meandering walk according to the mechanical characteristics of the robot.

When we define state $\xi \in \mathbb{R}^{2n+4}$ as $\xi^{\mathrm{T}} = [\,\dot{q}^{\mathrm{T}}\, q^{\mathrm{T}}\,]$, state $\xi_{\mathrm{str}} \in \mathbb{R}^{2n+4}$ of a straight walk can be written as $\xi_{\mathrm{str}}^{\mathrm{T}} = [\,\dot{q}_{\mathrm{str}}^{\mathrm{T}}\ q_{\mathrm{str}}^{\mathrm{T}}\,] = [\,v\ 0\ \cdots\ 0\ vt + x_{10}\ 0\ \cdots\ 0\,]$, where x_{10} is the state of x_1 at $t = 0$. Perturbed state $\xi \in \mathbb{R}^{2n+4}$ from a straight walk is defined as $\xi^{\mathrm{T}} = \xi_{\mathrm{str}}^{\mathrm{T}} + \delta\xi^{\mathrm{T}}$, where $\delta\xi = [\,\delta\dot{q}^{\mathrm{T}}\ \delta q^{\mathrm{T}}\,]^{\mathrm{T}} \in \mathbb{R}^{2n+4}$ is the perturbation. By linearizing the equation of motion (B.1) around state ξ_{str}, we obtain

$$K(q_{\mathrm{str}})\delta\ddot{q} = -P\delta q - D\delta\dot{q} - \Gamma_{\mathrm{leg}}(t)\{L_{\mathrm{leg}}(q_{\mathrm{str}},t)H\delta\dot{q} + L_{\mathrm{leg}\,t}(q_{\mathrm{str}},\dot{q}_{\mathrm{str}},t)H\delta q\}, \tag{C.1}$$

where

$$\Gamma_{\mathrm{leg}}(t) = H^{\mathrm{T}} L_{\mathrm{leg}}(q_{\mathrm{str}},t)^{\mathrm{T}} \Lambda_{\mathrm{leg}}(t),$$
$$L_{\mathrm{leg}\,t}(q,\dot{q},t) = \dot{L}_{\mathrm{leg}}(q,t).$$

Equation (C.1) yields that

$$\delta\dot{\xi} = A(t)\delta\xi, \tag{C.2}$$

where

$$A(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ I_{n+2} & O_{n+2} \end{bmatrix} \in \mathbb{R}^{(2n+4)\times(2n+4)},$$
$$A_{11}(t) = -K(q_{\mathrm{str}})^{-1}\{D + \Gamma_{\mathrm{leg}}(t)L_{\mathrm{leg}}(q_{\mathrm{str}},t)H\},$$
$$A_{12}(t) = -K(q_{\mathrm{str}})^{-1}\{P + \Gamma_{\mathrm{leg}}(t)L_{\mathrm{leg}\,t}(q_{\mathrm{str}},\dot{q}_{\mathrm{str}},t)H\},$$

and $I_{n+2} \in \mathbb{R}^{(n+2)\times(n+2)}$ and $O_{n+2} \in \mathbb{R}^{(n+2)\times(n+2)}$ are unit and zero matrices, respectively. Since vector $l_i^j(t)$ is periodic and its cycle is equivalent to step cycle τ, and vector $\dot{l}_i^j(t)$ is constant from (B.7) and (B.8), we achieve $A(t+\tau) = A(t)$.

REFERENCES

[1] J. Adolfsson, H. Dankowicz, and A. Nordmark, *3D passive walkers: Finding periodic gaits in the presence of discontinuities*, Nonlinear Dynam., 24 (2001), pp. 205–229.

[2] R. Altendorfer, N. Moore, H. Komsuoglu, M. Buehler, H. B. Brown, Jr., D. McMordie, U. Saranli, R. Full, and D. E. Koditschek, *RHex: A biologically inspired hexapod runner*, Autonomous Robots, 11 (2001), pp. 207–213.

[3] R. Altendorfer, D. E. Koditschek, and P. Holmes, *Stability analysis of a clock-driven rigid-body SLIP model for RHex*, Int. J. Robotics Res., 23 (2004), pp. 1001–1012.

[4] R. Altendorfer, D. E. Koditschek, and P. Holmes, *Stability analysis of legged locomotion models by symmetry-factored return maps*, Int. J. Robotics Res., 23 (2004), pp. 979–999.

[5] S. Aoi and K. Tsuchiya, *Locomotion control of a biped robot using nonlinear oscillators*, Autonomous Robots, 19 (2005), pp. 219–232.
[6] S. Aoi and K. Tsuchiya, *Adaptive behavior in turning of an oscillator-driven biped robot*, Autonomous Robots, to appear.
[7] S. Aoi and K. Tsuchiya, *Self-stability of a simple walking model driven by a rhythmic signal*, Nonlinear Dynam., 48 (2007), pp. 1–16.
[8] S. Aoi and K. Tsuchiya, *Stability analysis of a simple walking model driven by an oscillator with a phase reset using sensory feedback*, IEEE Trans. Robotics, 22 (2006), pp. 391–397.
[9] S. Aoi and K. Tsuchiya, *Bifurcation and chaos of a simple walking model driven by a rhythmic signal*, Internat. J. Non-Linear Mech., 41 (2006), pp. 438–446.
[10] F. Asano, Z. W. Luo, and M. Yamakita, *Biped gait generation and control based on a unified property of passive dynamic walking*, IEEE Trans. Robotics, 21 (2005), pp. 754–762.
[11] N. A. Bernstein, *The Coordination and Regulation of Movements*, Pergamon Press, Oxford, New York, 1967.
[12] J.-D. Boissonnat, O. Devillers, and S. Lazard, *Motion planning of legged robots*, SIAM J. Comput., 30 (2000), pp. 218–246.
[13] A. Castano, W.-M. Shen, and P. Will, *CONRO: Towards deployable robots with inter-robot metamorphic robots*, Autonomous Robots, 8 (2000), pp. 309–324.
[14] J. G. Cham, S. A. Bailey, J. E. Clark, R. J. Full, and M. R. Cutkosky, *Fast and robust: Hexapedal robots via shape deposition manufacturing*, Int. J. Robotics Res., 21 (2002), pp. 869–882.
[15] J. G. Cham, J. K. Karpick, and M. R. Cutkosky, *Stride period adaptation of a biomimetic running hexapod*, Int. J. Robotics Res., 23 (2004), pp. 141–153.
[16] R. Chatterjee, M. Nagai, and F. Matsuno, *Development of modular legged robots: Study with three-legged robot modularity*, in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan, 2004, pp. 1450–1455.
[17] S. H. Collins, A. L. Ruina, R. Tedrake, and M. Wisse, *Efficient bipedal robots based on passive-dynamic walkers*, Science, 307 (2005), pp. 1082–1085.
[18] A. Crespi, A. Badertscher, A. Guignard, and A. J. Ijspeert, *AmphiBot I: An amphibious snake-like robot*, Robotics and Autonomous Systems, 50 (2005), pp. 163–175.
[19] K. S. Espenschied, R. D. Quinn, R. D. Beer, and H. J. Chiel, *Biologically based distributed control and local reflexes improve rough terrain locomotion in a hexapod robot*, Robotics and Autonomous Systems, 18 (1996), pp. 59–64.
[20] M. Farkas, *Periodic Motions*, Springer-Verlag, Berlin, 1994.
[21] T. Fukuda and Y. Kawauchi, *Cellular robotic system (CEBOT) as one of the realizations of self-organizing intelligent universal manipulator*, in Proceedings of the IEEE International Conference on Robotics and Automation, IEEE Robotics and Automation Society, New York, 1990, pp. 662–667.
[22] Y. Fukuoka, H. Kimura, and A. Cohen, *Adaptive dynamic walking of a quadruped robot on irregular terrain based on biological concepts*, Int. J. Robotics Res., 22 (2003), pp. 187–202.
[23] R. J. Full, T. Kubow, J. Schmitt, P. Holmes, and D. Koditschek, *Quantifying dynamic stability and maneuverability in legged locomotion*, Integ. Comp. Biol., 42 (2002), pp. 149–157.
[24] M. Garcia, A. Chatterjee, A. Ruina, and M. Coleman, *The simplest walking model: Stability, complexity, and scaling*, ASME J. Biomech. Eng., 120 (1998), pp. 281–288.
[25] H. Geyer, A. Seyfarth, and R. Blickhan, *Spring-mass running: Simple approximate solution and application to gait stability*, J. Theoret. Biol., 232 (2005), pp. 315–328.
[26] R. Ghigliazza, R. Altendorfer, P. Holmes, and D. Koditschek, *A simply stabilized running model*, SIAM J. Appl. Dyn. Syst., 2 (2003), pp. 187–218.
[27] Y. Go, X. Yin, and A. Bowling, *Navigability of multi-legged robots*, IEEE/ASME Trans. Mechatronics, 11 (2006), pp. 1–8.
[28] M. Golubitsky, I. Stewart, P. L. Buono, and J. J. Collins, *A modular network for legged locomotion*, Phys. D, 115 (1998), pp. 56–72.
[29] M. Golubitsky, I. Stewart, P. L. Buono, and J. J. Collins, *Symmetry in locomotor central pattern generators and animal gaits*, Nature, 401 (1999), pp. 693–695.
[30] A. Goswami, B. Espiau, and A. Keramane, *Limit cycles in a passive compass gait biped and passivity-mimicking control laws*, Autonomous Robots, 4 (1997), pp. 273–286.

[31] A. Goswami, B. Thuilot, and B. Espiau, *A study of the passive gait of a compass-like biped robot: Symmetry and chaos*, Int. J. Robotics Res., 17 (1998), pp. 1282–1301.
[32] S. Grillner, *Control of locomotion in bipeds, tetrapods and fish*, in Handbook of Physiology, American Physiological Society, Bethesda, MD, 1981, pp. 1179–1236.
[33] S. Grillner, *Neurobiological bases of rhythmic motor acts in vertebrates*, Science, 228 (1985), pp. 143–149.
[34] A. J. Ijspeert, A. Crespi, and J. M. Cabelguen, *Simulation and robotics studies of salamander locomotion. Applying neurobiological principles to the control of locomotion in robots*, Neuroinformatics, 3 (2005), pp. 171–196.
[35] A. J. Ijspeert, A. Crespi, D. Ryczko, and J. M. Cabelguen, *From swimming to walking with a salamander robot driven by a spinal cord model*, Science, 315 (2007), pp. 1416–1420.
[36] S. Inagaki, H. Yuasa, and T. Arai, *CPG model for autonomous decentralized multi-legged robot system—generation and transition of oscillation patterns and dynamics of oscillators*, Robotics and Autonomous Systems, 44 (2003), pp. 171–179.
[37] S. Inagaki, H. Yuasa, T. Suzuki, and T. Arai, *Wave CPG model for autonomous decentralized multi-legged robot: Gait generation and walking speed control*, Robotics and Autonomous Systems, 54 (2006), pp. 118–126.
[38] K. Inoue, S. Ma, and C. Jin, *Neural oscillator network-based controller for meandering locomotion of snake-like robots*, in Proceedings of the IEEE International Conference on Robotics and Automation (New Orleans, 2004), IEEE Robotics and Automation Society, New York, 2004, pp. 5064–5069.
[39] A. Ishiguro, K. Ishimaru, K. Hayakawa, and T. Kawakatsu, *How should control and body dynamics be coupled?—A robotic case study*, in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, NV, 2003, pp. 1727–1732.
[40] A. Ishiguro, M. Shimizu, and T. Kawakatsu, *A modular robot that exhibits amoebic locomotion*, Robotics and Autonomous Systems, 54 (2006), pp. 641–650.
[41] D. L. Jindrich and R. J. Full, *Many-legged maneuverability: Dynamics of turning in hexapods*, J. Exp. Biol., 202 (1999), pp. 1603–1623.
[42] D. L. Jindrich and R. J. Full, *Dynamic stabilization of rapid hexapedal locomotion*, J. Exp. Biol., 205 (2002), pp. 2803–2823.
[43] H. Kimura, S. Akiyama, and K. Sakurama, *Realization of dynamic walking and running of the quadruped using neural oscillator*, Autonomous Robots, 7 (1999), pp. 247–258.
[44] H. Kimura, Y. Fukuoka, and K. Konaga, *Adaptive dynamic walking of a quadruped robot by using neural system model*, Advanced Robotics, 15 (2001), pp. 859–876.
[45] T. M. Kubow and R. J. Full, *The role of the mechanical system in control: A hypothesis of self-stabilization in hexapedal runners*, Phil. Trans. Roy. Soc. London Ser. B, 354 (1999), pp. 849–861.
[46] A. D. Kuo, *Energetics of actively powered locomotion using the simplest walking model*, ASME J. Biomech. Eng., 124 (2002), pp. 113–120.
[47] W. H. Lee and A. C. Sanderson, *Dynamic analysis and distributed control of the tetrobot modular reconfigurable robotic system*, Autonomous Robots, 10 (2001), pp. 67–82.
[48] M. A. Lewis, R. Etienne-Cummings, M. J. Hartmann, Z. R. Xu, and A. H. Cohen, *An in silico central pattern generator: Silicon oscillator, coupling, entrainment, and physical computation*, Biol. Cybern., 88 (2003), pp. 137–151.
[49] T. McGeer, *Passive dynamic walking*, Int. J. Robotics Res., 9 (1990), pp. 62–82.
[50] S. Mori, *Integration of posture and locomotion in acute decerebrate cats and in awake, free moving cats*, Prog. Neurobiol., 28 (1987), pp. 161–196.
[51] J. Nakanishi, J. Morimoto, G. Endo, G. Cheng, S. Schaal, and M. Kawato, *Learning from demonstration and adaptation of biped locomotion*, Robotics and Autonomous Systems, 47 (2004), pp. 79–91.
[52] G. N. Orlovsky, T. Deliagina, and S. Grillner, *Neuronal Control of Locomotion: From Mollusc to Man*, Oxford University Press, Oxford, UK, 1999.
[53] D. K. Pratihar, K. Deb, and A. Ghosh, *Optimal turning gait of a six-legged robot using a GA-fuzzy approach*, Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 14 (2000), pp. 207–219.

[54] I. Poulakakis, J. A. Smith, and M. Buehler, *Modeling and experiments of untethered quadrupedal running with a bounding gait: The Scout* II *Robot*, Int. J. Robotics Res., 24 (2005) pp. 239–256.

[55] I. Poulakakis, J. A. Smith, and M. Buehler, *On the dynamics of bounding and extensions towards the half-bound and the gallop gaits*, in Proceedings of the International Symposium on Adaptive Motion of Animals and Machines, Kyoto, Japan, 2003, ThA-I-2.

[56] R. D. Quinn, G. M. Nelson, R. J. Bachmann, D. A. Kingsley, J. T. Offi, T. J. Allen, and R. E. Ritzmann, *Parallel complementary strategies for implementing biological principles into mobile robots*, Int. J. Robotics Res., 22 (2003), pp. 169–186.

[57] S. Rossignol, *Neural Control of Stereotypic Limb Movements*, Oxford University Press, Oxford, UK, 1996.

[58] D. Rus and M. Vona, *Crystalline robots: Self-reconfiguration with unit-compressible modules*, Autonomous Robots, 10 (2001), pp. 107–124.

[59] U. Saranli, M. Buehler, and D. E. Koditschek, *RHex: A simple and highly mobile hexapod robot*, Int. J. Robotics Res., 20 (2001), pp. 616–631.

[60] U. Saranli, A. A. Rizzi, and D. E. Koditschek, *Model-based dynamic self-righting maneuvers for a hexapedal robot*, Int. J. Robotics Res., 23 (2004), pp. 903–918.

[61] J. Schmitt and P. Holmes, *Mechanical models for insect locomotion: Dynamics and stability in the horizontal plane* I. *Theory*, Biol. Cybern., 83 (2000), pp. 501–515.

[62] J. Schmitt and P. Holmes, *Mechanical models for insect locomotion: Dynamics and stability in the horizontal plane* II. *Application*, Biol. Cybern., 83 (2000), pp. 517–527.

[63] J. Schmitt and P. Holmes, *Mechanical models for insect locomotion: Stability and parameter studies*, Phys. D, 156 (2001), pp. 139–168.

[64] J. Schmitt, M. Garcia, R. Razo, P. Holmes, and R. J. Full, *Dynamics and stability of legged locomotion in the horizontal plane: A test case using insects*, Biol. Cybern., 86 (2002), pp. 343–353.

[65] M. W. Spong and F. Bullo, *Controlled symmetries and passive walking*, IEEE Trans. Automat. Control, 50 (2005), pp. 1025–1031.

[66] P. Staritz, S. Skaff, C. Urmson, and W. L. Whittaker, *Skyworker: A robot for assembly, inspection and maintenance of large scale orbital facilities*, in Proceedings of the IEEE International Conference on Robotics and Automation (Seoul, 2001), IEEE Robotics and Automation Society, New York, 2001, pp. 4180–4185.

[67] Y. Sugimoto and K. Osuka, *Walking control of quasi-passive-dynamic-walking robot Quartet* III *based on delayed feedback control*, in Proceedings of the 5th International Conference on Climbing and Walking Robots, Paris, France, 2002, pp. 123–130.

[68] G. Taga, Y. Yamaguchi, and H. Shimizu, *Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment*, Biol. Cybern., 65 (1991), pp. 147–159.

[69] G. Taga, *A model of the neuro-musculo-skeletal system for human locomotion* I. *Emergence of basic gait*, Biol. Cybern., 73 (1995), pp. 97–111.

[70] G. Taga, *A model of the neuro-musculo-skeletal system for human locomotion* II. *Real-time adaptability under various constraints*, Biol. Cybern., 73 (1995), pp. 113–121.

[71] K. Takakusaki, T. Habaguchi, J. Ohtinata-Sugimoto, K. Saitoh, and T. Sakamoto, *Basal ganglia efferents to the brainstem centers controlling postural muscle tone and locomotion: A new concept for understanding motor disorders in basal ganglia dysfunction*, Neuroscience, 119 (2003), pp. 293–308.

[72] K. Takakusaki, K. Saitoh, H. Harada, and M. Kashiwayanagi, *Role of basal ganglia-brainstem pathways in the control of motor behaviors*, Neurosci. Res., 50 (2004), pp. 137–151.

[73] T. Takuma and K. Hosoda, *Controlling the walking period of a pneumatic muscle walker*, Int. J. Robotics Res., 25 (2006), pp. 861–866.

[74] K. Tsuchiya, S. Aoi, and K. Tsujita, *Locomotion control of a multi-legged locomotion robot using oscillators*, in Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Hammamet, Tunisia, 2002, TA2G2.

[75] K. Tsujita, K. Tsuchiya, and A. Onat, *Adaptive gait pattern control of a quadruped locomotion robot*, in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, Hawaii, 2001, pp. 2318–2325.

[76] K. Tsujita, H. Toui, and K. Tsuchiya, *Dynamic turning control of a quadruped locomotion robot using oscillators*, Advanced Robotics, 19 (2005), pp. 1115–1133.

[77] A. J. van den Bogert, H. C. Schamhadt, and A. Crowe, *Simulation of quadrupedal locomotion using a rigid body model*, J. Biomech., 22 (1989), pp. 33–41.

[78] H. Wagner and R. Blickhan, *Stabilizing function of skeletal muscles: An analytical investigation*, J. Theoret. Biol., 199 (1999), pp. 163–179.

[79] M. Wisse, A. L. Schwab, R. Q. van der Linde, and F. C. T. van der Helm, *How to keep from falling forward: Elementary swing leg action for passive dynamic walkers*, IEEE Trans. Robotics, 21 (2005), pp. 393–401.

[80] M. Yim, Y. Zhang, K. Roufas, D. Duff, and C. Eldershaw, *Connecting and disconnecting for chain selfreconfiguration with polybot*, IEEE/ASME Trans. Mechatronics, 7 (2002), pp. 442–451.

[81] M. Yim, K. Roufas, D. Duff, Y. Zhang, C. Eldershawand, and S. Homans, *Modular reconfigurable robots in space applications*, Autonomous Robots, 14 (2003), pp. 225–237.

[82] E. Yoshida, S. Murata, A. Kamimura, K. Tomita, H. Kurokawa, and S. Kokaji, *A self-reconfigurable modular robot: Reconfiguration planning and experiments*, Int. J. Robotics Res., 21 (2002), pp. 903–916.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 2, pp. 378–394

Bifurcation and Bistability in a Model of Hematopoietic Regulation*

Caroline Colijn† and Michael C. Mackey‡

Abstract. Stem cells and their relationship with mature tissues are of increasing interest in the biomedical sciences, but the dynamics of stem cell/tissue interactions are not well understood. We give a generic stem cell/tissue model and examine the dynamics of a specific case of this, namely, a four-compartment model of blood cell production and regulation. We apply the findings to cyclical neutropenia and periodic chronic myelogenous leukemia, two diseases of the blood production system. We track the position of the Hopf bifurcation believed to give rise to blood cell oscillations in these diseases. Results account for the variable success of granulocyte-colony stimulating factor, a common treatment for cyclical neutropenia, in reducing oscillations. The model displays bistability of periodic solutions, presenting the opportunity to stabilize the system through a temporary perturbation that induces switching between locally stable solutions. It is found that oscillations can be suppressed by properly timed pulses of increased amplification in the platelet line. The medical interpretation of this result is that temporary administration of thrombopoietin may suppress the oscillations. Though it is neutrophil oscillations that characterize both diseases clinically, and though it is probably a destabilization in the neutrophil line that initiates the oscillations, pulses in the neutrophil amplification do not effectively suppress oscillations.

Key words. mathematical biology, delay differential equations, bifurcation, coupled oscillations, stem cells

AMS subject classifications. 37G10, 34C14, 34C23, 37L15

DOI. 10.1137/050640072

1. Introduction.

Stem cells are of great interest in the biomedical sciences due to their ability to develop into all of the cell types in the body. It has been suggested that therapies using pluripotent stem cells could present cures for currently incurable degenerative diseases such as Alzheimer's and diabetes, among others. Recently, new sources for stem cells in amniotic fluid [7] and fibroblast culture [24] have emerged; these findings could alleviate some of the political debate about the ethics of stem cell research. The spatial and temporal dynamics of stem cell proliferation and differentiation are not well understood, which presents the opportunity for mathematicians to contribute to an exciting and rapidly growing field with many open problems.

Stem cells, by definition, are pluripotential, and many stem cells (adult stem cells) are self-renewing. In other words, they can differentiate and then mature into diverse tissue types, and their population is able to sustain itself through proliferation. Consider a generic situation as follows: a small population of stem cells is capable of differentiating and then,

*Received by the editors September 9, 2006; accepted for publication (in revised form) by J. Keener March 6, 2007; published electronically June 1, 2007. This work was supported by MITACS (Canada) and the Natural Sciences and Engineering Research Council of Canada.
http://www.siam.org/journals/siads/6-2/64007.html
†Department of Mathematics and Centre for Nonlinear Dynamics, McGill University, 3655 Promenade Sir William Osler, Montreal, Canada H3G 1Y6 (ccolijn@gmail.com).
‡Departments of Physiology, Physics, and Mathematics, and Centre for Nonlinear Dynamics, McGill University, 3655 Promenade Sir William Osler, Montreal, Canada H3G 1Y6 (mackey@cnd.mcgill.ca).

through a series of cell divisions, giving rise to N distinct tissue types. The rate at which stem cells differentiate into a given cell type is a function of the existing populations of cells. Furthermore, it takes some time, typically a number of days, for the complete maturation process.

A generic model along these lines is

$$
\begin{aligned}
&\dot{q} = -kqh(q) + \mu q_\tau h(t_\tau) - q\sum_{i=1}^{N} h_i(X_1,\ldots,X_n),\\
&\frac{\partial x_i}{\partial t} + \frac{\partial x_i}{\partial a} = f_i(x_1,\ldots,x_n),\\
&x_i(0,t) = a_i q_{\tau_i} h_i(X_{1_{\tau_i}},\ldots,X_{n_{\tau_i}}),\\
&x_i(a,0) = x_i^0(a),
\end{aligned}
\tag{1}
$$

where

$$X_i(t) = \int_0^\infty x_i(a,t)da.$$

Here, q represents quiescent stem cells, namely, those that are not undergoing mitosis (proliferation). They can enter a proliferative phase, at rate $kh(q)$; they then return a time τ later having been multiplied by some factor $\mu > k$. We use the notation $q_\tau \equiv q(t-\tau)$ for delays. They may also differentiate into tissue types $i = 1,\ldots,N$. In (1) the regulation of stem cell differentiation into each tissue type is a function of the total tissue numbers X_i. The quiescent stem cells q therefore have a loss term at rate $h_i(X_1,\ldots,X_N)$ for each tissue. The tissue populations themselves are given by $x_i(a,t)$, where a represents the time since maturation and t is time.

A salient feature of stem cell/tissue interactions is given in the first boundary condition of (1): the factors a_i represent "amplification" resulting from the many stages of cell division between the stem cell compartment and the fully mature tissues. The number of mature cells, X_i, is typically several orders of magnitude larger than the number of stem cells; i.e., the a_i are large. The tissue numbers are coupled to each other not only through their interactions (given by f_i) but by their delayed common origin in a small pool of stem cells.

The dynamics of models such as (1) are in general difficult to determine. Furthermore, the tissue population dynamics may depend not only on time and age, as given here, but on spatial properties as well. However, to the extent possible, it is desirable to understand to what extent the stem cell dynamics are independent of the mature tissues, the strength of effective coupling between tissues, the resilience of the system to increased loss of stem or tissue cells, and of course the steady state(s), periodic solutions, and bifurcation structure of the system.

Perhaps the simplest version of a model that includes stem cells as well as several fully differentiated tissues is the hematopoietic (blood-producing) system. Here, the tissue types are the circulating blood cells and the stem cells are the hematopoietic stem cells in the bone marrow. Since blood circulates through the body on a time scale of minutes, on time scales relevant to stem cell differentiation (which takes days) there are no spatial effects to speak of. Furthermore, the feedback loops by which blood cell numbers are regulated are comparatively

well known and simple, and we can model differentiation as $h_i(X_1, \ldots, X_n) = h_i(X_i)$ so that the rate of entry into each cell lineage is a function of mature numbers in that lineage only.

Though in this context it is relatively simple, the hematological system demonstrates interesting observed dynamics. In several hematopoietic diseases, blood cell numbers oscillate significantly, with the same period of oscillation occurring in the neutrophils (white blood cells), platelets, and sometimes reticulocytes (red blood cell precursors). In cyclical neutropenia (CN), neutrophil numbers reach dangerously low levels, oscillating with a period of 19–21 days in humans and 11–16 days in dogs. Longer periods of up to 50 days have been observed [15]. Platelet levels oscillate around their mean value with the same period. In periodic chronic myelogenous leukemia (PCML), leukocyte levels oscillate far above normal values with very long periods, ranging from 40–80 days [11].

Previous modeling efforts and traditional biological research have made progress in understanding the dynamics of these diseases, but the precise nature and origins of the oscillations remain disputed. This is in part because the dynamics of the hematopoietic stem cells have not been well characterized. These are located inside the bone marrow, and so are comparatively inaccessible. So, unlike the circulating blood cells, good time series data for the hematopoietic stem cells (HSCs) are unavailable.

Bernard, Bélair, and Mackey [1] presented a mathematical model that coupled the HSCs and circulating neutrophil population dynamics. Oscillations arose from a Hopf bifurcation in the HSC compartment, and were consistent with some, but not all, observed features of neutrophil oscillations in CN. Colijn and Mackey [5] presented a model of blood cell production that included the HSCs, neutrophils, platelets, and erythrocytes, and found parameters that were most important in fitting model simulations to data.

In section 2 we analyze a version of (1) analogous to that given in [5]. We perform bifurcation analysis (section 3) with respect to the parameters that are most critical in CN, and compare the hematopoietic stem cell compartment alone with the full model. In section 3.3, for a point in parameter space characteristic of treated CN, we find three locally stable solutions: two periodic branches and the steady state. This allows the exploration in section 4 of several methods to perturb the system from an oscillating branch to the steady-state branch.

2. The model.

A nondimensional model of the hematopoietic production system is given by

$$
\begin{aligned}
\frac{dq}{dt} &= -qb_1h_q(q) + b_1\mu_1q_1h_q(q_1) - q\left\{b_2h_n(n) + b_3h_p(p) + b_4h_r(r)\right\},\\
\frac{dn}{dt} &= -\gamma_n n + a_n b_2 q_{\tau_{nm}} h_n(n_{\tau_{nm}}),\\
\frac{dp}{dt} &= -\gamma_p p + a_p b_3\left\{q_{\tau_{pm}}h_p(p_{\tau_{pm}}) - \mu_3 q_{\tau_{psum}}h_p(p_{\tau_{psum}})\right\},\\
\frac{dr}{dt} &= -\gamma_r r + a_r b_4\left\{q_{\tau_{rm}}h_r(r_{\tau_{rm}}) - \mu_4 q_{\tau_{rsum}}h_r(r_{\tau_{rsum}})\right\},
\end{aligned}
\tag{2}
$$

where q, n, r, and p are nondimensional stem cells, neutrophils, erythrocytes, and platelets, respectively. Subscripts indicate delays: $q_1 = q(t-1)$ and so on.

The functions h_q, h_n, h_r, and h_p are Hill functions given by

$$h_q = \frac{1}{1+q^4}, \quad h_n = \frac{\theta_1}{\theta_1+n}, \quad h_r = \frac{1}{1+r^{6.96}}, \quad h_p = \frac{1}{1+p^{1.29}},$$

where the values of the exponents were determined by previous modeling efforts in the respective cell lines [22, 18, 1]. This model is a nondimensional version of that given in [6].

As in (1), quiescent pluripotential HSCs q can leave the quiescent compartment and enter the proliferating compartment (at a rate $b_1 h_q(q)$). After a proliferation time τ_s, which in these units is 1, they re-enter the quiescent compartment having been multiplied by 2 and lost at a rate γ_s during the time for which they were proliferating ($\mu_1 = 2e^{-\tau_s \gamma_s}$). Alternatively, they can differentiate into each of the three peripheral cell lines, at rates $h_n(n)$, $h_p(p)$, and $h_r(r)$. All of these rates are negative feedback functions. In each circulating compartment there is a random loss rate (γ_n, γ_r, γ_p). The platelets and erythrocytes have an additional loss due to senescence, resulting after integration over the maturation structure of these compartments, in losses with delays τ_{psum} and τ_{rsum}.

This model is a simplified version of (1) in several ways: the regulatory functions h_i are functions of only one tissue type, the coupling terms f_i between the tissue types are simply random loss terms with no intertissue interaction, and we have not explicitly included maturation structure. The resulting system, (2), is a set of four coupled delay-differential equations with six delays. In the appendix, a list of the parameter values and their dimensional counterparts is given in Table 1, along with a list and description of the delays in the system.

3. Bifurcation analysis.

The dynamics of the stem cell compartment have been found to be critical in modeling CN and PCML [17, 14, 1, 6]. The stem cell parameters are not well constrained by current observation, and indeed, even the structure of the first equations of (1) and (2) are simplifications of a much more complex (and not fully understood) set of dynamics. However, given the parameters that must change in order for the system to mimic observed disease data [6], we have the opportunity to use bifurcation analysis to better understand the relevant dynamics of the stem cells. The parameters of interest are the stem cell death rate, which is inversely related to μ_1, the rate of re-entry to the proliferative compartment b_1, and the amount of differentiation out of the hematopoietic stem cell compartment.

The neutrophil compartment (second equation of (2)) is also important in CN and PCML. Changes in the amplification a_n in the neutrophil line are necessary to mimic CN, along with changes in the stem cells. We wish to characterize the effects of changing a_n in the comprehensive model for two reasons: we believe that reduced a_n is the central cause of CN, and G-CSF, the most common treatment for CN, raises a_n, often to above its normal steady-state value.

3.1. The stem cell compartment.

One mechanism that has been suggested for the onset of oscillations in neutropenia [15, 1, 6] is that there is an increase in apoptosis in the neutrophil line which destabilizes the stem cell compartment. Apoptosis is preprogrammed cell death; in this hypothesis, increased apoptosis affects cells during the maturation phase and fewer cells reach maturity. The negative feedback $h_n(n)$ responds, creating a greater demand for stem cells to differentiate into the neutrophil lineage. The stem cell compartment begins to oscillate and the oscillations are subsequently observed in the neutrophils, platelets, and

reticulocytes. An alternative mechanism suggested by [20] is that there is a failure in the peripheral neutrophil regulation, i.e., in the n compartment.

To investigate this further, we study the effects of parameter changes on the stem cell compartment when it is decoupled from the peripheral cell lines. This decoupled model for the stem cells alone is given by

$$\frac{dq}{dt} = -\frac{b_1}{1+q^4}q + \mu_1 \frac{b_1}{1+q_1^4}q_1 - \delta q, \tag{3}$$

where δ is a constant summarizing the total differentiation out of the stem cell compartment and into the peripheral lines. If apoptosis were to increase in any of the circulating cell lines, the negative feedback would respond by increasing the differentiation out of the stem cell compartment, corresponding to an increase in δ.

Equation (3) has a unique nontrivial positive steady-state solution q^* given by

$$q^*(1+q^{*4}) = \frac{b_1(\mu_1 - 1)}{\delta}.$$

In the range of interest here, $q^* \sim 2$ or more, so that $q^{*4} >> q^*$ and we can write $q^* \sim (\frac{\lambda_1}{\delta})^{1/5}$, where $\lambda_1 = b_1(1-\mu_1)$. Equation (3) may be linearized about its steady state, with $z = q - q^*$, to give

$$\dot{z} = \alpha z - \Lambda z_1$$

with

$$\alpha = b_1 h'(q^*) - \delta \ \sim \ \delta \frac{5-\mu_1}{\mu_1 - 1}$$

and

$$\Lambda = b_1 h'(q^*) \ \sim \ \frac{4\delta}{\mu_1 - 1}.$$

Both α and Λ are positive. The characteristic equation for the eigenvalues σ is

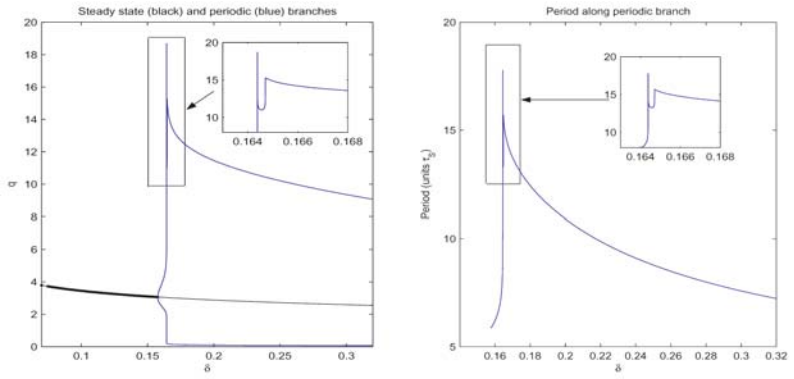
$$\sigma = \alpha - \Lambda e^{-\sigma}.$$

Letting $\sigma = i\omega$, the boundary at which instability occurs is defined by

$$\omega = \alpha \tan\omega, \quad \Lambda_1 = \frac{\omega}{\sin\omega},$$

and instability arises when $\Lambda > \Lambda_1$. As δ increases, α increases. For small α, the solution to $\omega = \alpha \tan\omega$ lies in the interval $[0, \pi/2)$, where $\Lambda_1 = \frac{\omega}{\sin\omega}$ is positive. For $\alpha > 1$, however, $\omega \in [\pi, 3\pi/2)$ and Λ_1 is negative. This ensures that $\Lambda > \Lambda_1$. We therefore expect a Hopf bifurcation when $\alpha = 1$, or $\delta = 1/6$ (when $\mu_1 = 1.6$, the normal value).

This confirms that an increase in apoptosis in one of the circulating cell lines can initiate oscillations in the stem cells. We compute the stability of the periodic solution using DDE-BIFTOOL [9, 10], and it is stable, though some of the Floquet multipliers are very close to 1 in magnitude, so that we may expect long transient approaches to the orbit. Note that α, Λ, and therefore the position of the Hopf bifurcation do not depend on b_1, the rate of re-entry to the proliferative compartment (see Figure 2). However, below a critical value of b_1 the



(a) q vs. δ (b) Period vs. δ

Figure 1. *Steady-state and periodic oscillations in the decoupled stem cell compartment. Where the steady-state branch is shown as a thick line (left side), it is stable, and where it is thin (right side), it is unstable. The maximum and minimum values of the oscillations in the quiescent (q) cells are shown for the resulting periodic branch* (a)*, and the period is shown in* (b).

stem cells cannot maintain a positive steady state and the trivial solution is the only solution to (3).

Figure 1 shows the steady-state branch of solutions (thick line: stable; thin line: unstable) and the loss of stability at the bifurcation $\delta \sim 1/6$. Compared to the steady-state value of $\delta = 1/8$, the change required to initiate oscillations is not large, and certainly not unrealistic biologically. To cause such a change in the differentiation function, the value of n need only decrease by a factor of $1/3$ from the healthy value. The period along the branch is shown in Figure 1 in the right-hand plot and corresponds to the observed period of 13–50 days (4–16 units). Periods longer than those normally observed (> 10 units) occur only in a small part of the parameter space, in correspondence with observation.

We now turn to the parameters μ_1 and b_1, tracking the position of the Hopf bifurcation as b_1, δ, and μ_1 change. Figure 2 shows the results. In Figure 2(a), the "healthy" steady-state solution is in the upper stable region shown on the plot. As μ_1 decreases (the death rate in the stem cell compartment rises), oscillations begin, and then cease for values of approximately $\mu_1 < 1$. The steady-state solution in this lower region is the trivial solution, $q \equiv 0$, because if the death rate γ_S rises too much, the stem cells can no longer maintain their population. The colors shown in the oscillating regions of the plots in Figure 2 represent the period of oscillations.

Figure 2(b) shows the stability curve and periods of oscillation resulting from the Hopf bifurcation as δ increases. As the death rate in the stem cell compartment falls, μ_1 rises and the stem cells are better able to maintain a stable steady-state equilibrium in response to a need for more circulating cells (an increase in δ). In Figure 2(c), there is a threshold value of δ below which the system is always stable no matter what the value of b_1, but at higher values of δ, an increase in b_1 can destabilize the system and result in the initiation of oscillations. However, since the healthy value of b_1 (22.4) is much above this range, we conclude that oscillations are caused by an increase in δ rather than an increase in b_1.

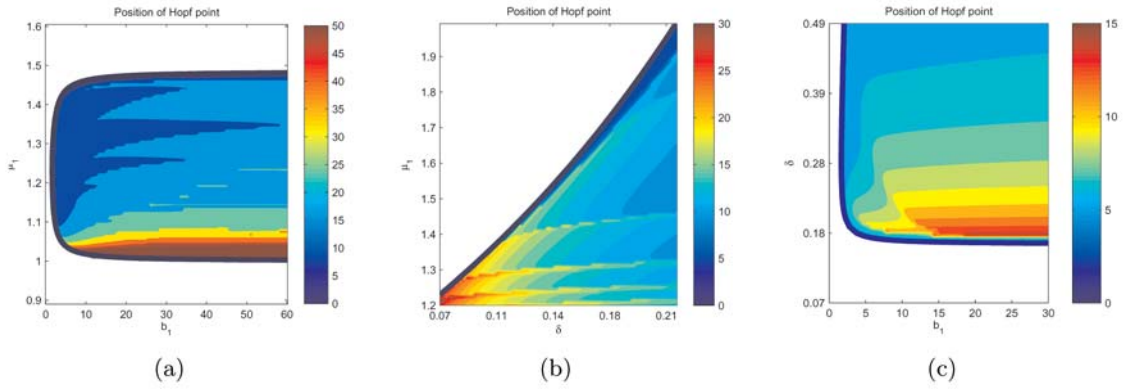


Figure 2. *Hopf bifurcation, position dependent on b_1, μ_1, and δ in the decoupled stem cell model of (3). The colored regions are the regions of Hopf instability, and their boundaries give the location of the bifurcation. The color indicates the period of oscillation.*

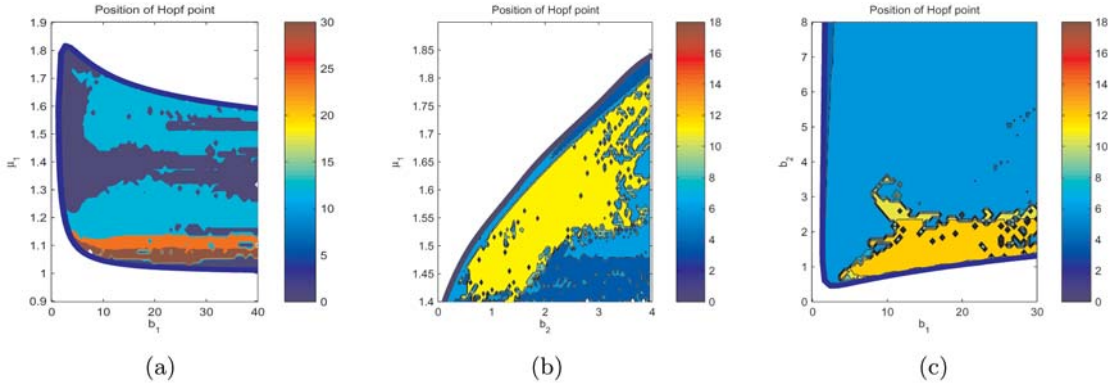


Figure 3. *Hopf bifurcation dependent on b_1, μ_1, and b_2 in the full model of (2). Thick solid lines are the location of the Hopf instability, colored regions are oscillatory, and the color indicates the period of oscillation.*

3.2. The full model: Hopf bifurcations. We now wish to find the corresponding Hopf bifurcation in the full model, if it exists, and compare its location to that in the stem cell compartment alone, under corresponding parameter changes. Due to the complexity of the model, this is done numerically. Also, δ is not constant in the full model; here, we can explicitly raise b_2 (increase output from the hematopoietic stem cells q) and/or decrease a_n. The bifurcation occurs as a_n decreases below $a_n = 21.5$, and is again supercritical. Figure 3 shows its position under the analogous parameter changes to Figure 2.

The qualitative behavior of the full model is similar to that of the stem cell compartment alone, and there are biological interpretations for the differences between the two. For example, in Figure 3(a) oscillations begin as μ_1 is decreased, and then give way to the trivial solution below $\mu_1 = 1$. In the full model there is a region in the upper-left portion of the plot where an increase in b_1 can stabilize oscillations: an increased rate of re-entry to the proliferative phase can compensate for increased stem cell apoptosis, unlike in Figure 2(a). In Figure 3(b), the concavity of the Hopf curve is reversed, but otherwise the Hopf position as μ_1 and b_2 are

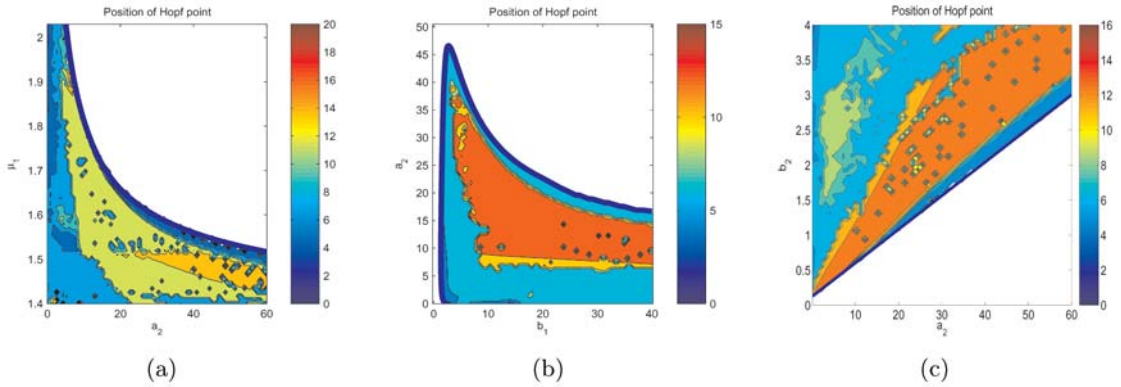


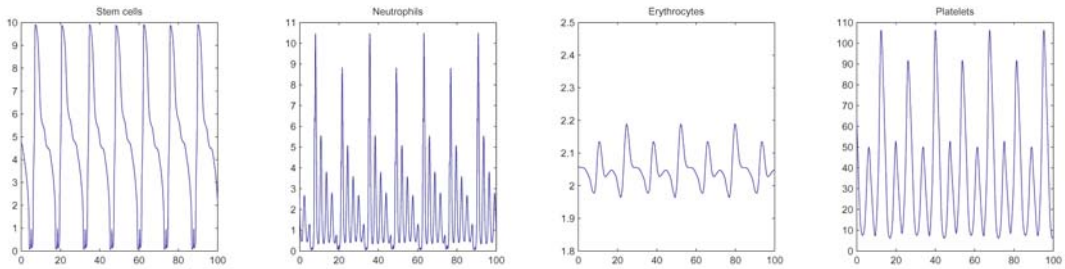
Figure 4. *Hopf position as a_n changes. Colored regions are oscillatory, and the color indicates the period of oscillation.*

changed is qualitatively the same. Figure 3(c) shows that unlike in the stem cell model alone, an increase in b_1 can again compensate for an increase in b_2 and stabilize oscillations. The full model is better able to recover from destabilizing changes than the stem cell model without the coupling to the peripheral cell lines.

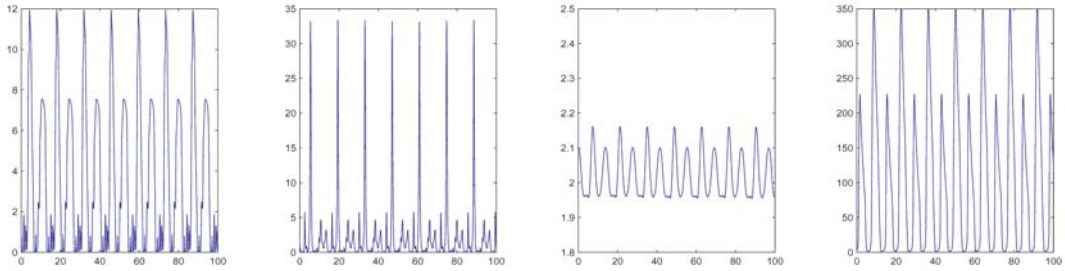
Decreasing b_2 and increasing μ_1, based on these results, are probably the two best methods of stabilizing existing oscillations. However, the most common treatment for CN is the administration of granulocyte-colony stimulating factor (G-CSF) [25, 13, 2, 19], which is known to increase neutrophil levels by reducing apoptosis in the neutrophil precursors [13], and is therefore modeled by an increase in the parameter a_n. We track the Hopf bifurcation with respect to a_n; the results are shown in Figure 4. In each plot of the figure it is clear that increasing a_n has the desired effect of stabilizing the oscillations, though significant increases may be necessary if μ_1 or b_1 is small. If they are too small, oscillations may not be stabilized at all.

In clinical data, G-CSF sometimes abolishes oscillations but may actually increase their amplitude [16]. The bifurcation analysis accounts for this. We would hypothesize that when G-CSF does not abolish oscillations, it is because there is not enough capacity for self-maintenance in the stem cells: their apoptosis rate is too high, and/or their proliferation rate is too low. Furthermore, the results shown in Figure 4(b) show that increasing b_1 may stabilize oscillations if a_2 is high enough; clinically, stem cell factor would raise b_1 and is, in fact, sometimes used to treat CN.

In this section, we have found that oscillations begin if the death rate in the stem cell compartment rises (μ_1 decreases), if there is increased differentiation out of the stem cells (b_2 increases), if there is a reduction in the proliferation rate in the stem cell compartment (b_1 decreases), or if there is increased apoptosis in the neutrophil line (a_n decreases). In some regions, an increase in b_1 may recover stability, but this recovery is not possible in the reduced model containing only the stem cells. The otherwise similar natures of the plots shown in this and the previous section lend credibility to the claim that destabilizations in the stem cell compartment are at the heart of oscillatory behavior in the hematological system.



(a) Low-amplitude branch (branch 1) with period 28 dimensionless time units



(b) High-amplitude branch (branch 2) with period 14 dimensionless time units

Figure 5. *Two locally stable periodic branches when $a_n = 58$ and all other parameters are set to their default values.*

3.3. Multistability. Not only does G-CSF not always suppress oscillations, it may even initiate them [16]. Furthermore, in a model containing only the stem cells and the neutrophils, Bernard, Bélair, and Mackey [1] found that there was a small range of a_n where there is bistability of two periodic solutions. Motivated by these results, we explore the existence of multistability when a_n is increased from its normal value (35.6) to 58. This is a simple way to mimic a patient undergoing G-CSF treatment. All the other parameters are left at the values given in Table 1; this point in parameter space will be referred to as point P.

At point P, the steady-state solution is locally stable. However, we also find two locally stable periodic orbits at P (see Figure 5). The low-amplitude orbit (Figure 5(a)) has a period of 28 units, and the orbit shown in Figure 5(b) has a period of 14 units with a higher-frequency oscillation at a period of 7 units, corresponding to the secondary bump in the stem cell profile. A wide range of simulations using different initial conditions failed to uncover any other locally stable periodic solutions.

The low-amplitude orbit is more reminiscent of treated neutropenia than the high-amplitude orbit, where neutrophil levels at maximum are more than 30 times their steady-state value (1 in these units). Also, the platelet levels in the high-amplitude branch are much higher than are observed clinically, and for more of the oscillation, than in the low-amplitude orbit.

4. Branch switching. The fact that multistability exists for this biologically reasonable choice of parameter values naturally leads to the question of how the oscillations might be stabilized; stabilization would correspond clinically to an end of disease symptoms. In this

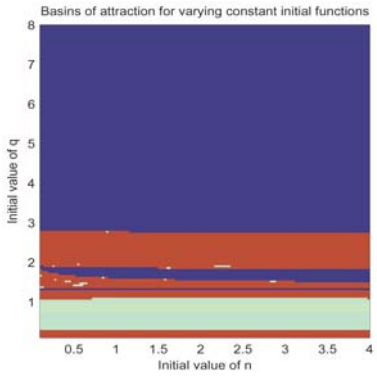


Figure 6. *Basins of attraction for varying initial q and n values. Blue corresponds to the steady-state branch, red to the low-amplitude branch, and green to the high-amplitude branch.*

section we investigate the possibility of changing the initial function to guide which behavior the system approaches, and then we examine the effects of temporary parameter changes on the system when it is oscillating on the low-amplitude branch in Figure 5. In the computations, we use a Runge–Kutta integration scheme with step sizes ranging from 0.005 to 0.05 (most commonly 0.01) dimensionless time units, in Bard Ermentrout's software *xppaut*.[1]

4.1. Effect of initial functions. Because the system in (2) has delays, it is necessary to specify initial functions for each of the four variables, during the period $[-\tau_{\max}, 0]$. We first set all variables constant on $[-\tau_{\max}, 0]$ and examine the effect of the choice of these constants. Figure 6 shows the results when only the initial values of q and n on $[-\tau_{\max}, 0]$ (the initial numbers of stem cells and neutrophils, respectively) are changed; each branch is assigned a color: blue for the stable steady-state solution, red for the low-amplitude orbit, and green for the high-amplitude orbit. The same color scheme applies to figures showing basins of attraction in this and subsequent sections. The lack of other colors (and hence of other periods) is evidence that these three solutions are the only locally stable solutions at this point in parameter space.

The most prominent feature of Figure 6 is that the initial value of n has almost no effect on the eventual behavior of the system, even when n is increased up to three times the normal value of 1. However, the initial population of stem cells has a significant effect between 0 and 4, though the value at the steady state is 3.2. This is further indication that oscillations in CN and PCML, and in the hematological system in general, are most strongly connected to the dynamics of the stem cell compartment.

This result leads to the question of what the effects of nonconstant initial functions for q, particularly periodic initial functions, would be. We next allow the initial function for q to change, leaving the initial functions for n, r, and p constant at their healthy, steady-state values. On $[-\tau_{\max}, 0]$, q is given by

$$q(t) = A_q \sin\left(\tfrac{2\pi}{T_q} t\right) + M_q, \tag{4}$$

[1] http://www.math.pitt.edu/~bard/xpp/xpp.html

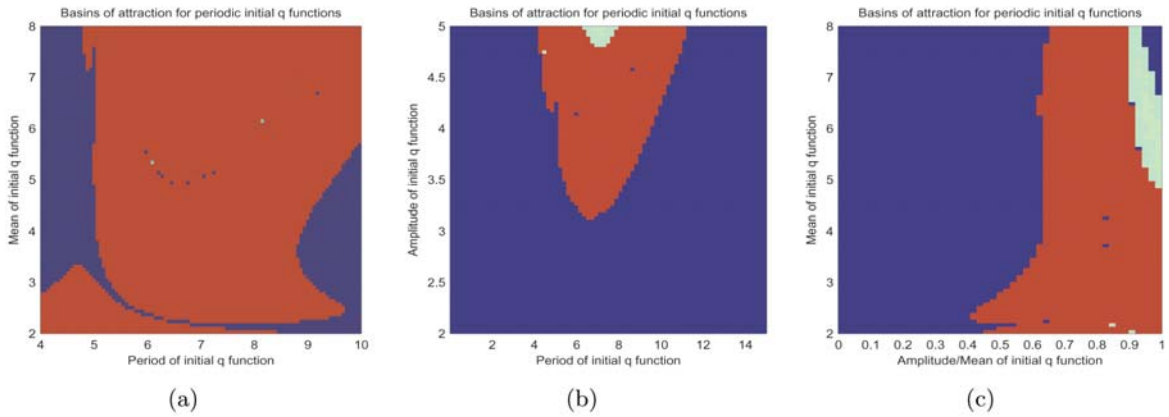


Figure 7. *Basins of attraction for various periodic initial functions. Only q is varying; the other variables are constant on $[-\tau_{\max}, 0]$. Blue is steady-state, red is the low-amplitude branch, and green is the high-amplitude branch.*

so that the three parameters defining this function are the amplitude of the oscillations, A_q, the period, T_q, and the mean, M_q. Figure 7 shows the resulting basins of attraction. In Figure 7(a), note that for most of the plotted range of periods and means the solution stabilizes to the low-amplitude branch, where in Figure 7(b) most of the solutions stabilize to the steady-state branch. Figure 7(c) shows the basins of attraction when the mean M_q and the fraction A_q/M_q are varied; note that a value of A_q/M_q greater than 1 is not possible because it would result in negative q values in the initial function for q on $[-\tau_{\max}, 0]$.

Figures 6 and 7 show the manner in which the initial function determines whether the system ultimately oscillates (mimicking CN) or stabilizes (mimicking a healthy individual). However, the stem cells, located in the bone marrow, are difficult to identify and manipulate, and any such manipulation would presumably also change several parameters in the system (and not only the values of q). It is therefore difficult to translate these results into realistic medical interventions. In the following section we investigate temporary interventions and their effect on the stability of the system.

4.2. Branch jumping by short-term parameter changes. We explore the possibility of branch jumping, i.e., perturbing the system so that it moves from one solution to another, using temporary pulses in the system parameters a_n, a_r, and a_p. These represent amplification in the neutrophil, erythrocyte, and platelet lines, respectively, and they are chosen because it is known that the administration of regulatory cytokines increases these amplification factors via the inhibition of apoptosis: G-CSF increases the neutrophil amplification, while erythropoietin and thrombopoietin have similar effects on the erythrocyte and platelet lines, respectively [3, 21, 23]. Thrombopoietin and erythropoietin are not used to treat neutropenia, because they are primilary associated with platelet and erythrocyte dynamics and neutropenia is characterized by decreased neutrophil levels. Furthermore, the administration of thrombopoietin causes an immune response [8], rendering it of limited value. However, there may be other clinical interventions which would raise the platelet amplification a_p in the manner we investigate here.

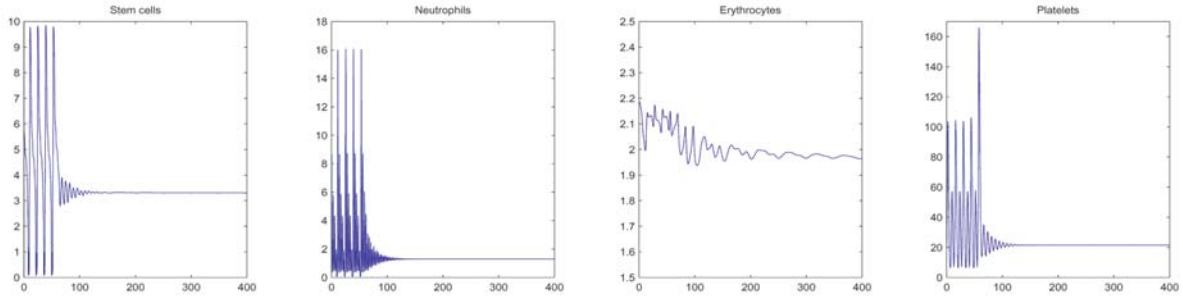


Figure 8. *Stabilization of the low-amplitude branch with a pulse in a_p of amplitude* 40 *and duration* 4.75*, mimicking thrombopoietin administration, delivered at* $t = 50$*.*

To simulate the effects of a brief administration of the relevant cytokines, we have computed the basins of attraction for parameter functions of the form

$$a_i = a_i^{\text{base}} + a_i^{\text{amp}} H(t - t_1) H(t_2 - t), \tag{5}$$

where $H(t)$ is the Heaviside function and $i \in \{n, r, p\}$. This function represents a change in the parameter a_i from a_i^{base}, a base value, to $a_i^{\text{base}} + a_i^{\text{amp}}$ during the time interval $[t_1, t_2]$. For each basin computation, we choose an initial function well within the basin of attraction of the periodic low-amplitude solution (Figure 5(a)). We define the phase to be 0 at time $t_1 = 200$ to consistently examine the effect of the phase time at which the pulse is delivered. Because each computation is performed from the same initial condition at the same parameter set, this is unambiguous. Figure 8 shows an example of a successful stabilization using a pulse of the parameter a_p, according to (5). Note that the erythrocytes are considerably slower in reaching equilibrium than the other variables, due to the long delay τ_{RS} in the erythrocyte compartment.

Because G-CSF is the most common treatment for CN, one would suspect that pulses of increased amplification a_n would have a significant effect on the qualitative behavior of the solutions. However, oscillations are rarely suppressed using pulses in a_n. Figure 9 shows the results of exploring various durations and amplitudes of the pulses. Note that the regions in which the stabilization attempt was successful (blue areas in the figure) are small, and that the amplitude of the pulse needs to be high. The normal value of a_n is 35.6, and so the values required for stabilization represent an increase by a factor of 6 or more. This would translate into a high dose of G-CSF, which would have to be precisely timed to stabilize the system.

Figure 10 shows similar basin of attraction plots for pulses in a_r while varying the amplitude, duration, and phase of the pulse. Again, while stabilization is possible using a pulse in a_r, regions where the solution stabilizes are small and the amplitude of the pulse is large compared to the healthy value of a_r (~ 1). However, the regions of stabilization are larger and the pulses do not need to be as precisely timed as those in a_n. This gives the somewhat counterintuitive result that stabilization of oscillations in CN and PCML may be more easily accomplished with erythropoietin administration than with G-CSF, though the dosage would have to be high to obtain the necessary large increases in a_r.

The most promising stabilization results we have obtained have been with pulses using

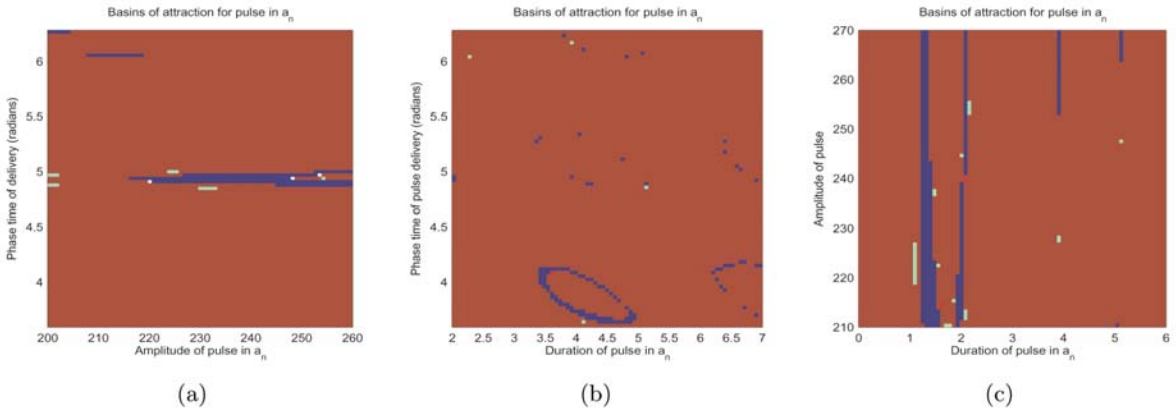


Figure 9. *Basins of attraction for an a_n pulse, simulating G-CSF administration.*

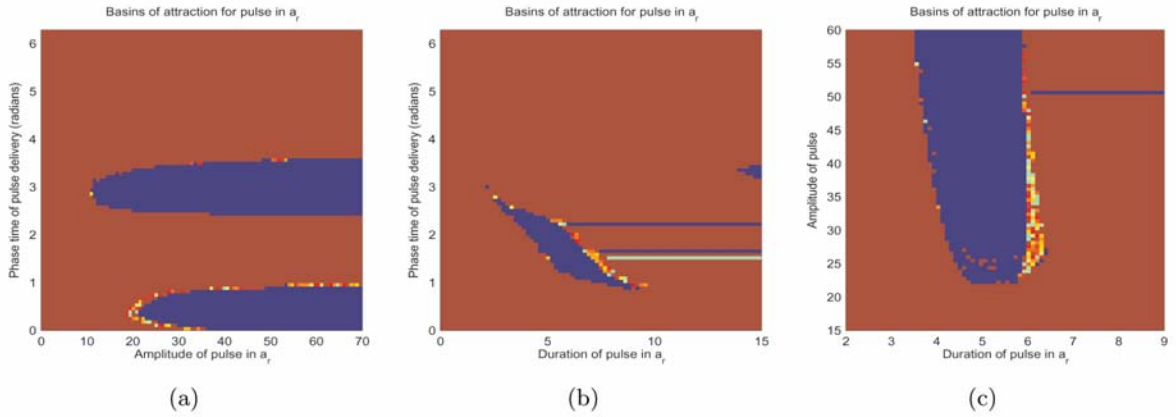


Figure 10. *Basins of attraction for pulses in a_r, simulating erythropoietin administration.*

the parameter a_p; Figure 11 shows the results. The blue regions, where the oscillations have been stabilized by the pulse, are comparatively large. This indicates that the stabilization is a robust phenomenon at this point (P) in parameter space.

The normal value of a_p is approximately 58 and the minimum additional pulse required is only about 20, so that a comparatively small and achievable pulse size is sufficient. Figure 11(a) shows clearly that some points in the oscillation are more advantageous times at which to deliver the pulse, namely, at phases 0, $\pi/2$, π, and $3\pi/2$. Figure 11(b) shows that changing the pulse duration shifts these bands. In Figure 11(c), it is interesting to note that increasing the duration of the pulse does not necessarily increase the likelihood that stabilization will occur. While it does appear that increasing the amplitude of the pulse makes stabilization more likely, it is not usually the best way to stabilize the system. Changing the duration or timing of the pulse is more effective.

With reference to (2), we can partially account for the somewhat surprising fact that perturbations to the platelets are best able to suppress oscillations. Any increase in a_n, a_p, or a_r will cause a decrease in δ, which we know from section 3.1 can stabilize the stem cell

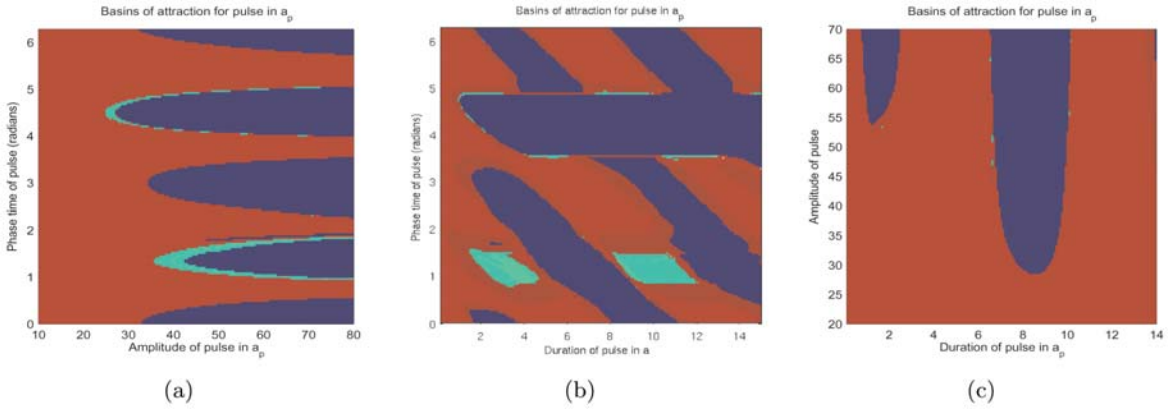


Figure 11. *Basins of attraction for pulses in a_p, simulating thrombopoietin administration.*

compartment (at least, when considered on its own). Suppose that when an amplification a_i increases, the corresponding compartment n, r, or p responds by trying to "track" the change and reach a new pseudosteady state. Then we can write

$$\begin{aligned}
\frac{\partial \delta}{\partial a_n} &= \frac{\partial \delta}{\partial n}\frac{\partial n}{\partial a_n} \approx \frac{-b_2^2\theta_1^2 q}{(\theta+n)^2(\gamma_n\theta+1)} = O(10^{-2}),\\
\frac{\partial \delta}{\partial a_p} &= \frac{\partial \delta}{\partial p}\frac{\partial p}{\partial a_p} \approx \frac{-b_3 s_p p^{s_p} q(1-\mu_3)}{(1+p^{s_p})^2(1_+ s_p)p^{s_p}} = O(10^{-3}),\\
\frac{\partial \delta}{\partial a_p} &= \frac{\partial \delta}{\partial p}\frac{\partial p}{\partial a_p} \approx \frac{-b_3 s_p p^{s_p} q(1-\mu_3)}{(1+p^{s_p})^2(1_+ s_p)p^{s_p}} = O(10^{-5}).
\end{aligned} \tag{6}$$

A change in a_p is far more effective than a change in a_r at changing δ. In addition, the time scales for the n and p equations ($1/\gamma_n$ and $1/\gamma_p$) are much faster than for the r equation, as $1/\gamma_r \sim O(10^3)$. The value of r thus does not respond quickly, and even if it did, (6) indicates that the response would not be as strong as the platelet response. While the neutrophil response time is adequate, the neutrophils are not in a pseudosteady state on the low-amplitude branch; rather, when q is high they undergo high-frequency transient oscillations (see Figure 5(a) and the discussion in [4]). They therefore do not have a consistent effect on δ when a_n is raised.

In addition, it has been found that q undergoes relaxation oscillations and therefore has slow and fast phases [12, 4]. This, together with the details of the response of p to an increase in a_p, probably accounts for the sensitivity of the results on the phase of oscillation; when q is in the lower half of the slow portion of the oscillation, decreasing δ may have more effect than when q is high.

5. Conclusions. Despite the potential complexity of stem cell/tissue models, the model analyzed here for hematopoietic production and regulation shows some interpretable dynamics. The location and behavior of the Hopf bifurcation believed to give rise to CN in the full model occurs in the stem cell model when it is decoupled from the mature tissues, suggesting a possible avenue for analysis of more complicated versions of the model given in (1). Inter-

estingly, the inclusion of the tissue types is a stabilizing influence: the full model of (2) is better able to recover from destabilizing changes than the decoupled model of (3).

The results presented here support the hypothesis that oscillations in dynamical hematological diseases are driven by oscillations in the stem cell compartment. We find that oscillations consistent with those observed for CN are initiated when an increase in demand for circulating blood cells causes the stem cell compartment to undergo a Hopf bifurcation. Our analysis accounts for the variable success of G-CSF in suppressing oscillations in CN. When G-CSF fails to suppress oscillations, this may be due to a lack of regenerative capacity in the hematopoeitic stem cells—a death rate that is too high, or too little re-entry into the proliferative compartment, or a combination. Our results also account for why stem cell factor may be a promising treatment for CN.

For G-CSF-treated CN, we find two locally stable periodic orbits and a locally stable steady state. The choice of solution depends most strongly on the initial values of the stem cells. We find that, using short-term pulses of increased amplification in any of the peripheral lines, it is possible to suppress oscillations with temporary measures that have clear medical interpretations. This is most easily accomplished with pulses of increased a_p, the amplification in the platelet line, mimicking temporary administration of thrombopoietin. While thrombopoietin may not be promising due to the initiation of an immune response that targets it, other interventions that temporarily raise platelet numbers could be considered. The timing of the pulse is important, but because the regions in parameter space where the stabilization is successful are quite large, finding an appropriate time is not difficult. These results suggest that combining G-CSF with short-term platelet-enhancing drugs may be a promising approach to abolishing oscillations. These somewhat surprising effects of tissue level perturbations on the full model may be understood in terms of their effects on the decoupled stem cell model: perturbations that would stabilize the decoupled stem cell model are more able to switch the full model from oscillatory to stable dynamics than those that would not.

Appendix.

Table 1

Definitions and typical values of the dimensionless parameters of the model. Dimensional parameters correspond to those given in [6].

Symbol	Description	Dimensional equivalent	Typical value
b_1	stem cell proliferation	$\tau_s k_0$	22.4
b_2	neutrophil differentiation factor	$\tau_s f_0$	1.1
b_3	platelet differentiation factor	$\tau_S \bar{\kappa}_P$	3.3
b_4	erythrocyte differentiation factor	$\tau_S \bar{\kappa}_R$	3.3
θ_1	Hill function parameter	θ/N^*	0.055
s_p	Hill exponent	m	1.29
s_r	Hill exponent	r	6.96
μ_1	proliferation with loss factor	$2e^{-\gamma_S \tau_S}$	1.6
μ_3	loss factor	$e^{-\gamma_P \tau_{PS}}$	0.24
μ_4	loss factor	$e^{-\gamma_R \tau_{RS}}$	0.9
γ_n	neutrophil death rate	$\tau_S \gamma_N$	6.7
γ_r	erythrocyte death rate	$\tau_S \gamma_R$	0.003
γ_p	platelet death rate	$\tau_S \gamma_P$	0.42
a_n	neutrophil amplification	$A_N \theta_2 / N_1^*$	35.63
a_p	platelet amplification	$A_P \theta_2 K_P^{1/s_p}$	58
a_r	erythrocyte amplification	$A_R \theta_2 K_R^{1/s_r}$	1.1
τ_s	stem cell proliferation time	τ_s/τ_S	1
τ_{nm}	neutrophil maturation time	τ_{NM}/τ_S	1.25
τ_{pm}	platelet maturation time	τ_{PM}/τ_S	2.5
τ_{ps}	platelet aging time to senescence	τ_{PS}/τ_S	5.9
τ_{psum}	$\tau_{pm} + \tau_{ps}$		8.4
τ_{rm}	erythrocyte maturation time	τ_{RM}/τ_S	2.1
τ_{rs}	erythrocyte aging time to senescence	τ_{RS}/τ_S	42.9
τ_{rsum}	$\tau_{rm} + \tau_{rs}$		45
q^*	stem cell steady-state value	Q^*/θ_2	3.6
n^*	neutrophil steady-state value	N^*/N^*	1
p^*	platelet steady-state value	$P^* K_p^{1/s_p}$	21
r^*	erythrocyte steady-state value	$R^* K_r^{1/s_r}$	2.2

REFERENCES

[1] S. Bernard, J. Bélair, and M. Mackey, *Oscillations in cyclical neutropenia: New evidence based on mathematical modeling*, J. Theor. Biol., 223 (2003), pp. 293–298.

[2] M. Bonilla, A. Gillio, M. Ruggeiro, N. Kernan, J. Brochstein, M. Abboud, L. Fumagalli, M. Vincent, J. Gabrilove, L. M. Souza, K. Welte, and R. O'Reilly, *Effects of recombinant human granulocyte colony-stimulating factor on neutropenia in patients with congenital agranulocytosis*, New England J. Medicine, 320 (1989), pp. 1574–1580.

[3] O. Borge, V. Ramsfjell, L. Cui, and S. Jacobsen, *Ability of early acting cytokines to directly promote survival and suppress apoptosis of human primitive CD34+CD38- bone marrow cells with multilineage potential at the single cell level: Key role of thrombopoietin*, Blood, 90 (1997), pp. 2282–2292.

[4] C. Colijn, A. Fowler, and M. Mackey, *High frequency spikes in long period blood cell oscillations*, J. Math. Biol., 53 (2006), pp. 499–519.

[5] C. Colijn and M. Mackey, *A mathematical model of hematopoiesis:* I. *Periodic chronic myelogenous leukemia*, J. Theor. Biol., 237 (2005), pp. 117–132.

[6] C. Colijn and M. Mackey, *A mathematical model of hematopoiesis:* II. *Cyclical neutropenia*, J. Theor. Biol., 237 (2005), pp. 133–146.

[7] P. de Coppi, G. Bartsch, Jr., M. Siddiqui, T. Xu, C. C. Santos, L. Perin, G. Mostoslavsky, A. C. Serre, E. Y. Snyder, J. J. Yoo, M. E. Furth, S. Soker, and A. Atala, *Isolation of amniotic stem cell lines with potential for therapy*, Nature Biotechnology, 25 (2007), pp. 100–106.

[8] D. C. Dale, J. L. Nichol, D. A. Rich, D. M. Best, S. J. Slichter, W. Sheridan, and P. Hunt, *Chronic thrombocytopenia is induced in dogs by development of cross-reacting antibodies to the MpL ligand*, Blood, 90 (1997), p. 3456.

[9] K. Engelborghs, T. Luzyanina, and D. Roose, *Numerical bifurcation analysis of delay differential equations using DDE-BIFTOOL*, ACM Trans. Math. Software, 28 (2002), pp. 1–21.

[10] K. Engelborghs, T. Luzyanina, and G. Samaey, *DDE-BIFTOOL v.* 2.00*: A MATLAB Package for Bifurcation Analysis of Delay Differential Equations*, Tech. rep., Department of Computer Science, K. U. Leuven, Leuven, Belgium, 2001.

[11] P. Fortin and M. Mackey, *Periodic chronic myelogenous leukemia: Spectral analysis of blood cell counts and etiological implications*, Brit. J. Haematol., 104 (1999), pp. 336–245.

[12] A. C. Fowler and M. C. Mackey, *Relaxation oscillations in a class of delay differential equations*, SIAM J. Appl. Math., 63 (2002), pp. 299–323.

[13] W. P. Hammond, T. H. Price, L. M. Souza, and D. C. Dale, *Treatment of cyclic neutropenia with granulocyte colony stimulating factor*, New England J. Medicine, 320 (1989), pp. 1306–1311.

[14] C. Haurie, D. C. Dale, R. Rudnicki, and M. C. Mackey, *Modeling complex neutrophil dynamics in the grey collie*, J. Theor. Biol., 204 (2000), pp. 504–519.

[15] C. Haurie, M. C. Mackey, and D. C. Dale, *Cyclical neutropenia and other periodic hematological diseases: A review of mechanisms and mathematical models*, Blood, 92 (1998), pp. 2629–2640.

[16] C. Haurie, M. C. Mackey, and D. C. Dale, *Occurrence of periodic oscillations in the differential blood counts of congenital, idiopathic and cyclical neutropenic patients before and during treatment with G-CSF*, Exper. Hematol., 27 (1999), pp. 401–409.

[17] M. C. Mackey, *Dynamic haematological disorders of stem cell origin*, in Biophysical and Biochemical Information Transfer in Recognition, J. G. Vassileva-Popova and E. V. Jensen, eds., Plenum Publishing, New York, 1979, pp. 373–409.

[18] J. Mahaffy, J. Bélair, and M. Mackey, *Hematopoietic model with moving boundary condition and state dependent delay: Applications in erythropoiesis*, J. Theor. Biol., 190 (1998), pp. 135–146.

[19] A. R. Migliaccio, G. Migliaccio, D. C. Dale, and W. P. Hammond, *Hematopoietic progenitors in cyclic neutropenia: Effect of granulocyte colony stimulating factor in vivo*, Blood, 75 (1990), pp. 1951–1959.

[20] A. Morley, *Periodic diseases, physiological rhythms and feedback control–a hypothesis*, Aust. Ann. Med., 3 (1970), pp. 244–249.

[21] A. Ritchie, S. Vadhan-Raj, and H. Broxmeyer, *Thrombopoietin suppresses apoptosis and behaves as a survival factor for the human growth factor-dependent cell line, M*07*e*, Stem Cells, 14 (1996), pp. 330–336.

[22] M. Santillan, J. Mahaffy, J. Bélair, and M. Mackey, *Regulation of platelet production: The normal response to perturbation and cyclical platelet disease*, J. Theor. Biol., 206 (2000), pp. 585–603.

[23] M. Silva, D. Grillot, A. Benito, C. Richard, G. Nunez, and J. Fernandez-Luna, *Erythropoietin can promote erythroid progenitor survival by repressing apoptosis through bcl*-1 *and bcl*-2, Blood, 88 (1996), pp. 1576–1582.

[24] K. Takahashi and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors*, Cell, 126 (2006), pp. 663–676.

[25] D. G. Wright, R. F. Kenney, D. H. Oette, V. F. LaRussa, L. A. Boxer, and H. L. Malech, *Contrasting effects of recombinant human granulocyte-macrophage colony-stimulating factor (CSF) and granulocyte CSF treatment on the cycling of blood elements in childhood-onset cyclic neutropenia*, Blood, 84 (1994), pp. 1257–1267.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 2, pp. 395–402

Application of the 0-1 Test for Chaos to Experimental Data*

Ian Falconer†, Georg A. Gottwald‡, Ian Melbourne§, and Kjetil Wormnes¶

Abstract. A reliable and efficient method of distinguishing between chaotic and nonchaotic behavior in noise-contaminated, but essentially stationary and deterministic, time series data has far reaching applications. Recently, we proposed a new method of detecting chaos which applies directly to the time series data and does not require phase space reconstruction. To illustrate the effectiveness of the method for experimental data, we analyze data from a bipolar motor.

Key words. testing for chaos, deterministic dynamical systems, experimental data, nonlinear time series

AMS subject classifications. 37M10, 65P20, 37D45

DOI. 10.1137/060672571

The standard technique of distinguishing between regular and chaotic dynamics in deterministic time series data is to calculate the maximal Lyapunov exponent [1, 7, 14]. In the case of physical and biological applications, where the underlying dynamics is unknown, phase space reconstruction [1, 7, 16, 18, 19] is a necessary first step in computing Lyapunov exponents. However, there are problems inherent in phase space reconstruction, as discussed in detail in [6, 17].

Recently, we proposed a new method [9] of detecting chaos which applies directly to the time series data and does not require phase space reconstruction. Moreover, the dimension and origin of the dynamical system are irrelevant. The input is the time series data and the output is 0 or 1 depending on whether the dynamics is nonchaotic or chaotic. This 0-1 test for chaos is equally applicable to maps, ordinary and partial differential equations, and generally to data sets arising from deterministic systems. Of utmost importance is the effectiveness of the method for experimental data. To illustrate our method we analyze data coming from an experimental set up of a bipolar motor in an alternating magnetic field.

We first describe our test. Consider a scalar observable $\Phi(n)$. In an experiment, $\Phi(n)$ is a discrete set of measurement data. Choose $c > 0$ and define

$$p(n) = \sum_{j=1}^{n} \Phi(j) \cos jc \tag{1}$$

for $n = 1, 2, \ldots$. We claim that (i) $p(n)$ is bounded if the underlying dynamics is nonchaotic

*Received by the editors October 17, 2006; accepted for publication (in revised form) by A. Stuart March 28, 2007; published electronically June 1, 2007.
http://www.siam.org/journals/siads/6-2/67257.html
†School of Physics, University of Sydney, NSW 2006, Australia (i.falconer@physics.usyd.edu.au).
‡Mathematics and Statistics, University of Sydney, NSW 2006, Australia (gottwald@maths.usyd.edu.au). This author is partly supported by ARC grants DP0452147 and DP0667065 and by EPSRC grant GR/S22714/01.
§Mathematics and Statistics, University of Surrey, Guildford, Surrey GU2 7XH, UK (ism@math.uh.edu). This author is partly supported by EPSRC grant GR/S22714/01.
¶School of Engineering, University of Sydney, NSW 2006, Australia (wormnes@gmail.com).

(e.g., periodic or quasiperiodic) and (ii) $p(n)$ behaves asymptotically like Brownian motion if the underlying dynamics is chaotic. Later, we briefly explain the justification behind these claims. For the moment, we suppose that the claims are correct and show how to proceed.

The form of our test (1) which involves only the observable $\Phi(n)$ highlights the universality of the test; the origin and nature of the data fed into the diagnostic system (1) are irrelevant for the test. The method is independent of the observable; almost any choice of Φ will suffice [9]. Moreover, the dimension of the underlying dynamical system does not pose practical limitations on the method as is the case for traditional methods involving phase space reconstruction [17]. Of course, it is necessary (as with all other methods) that (i) initial transients have died out so that the trajectories are on (or close to) the attractor at time zero, and (ii) the time series is long enough to allow for asymptotic behavior of $p(n)$. In particular, it is required that the data is essentially stationary as well as deterministic. High-dimensionality of the attractor may impact (ii), since the time series must be sufficient in length that the dynamics explores enough of the attractor; but this is an intrinsic part of the problem and not a drawback of the method.

To determine the character of the growth of the function $p(n)$ defined in (1), i.e., its diffusive behavior, it is natural to look at the mean square displacement of $p(n)$, defined to be $M(n) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} [p(j+n) - p(j)]^2$. If the behavior of $p(n)$ is asymptotically Brownian, i.e., the underlying dynamics is chaotic, then $M(n)$ grows linearly in time; if the behavior is bounded, i.e., the underlying dynamics is nonchaotic, then $M(n)$ is also bounded. The asymptotic growth rate K of $M(n)$ can be numerically determined by means of linear regression of $\log M(n)$ versus $\log n$. This allows for a clear distinction of a nonchaotic and a chaotic system as either $K = 0$ or $K = 1$.

To justify the 0-1 test for chaos, note that the function $p(n)$ can be viewed as a component of the solution to the skew product system

$$\begin{aligned} \theta(n+1) &= \theta(n) + c, \\ p(n+1) &= p(n) + \Phi(n)\cos\theta(n), \\ q(n+1) &= q(n) + \Phi(n)\sin\theta(n). \end{aligned} \tag{2}$$

Here (θ, p, q) represent coordinates on the Euclidean group $E(2)$ of rotations θ and translations (p, q) in the plane. It has been shown [13] that typically the dynamics on the group extension is sublinear and is (i) bounded if the underlying dynamics is nonchaotic and (ii) unbounded (but sublinear) if the underlying dynamics is chaotic. Moreover, the p and q components each behave asymptotically like Brownian motion on the line if the chaotic attractor is uniformly hyperbolic [8]. A nondegeneracy result [13] ensures that for chaotic dynamics the variance of the Brownian motion is nonzero for almost all choices of $c > 0$ and observable Φ. Recent work [12, 5] indicates that these statements remain valid for large classes of nonuniformly hyperbolic systems, such as Hénon-like attractors. Roughly speaking, to obtain $K = 1$ it suffices that the underlying system has a faster than quadratic rate of decay of correlations [2, 4, 9].

One might ask why it is not better to work with, instead of the $E(2)$-extension, the simpler $\mathbb{R}$-extension $p(n+1) = p(n) + \Phi(n)$, which can again be used to detect for chaos [9]. However, $p(n)$ grows linearly and it is necessary to subtract the linear term before computing $M(n)$.

Failure to do so produces the growth rate $K = 2$ regardless of whether the dynamics is regular or chaotic. The inclusion of the rotation θ kills off the linear growth.

In previous work [9] we illustrated the 0-1 test for chaos with numerically obtained data sets by simulating maps, ordinary differential equations, and partial differential equations, thereby showing the applicability to high-dimensional systems. Furthermore, we showed that a modified version of the test works very favorably for noisy systems when compared to traditional phase space reconstruction methods [10]. Here, we demonstrate the effectiveness of the modified test for experimental data.

To obtain experimental data, we constructed a bipolar motor, where a dipole magnet is suspended in a spatially uniform, linearly polarized oscillating magnetic field [3]. The motor consists of a short cylindrical ferrite permanent magnet located between a pair of Helmholtz coils so that it is free to rotate about an axis normal to the magnetic field of the coils, with the magnetic axis of the magnet perpendicular to the axis of rotation (see Figure 1). The magnetic moment of this magnet was determined to be 3.70 joule/tesla by measuring the oscillation of the magnet in a fixed magnetic field. A magnetic field varying sinusoidally with frequency ω and peak value 4.91×10^{-3} tesla was generated by driving the Helmholtz coils with a Pasco PI-9587C low-frequency digital function generator/amplifier combination. We use the frequency ω as a parameter. To measure the angular position ϑ of the dipole magnet, we digitally recorded with a 25 frames/second digital video camera a spot painted on an aluminium disc mounted coaxially at the top of the rotor axle.

This experiment is the realization of the forced ordinary differential equation $\ddot{\vartheta} + \gamma\dot{\vartheta} + (\mu B/I)\sin\vartheta\cos\omega t = 0$, where ϑ is the angle between the dipole moment μ and the magnetic field B, and I is the moment of inertia. Here, the damping is assumed to be proportional to the angular velocity with a constant rate γ. The observable Φ in the experiment corresponds to the function $\cos\vartheta$ (up to a constant phase).

Using $\Phi = \cos\vartheta$ as our observable, we took recordings for 9 minutes, which amounts to 13500 data points. We measured the autocorrelation of the data and found it to be oversampled; the e-folding time [11] was $\tau = 5$, and we therefore created a (shorter) time series consisting of 2700 data points. (We note that when the data is collected over a small interval of time, it is important that the data is not oversampled, since the mean square displacement scales as $M(n) = Vn + o(n)$, where $V = 0$ in the regular case and $V \ll 1$ in the oversampled chaotic case.) We found that our test also works with a data set recorded over 3 minutes, which amounts to $N = 900$ data points after taking every 5th data point.

In Figure 2, we show experimentally obtained phase plots exhibiting periodic and chaotic dynamics for different values of the frequency ω. One may question the need for our test as one can clearly distinguish chaotic from regular dynamics by looking at the phase plots in Figure 2. However, this is possible only for low-dimensional systems, and even then it is problematic for systems with quasiperiodic dynamics. We showed previously [10] that our test distinguishes between quasiperiodic and chaotic dynamics in high-dimensional systems. In this paper, it is our aim to show that the test works for real world problems which are inevitably contaminated by noise.

To illustrate the mathematics behind the test, we show in Figure 3 the bounded/Brownian-like dynamics of the translation components (p, q) of the $E(2)$ extension (2) corresponding to the periodic/chaotic dynamics in Figure 2. We made the arbitrary choice $c = 1.95$ in

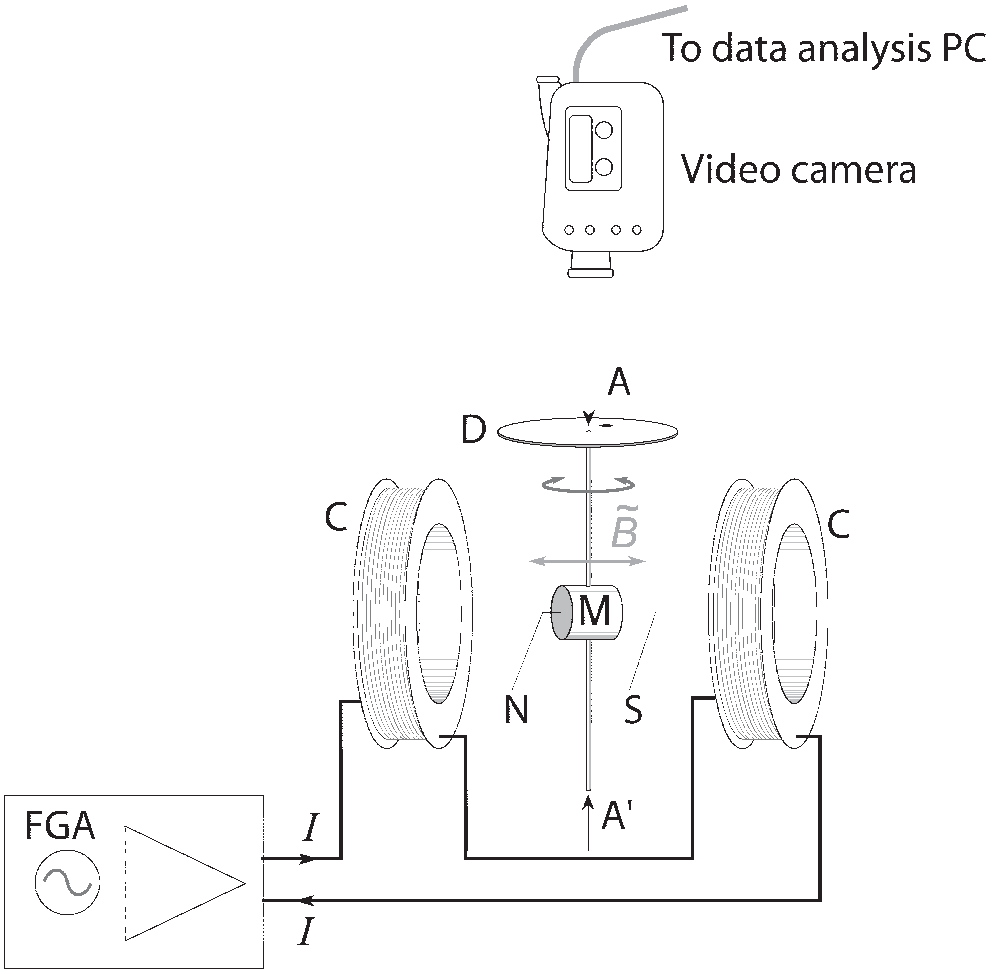


Figure 1. *Schematic diagram of the experimental setup of the bipolar motor. C: Helmholtz coil pair used to provide an oscillating magnetic field B. (For clarity, the coils are shown farther apart than they would be in the Helmholtz configuration, where the (mean) coil separation is set equal to the (mean) coil radius.) FGA: Pasco PI-9587C function generator/amplifier. M: Ferrite permanent magnet, where N and S indicate the orientation of its magnetic moment. D: Indicator disc for angular position of the ferrite magnet. The position of the black spot on this white-painted disk is recorded by the video camera. AA′: Axis of rotation of the ferrite magnet.*

computing the trajectories shown in Figure 3. To produce this figure we took the data set consisting of 2700 points.

In Figure 4, we show the mean square displacement corresponding to Figure 3 (again for $c = 1.95$ but now using the data set consisting of $N = 900$ points) clearly showing the bounded

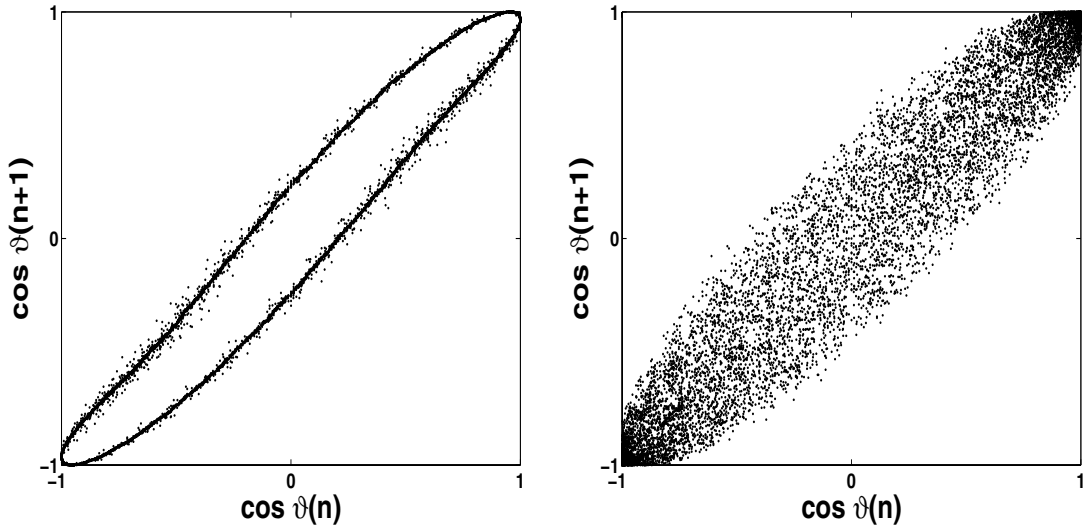


Figure 2. *Experimentally obtained phase plots consisting of* 13500 *data points for different values of the forcing frequency* ω*. The moment of inertia is* $I = 4.42 \times 10^{-4}\mathrm{kg\,m^2}$*, the magnetic moment is* $\mu = 3.70\mathrm{J/T}$*, the maximal magnetic field is* $B = 4.91 \times 10^{-3}\mathrm{T}$*, and the damping coefficient is* $\gamma = 0.046\mathrm{Hz}$*. The forcing frequencies are (left)* $\omega = 0.9\mathrm{Hz}$ *and (right)* $\omega = 0.6\mathrm{Hz}$*.*

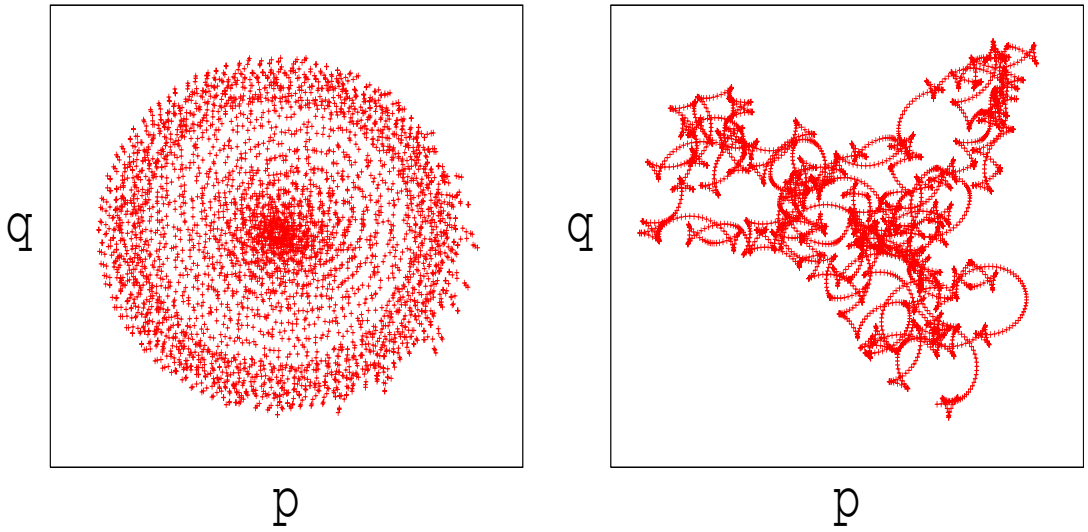


Figure 3. *The dynamics of the translation components* (p, q) *of the* $E(2)$ *extension. (Left) Bounded trajectories are shown corresponding to periodic dynamics at* $\omega = 0.9\mathrm{Hz}$*. (Right) Brownian-like trajectories are shown corresponding to chaotic dynamics at* $\omega = 0.6\mathrm{Hz}$*. Both cases show* 2700 *data points and are calculated for* $c = 1.95$*.*

behavior for regular dynamics and the linear growth for chaotic dynamics. In practice [10], we take $n = 10, 11, 12, \ldots, 90$ to ensure that $0 \ll n \ll N = 900$ in the definition of $M(n)$. The asymptotic growth rate K of the mean square displacement is determined by fitting a straight line to the curve $\log M(n)$ versus $\log n$ through minimizing the absolute deviation. This has

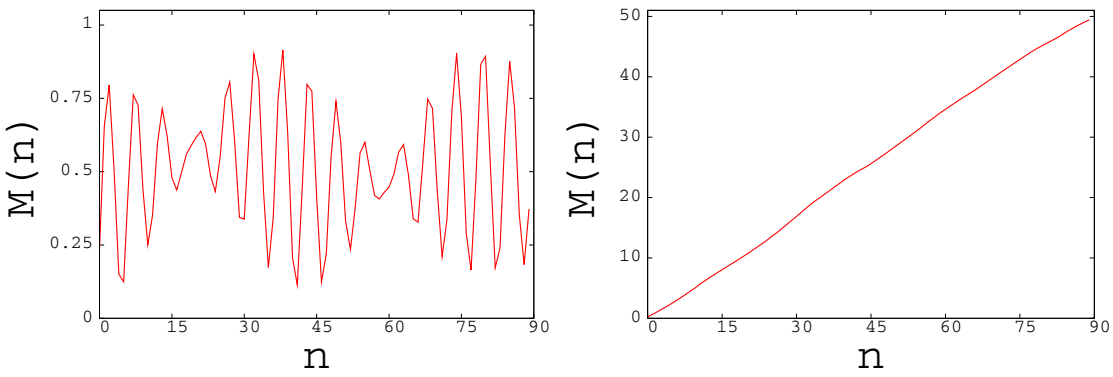


Figure 4. *Mean square displacement $M(n)$ as a function of n for 900 data points shown from $n = 0$ up to $n = 90$. (Left) Periodic case $\omega = 0.9$Hz corresponding to Figure 3 (left). (Right) Chaotic case $\omega = 0.6$Hz corresponding to Figure 3 (right). Both cases are calculated for $c = 1.95$.*

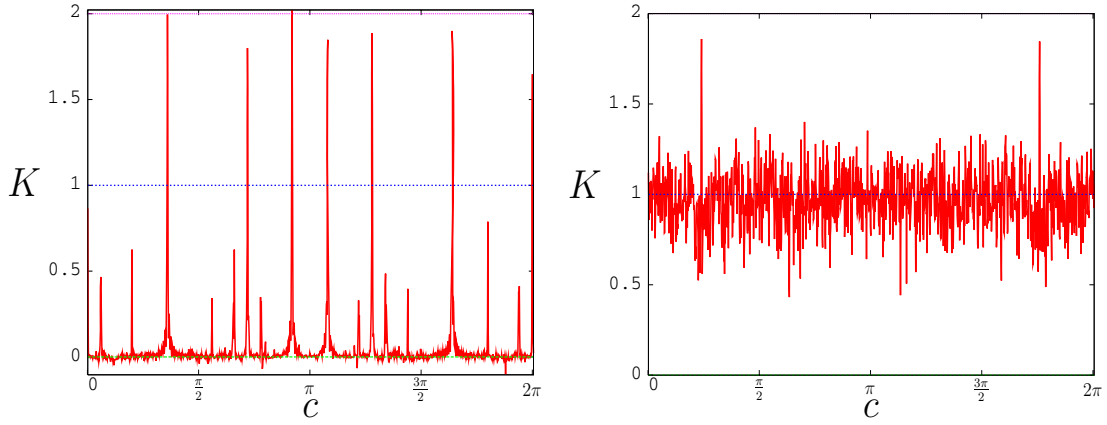


Figure 5. *Dependence of the asymptotic growth rate K on the frequency c. (Left) Regular dynamics at 0.9Hz and (right) chaotic dynamics at 0.6Hz.*

the advantage, compared to a least square fit, that outliers are weighted less [15], which is desirable as the linear behavior of $M(n)$ is valid only for $n \gg 1$ and can deviate strongly for small n.

Since the test is 2π-periodic in c, we choose c from $(0, 2\pi)$. In Figure 5, we plot K as a function of c where c is sampled uniformly with $\Delta c = 0.005$ between 0.005 and 6.28. The figure shows that the test is essentially independent of the specific choice of c and that we are justified in taking randomly chosen values of c. Only exceptional *resonant* values of c yield values of K which do not fit the picture of $K = 0$ for regular dynamics and $K = 1$ for chaotic dynamics. Resonances with a corresponding value of $K = 2$ occur if the frequency c is commensurate with a nonlinear frequency of the underlying dynamical system (particularly for values $c = 0$ and $c = 2\pi$ for data with nonzero mean). Figure 5 illustrates this resonance phenomenon. To avoid these resonances distorting the "true" value of K, we perform the test for several different values of c and then take the median value of K. We use the median rather than the mean, since the median gives less weight to the strong outliers stemming from

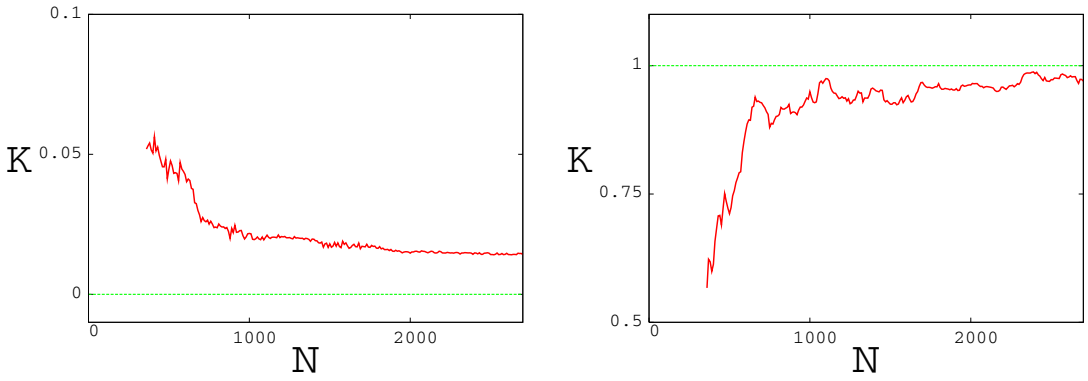


Figure 6. *Dependence of the asymptotic growth rate K on the length of the time series N. (Left) Periodic case $\omega = 0.9$Hz. (Right) Chaotic case $\omega = 0.6$Hz. We used the data set corresponding to the full 9 minutes of recording, taking every 5th data point, so the maximum value of N is 2700.*

resonances. In practice, 100 random choices of c suffice [10].

Using the experimental data and the method described above, we computed $K = 0.02$ for the case $\omega = 0.9$Hz and $K = 0.92$ for the case $\omega = 0.6$Hz, clearly indicating regular dynamics for the first parameter value and chaotic dynamics for the second parameter value.

In Figure 6, we show the dependence of the median value for K as a function of N for both cases. One can clearly see the improvement of the test with increasing N. Note that for the regular case in Figure 6 (left), K is not converging to 0. This is due to the noise present in our data. However, values of 0.02 are sufficiently close to 0 to allow for a binary distinction with chaotic values around of K which exceed 0.9. (See [10] for a discussion on data contaminated by measurement noise for our test.)

We have established a simple, inexpensive test for chaos and tested its efficiency with an analysis of experimental data obtained from a bipolar motor. The computational effort is of low cost, both in terms of programming efforts and in terms of actual computation time. This is a binary test that distinguishes purely between nonchaotic and chaotic dynamics by means of the quantity K taking values close to either 0 or 1. The most powerful aspects of our method are that there are no practical limitations arising from the nature and dimension of the deterministic data under consideration (except the intrinsic problem, mentioned earlier, that the time series needs to be sufficiently long that the dynamics explores enough of the attractor) and that the test is robust to contamination by noise.

Acknowledgment. The authors are grateful for discussions with Ramon Xulvi-Brunet.

REFERENCES

[1] H. D. I. Abarbanel, R. Brown, J. J. Sidorovich, and L. S. Tsimring, *The analysis of observed chaotic data in physical systems*, Rev. Modern Phys., 65 (1993), pp. 1331–1392.

[2] P. Ashwin, I. Melbourne, and M. Nicol, *Hypermeander of spirals; local bifurcations and statistical properties*, Phys. D, 156 (2001), pp. 364–382.

[3] M. J. Ballico, M. L. Sawley, and F. Skiff, *The bipolar motor: A simple demonstration of deterministic chaos*, Amer. J. Phys., 58 (1990), pp. 58–61.

[4] V. N Biktashev and A. V. Holden, *Deterministic Brownian motion in the hypermeander of spiral waves*, Phys. D, 116 (2001), pp. 342–382.
[5] H. Bruin, M. Holland, and I. Melbourne, *Subexponential decay of correlations for compact group extensions of nonuniformly expanding systems*, Ergodic Theory Dynam. Systems, 25 (2005), pp. 1719–1738.
[6] M. Casdagli, S. Eubank, J. D. Farmer, and J. Gibson, *State space reconstruction in the presence of noise*, Phys. D, 51 (1991), pp. 52–98.
[7] J.-P. Eckmann, S. O. Kamphurst, D. Ruelle, and S. Ciliberto, *Liapunov exponents from time series*, Phys. Rev. A (3), 34 (1986), pp. 4971–4979.
[8] M. Field, I. Melbourne, and A. Török, *Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions*, Ergodic Theory Dynam. Systems, 23 (2003), pp. 87–110.
[9] G. A. Gottwald and I. Melbourne, *A new test for chaos in deterministic systems*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 460 (2004), pp. 603–611.
[10] G. A. Gottwald and I. Melbourne, *Testing for chaos in deterministic systems with noise*, Phys. D, 212 (2005), pp. 100–110.
[11] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed., Cambridge University Press, Cambridge, UK, 2004.
[12] I. Melbourne and M. Nicol, *Statistical properties of endomorphisms and compact group extensions*, J. London Math. Soc. (2), 70 (2004), pp. 427–446.
[13] M. Nicol, I. Melbourne, and P. Ashwin, *Euclidean extensions of dynamical systems*, Nonlinearity, 14 (2001), pp. 275–300.
[14] T. S. Parker and L. O. Chua, *Practical Numerical Algorithms for Chaotic Systems*, Springer-Verlag, New York, 1989.
[15] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, UK, 1992.
[16] M. Sano and Y. Sawada, *Measurement of the Lyapunov spectrum from a chaotic time series*, Phys. Rev. Lett., 55 (1985), pp. 1082–1085.
[17] T. Schreiber and H. Kantz, *Noise in chaotic data: Diagnosis and treatment*, Chaos, 5 (1995), pp. 133–142.
[18] F. Takens, *Detecting strange attractors in turbulence*, in Dynamical Systems and Turbulence, Lecture Notes in Math. 898, Springer-Verlag, Berlin, 1981, pp. 366–381.
[19] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, *Determining Lyapunov exponents from a time series*, Phys. D, 16 (1985), pp. 285–317.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 2, pp. 403–440

Unfolding the Cusp-Cusp Bifurcation of Planar Endomorphisms*

Bernd Krauskopf†, Hinke M. Osinga†, and Bruce B. Peckham‡

Abstract. In many applications of practical interest, for example, in control theory, economics, electronics, and neural networks, the dynamics of the system under consideration can be modeled by an endomorphism, which is a discrete smooth map that does not have a uniquely defined inverse; one also speaks simply of a noninvertible map. In contrast to the better known case of a dynamical system given by a planar diffeomorphism, many questions concerning the possible dynamics and bifurcations of planar endomorphisms remain open. In this paper we make a contribution to the bifurcation theory of planar endomorphisms. Namely, we present the unfoldings of a codimension-two bifurcation, which we call the cusp-cusp bifurcation, that occurs generically in families of endomorphisms of the plane. The cusp-cusp bifurcation acts as an organizing center that involves the relevant codimension-one bifurcations. The central singularity is an interaction of two different types of cusps. First, an endomorphism typically folds the phase space along curves J_0 where the Jacobian of the map is zero. The image J_1 of J_0 may contain a cusp point, which persists under perturbation; the literature also speaks of a map of type $Z_1 < Z_3$. The second type of cusp occurs when a forward invariant curve W, such as a segment of an unstable manifold, crosses J_0 in a direction tangent to the zero eigenvector. Then the image of W will typically contain a cusp. This situation is of codimension one and generically leads to a loop in the unfolding. The central singularity that defines the cusp-cusp bifurcation is, hence, defined by a tangency of an invariant curve W with J_0 at the preimage of the cusp point on J_1. We study the bifurcations in the images of J_0 and the curve W in a neighborhood of the parameter space of the organizing center—where both images have a cusp at the same point in the phase space. To this end, we define a suitable notion of equivalence that distinguishes between the different possible local phase portraits of the invariant curve relative to the cusp on J_1. Our approach makes use of local singularity theory to derive and analyze completely a normal form of the cusp-cusp bifurcation. In total we find eight different two-parameter unfoldings of the central singularity. We illustrate how our results can be applied by showing the existence of a cusp-cusp bifurcation point in an adaptive control system. We are able to identify the associated two-parameter unfolding for this example and provide all the different phase portraits.

Key words. discrete-time system, noninvertible planar map, invariant curve, unstable manifold, codimension-two bifurcation

AMS subject classifications. 37D10, 37M20, 65P30

DOI. 10.1137/060672753

*Received by the editors October 18, 2006; accepted for publication by T. Kaper March 3, 2007; published electronically June 15, 2007.

http://www.siam.org/journals/siads/6-2/67275.html

†Bristol Centre for Applied Nonlinear Mathematics, Department of Engineering Mathematics, University of Bristol, Bristol BS8 1TR, UK (b.krauskopf@bristol.ac.uk, h.m.osinga@bristol.ac.uk). These authors were both supported by EPSRC Advanced Research Fellowship grants.

‡Department of Mathematics and Statistics, University of Minnesota Duluth, Duluth, MN 55812 (bpeckham@d.umn.edu). This author acknowledges support from the National Science Foundation (grant DMS-9973926) and the hospitality and support of the Bristol Centre for Applied Nonlinear Mathematics during his sabbatical in Fall 2004.

1. Introduction. Many situations of practical interest are modeled mathematically by a map on a suitable phase space. In this case time is thought to be discrete and the time evolution of an initial point is given by the iterates of the map. We consider here the case that the dynamics is generated by an *endomorphism*, that is, by a smooth map that does not have a (uniquely defined) inverse. As is common in the literature, we simply speak of a noninvertible map. Note that the theory of endomorphisms of the real line is well developed, with the logistic map being the most famous example. However, much less is known about the possible dynamics and bifurcations of endomorphisms on $\mathbb{R}^n$ for $n \geq 2$. We are concerned here with the case of a noninvertible planar map that maps $\mathbb{R}^2$ to itself. Such maps arise naturally as models in several areas of application, including control theory [1, 8, 9], economics [2, 3], radiophysics [23], and neural networks [34]. Noninvertibility easily occurs in applications that feed back sampled data using too large a sampling time. This is effectively equivalent to the case that a vector field model is integrated with too large an integration step [22].

In a region where a planar endomorphism has a well-defined branch of an inverse it may display all the dynamical complexity of a planar *diffeomorphism*, that is, of a smooth planar map with a smooth inverse. As is now well known, planar diffeomorphisms typically show complicated dynamics, including chaos; see, for example, [13, 30] as entry points to the extensive literature. Famous examples of planar diffeomorphisms are the Hénon map [16] and the Ikeda map [17], as well as Poincaré maps of periodically driven systems, such as the forced pendulum, Van der Pol, and Duffing oscillators [13, 38].

The main question in the study of planar endomorphisms is what extra dynamical features may occur beyond what is known for planar diffeomorphisms. While there have been quite a number of studies of specific planar endomorphisms, much less is known about the generic dynamics and bifurcations of planar endomorphisms that must be expected in a typical example. In short, there is, as yet, no systematic bifurcation theory for planar endomorphisms.

To be specific, consider a family of endomorphisms of the plane

$$\begin{aligned} f : \mathbb{R}^2 \times \mathbb{R}^m \to \quad & \mathbb{R}^2, \\ (x, \lambda) \quad \mapsto \quad & f(x, \lambda), \end{aligned} \tag{1.1}$$

where $\lambda \in \mathbb{R}^m$ is an m-dimensional parameter, and f is a smooth map. We consider the case that f is not a diffeomorphism, which means that the Jacobian Df of f has a nonzero kernel. We define the (nonempty) *singular locus*

$$J_0 := \ker(Df) = \left\{ x \in \mathbb{R}^2 \mid Df\,(x) \text{ is singular} \right\}. \tag{1.2}$$

The image $J_1 := f(J_0)$ is called the *critical locus*. In the literature the critical locus J_1 is also referred to as *critical curve* [29] or *Ligne Critique (LC)* [26], while J_0 is also referred to as the *curve of merging preimages* [29] or LC^{-1} [26]. Technically, $LC \subset J_1$ and $LC^{-1} \subset J_0$, but for this paper it is sufficient to consider them equal. This is justified by their properties, which follow immediately from the implicit function theorem.

Proposition 1.1 (generic properties of J_0 and J_1). *Generically, that is, for a generic family f and parameter λ in general position, the singular locus J_0 is a smooth curve where* $\dim(\ker(Df)) = 1$. *The eigenvector of the eigenvalue* 0 *is transverse to J_0 except at isolated points, called precusp points, where it is tangent to J_0. Therefore, the critical locus J_1 consists*

of smooth curve segments that meet at isolated cusp points, which are the images of the precusp points.

Generally speaking, noninvertibility gives rise to regions with different numbers of inverses. The critical locus J_1 divides the phase plane $\mathbb{R}^2$ into regions with a constant number of preimages. These regions are usually labeled by Z_k, where k is the number of preimages in that region [10, 26]. Generically, that is, except at the cusp points, the map f folds the phase plane along the smooth curve J_0, which is mapped to J_1. Therefore, the number of preimages differs locally by two on either side of J_1 [4]. As one moves from one region into the next by crossing a fold curve the number of preimages changes to $k \pm 2$.

In the interior of a region Z_k for $k \geq 1$ one can select a single branch of the inverse, so that the map f is locally a diffeomorphism. New phenomena, which do not occur for diffeomorphisms, may arise when a (forward) invariant object, such as a fixed point, periodic orbit, or invariant manifold, interacts with the boundary of the region of definition of the inverse branch. Therefore, the bifurcation theory of endomorphisms can be thought of as the study of the interaction of dynamical objects and their images with J_0 and J_1, respectively.

The literature on noninvertible planar maps consists largely of case studies in specific examples that reveal specific phenomena and a number of codimension-one bifurcations. A lot of attention has been devoted to the structure of basins of attraction. Because a basin may consist of disconnected regions or be multiply connected, it is often referred to as a "sea" possibly with "islands" in it, or as "land" with "lakes" [19, 26]; connectivity properties are characterized with an "island number" or "lake number" [26]. However, as was shown in [6, 7], changes to the basin boundary are due to two types of tangencies, called inner and outer tangencies, of a stable set (the generalization of the stable manifold of a saddle point) with J_1.

Another topic that received a lot of interest as typical for noninvertible maps are self-intersections and associated "loops" of (forward) invariant sets; see [10, 11, 18, 24, 27], and [22], where loops are referred to as "antennae." The main interest is in bifurcations leading to the destruction of an invariant curve (also called "IC" or "torus"), which is the closure of the unstable manifolds of a suitable periodic orbit if the dynamics on the curve is phase locked. To define an unstable manifold of an endomorphism consider a generic saddle point p of f (or a suitable iterate of f). Genericity of p means in particular that $p \notin J_0$. Therefore, there exists the local unstable manifold $W^u_{loc}(p)$ associated with the unique inverse branch that fixes p. The *global unstable manifold* $W^u(\mathbf{x}_0)$ can then be defined as

$$W^u(p) = \bigcup_{n=1}^{\infty} f^n(W^u_{loc}(p)). \tag{1.3}$$

Note that forward images under f are unique, so that $W^u(p)$ is indeed forward invariant. Furthermore, $W^u(p)$ is generically an immersed manifold; see, for example, [37]. Even though $W^u(p)$ may have structurally stable transverse self-intersections, it is justified to speak of $W^u(p)$ as the global unstable manifold. In particular, $W^u(p)$ can be computed numerically by any algorithm that works for diffeomorphisms; see also [6, 7]. Indeed, depending on how an unstable manifold or other invariant curve crosses J_0, a loop in its image may be the result. The codimension-one bifurcation that creates a small loop was analyzed in [10, 11, 24];

we refer to it as the *loop-creation bifurcation* and it is discussed in more detail in section 2. Due to forward invariance of $W^u(p)$, typically an infinite number of loops is created in the loop-creation bifurcation, which may give rise to a rather spectacular loss of smoothness of an invariant curve. In combination with other mechanisms (as known for diffeomorphisms) of the break-up of tori, one finds phenomena such as the "appearance of loops on an unstable manifold, and the reappearance of an attractor, this time chaotic with loops" [11, p. 107]; this type of attractor has been called a "weakly chaotic ring" [10, 27].

In this paper we make a contribution to the bifurcation theory of planar endomorphisms by providing the unfolding of a codimension-two bifurcation, which we call the *cusp-cusp bifurcation*. Our study was motivated by the properties of a well-referenced example of a noninvertible system, namely, the discrete-time adaptive control system discussed in [1, 8, 9]. More specifically, in [10] a loop of the unstable manifold of a saddle point is found that surrounds a cusp point C_1 on J_1. Several other codimension-one bifurcations are found near this situation, including a loop-creation bifurcation and the passing of the invariant curve through the cusp point. These bifurcations and the generic phase portraits near them are illustrated in [10] near the cusp point C_1 on J_1 as well as near the precusp point C_0 on J_0, where $f(C_0) = C_1$. We identify the cusp-cusp bifurcation as the organizing center of the observed dynamics in the sense that the known codimension-one bifurcations occur in its unfolding.

Specifically, at a cusp-cusp bifurcation there is a quadratic tangency of an invariant curve W with the singular locus J_0 exactly at the precusp point C_0. This situation corresponds to a cusp of $f(W)$ exactly at the cusp C_1 on J_1—hence the name of this bifurcation; see Figure 4(a) and Figure 20(a) and (b). The situation is of codimension two because one parameter is needed to ensure a tangency of W with J_0, while another parameter must be adjusted to ensure that this tangency occurs at C_0. Recall that the existence of an isolated cusp point on J_1 is a generic property, that is, a cusp point is stable under small parameter variations. Therefore, we assume that the unfolding parameters of the cusp-cusp bifurcation do not lead to bifurcations of the cusp point (such as the disappearance of the cusp point in a swallowtail bifurcation).

We derive a normal form of the cusp-cusp bifurcation by considering the interaction of a parabola (representing W locally) under the action of a normal-form map given by projections via the graph of the function f in $\mathbb{R}^4$ near a generic cusp point. This surface can best be viewed in a projection of $\mathbb{R}^4$ onto $\mathbb{R}^3$, where it takes the form of a generic two-dimensional surface with a cusp as known from singularity theory [4]; see Figure 3. In this setup both the cusp point C_1 and the precusp point C_0 are at the origin and the curve J_0 is a fixed parabola with C_0 as its maximum. The unfolding or primary parameters in the normal form are the vertical and horizontal positions of the maximum or minimum of the parabola representing W, while its "steepness" (relative to the curve J_0) plays the role of a higher-order or secondary parameter in the normal form. Depending on its value we distinguish eight different kinds of cusp-cusp bifurcations with associated two-parameter unfoldings. All unfoldings (cases I to VIII) are presented as two-parameter diagrams with associated representative phase portraits, that is, configurations of W near J_0 and $f(W)$ near J_1. We also show that the phase portraits found in [10] are organized by a cusp-cusp bifurcation of type VII.

Singularity theory is the main tool in our analysis of the cusp-cusp bifurcation. It has

proven its use in the bifurcation analysis of invertible maps and vector fields, where classic singularities occur in the product of phase space and parameter space; see [12] and also [5] for a recent example. The question in the present setting is how a noninvertible planar map is folding the phase plane over itself. In other words, one needs to study projections via a surface, namely, the smooth image of the domain, in a higher-dimensional space. This point of view has been adopted before, for example, in [39]. The action of a noninvertible planar map near any of the fold curves that make up J_1 can be understood by considering projections via a surface in $\mathbb{R}^3$ with a generic fold line [14, 15]. Planar quadratic maps with bounded critical sets are considered in [28]. The papers [31, 33] study connections between the complex quadratic family and more general two-parameter families of maps of the plane $\mathbb{R}^2$. In [7] projections via a folded surface were used to identify two types of codimension-one tangency bifurcations of a stable set with J_1 that are responsible for rearrangements of the basins of attractions of a noninvertible map. The study presented here is very much in the same spirit. Note that near folds and cusps, the situations considered here, the surface via which one projects can be embedded in $\mathbb{R}^3$, rather than in $\mathbb{R}^4$, which greatly helps with visualizing the local map.

Another important ingredient of our study is a suitable notion of equivalence that allows one to define codimension-one bifurcations in this context. We consider here equivalence as given by the position of the oriented invariant curve relative to J_1; see Definition 3.4.

The paper is organized as follows. In section 2 we discuss some mathematical background, including the normal form of a smooth map near a cusp point. In section 3 we introduce the normal-form setting of the cusp-cusp bifurcation, as well as our notion of equivalence of invariant curves relative to J_1. In section 4 we discuss the codimension-one bifurcations that occur near a cusp-cusp bifurcation, and in section 5 we present all cases of two-parameter unfoldings. The adaptive control system from [10] is studied in section 6 to show how our results manifest themselves in practice. We conclude in section 7, where we also point out directions for future research. The numerical methods that we use in section 6 are discussed in Appendix A.

2. Mathematical setting and cusp normal-form map.

In order to understand the folding properties of f we consider the line field $\mathcal{E}$ associated with J_0, defined by

$$\mathcal{E} = \{l(e_0(x), x) \mid x \in J_0\}, \tag{2.1}$$

where $l(v, p)$ is the line through the point p given by the vector v, and $e_0(x)$ is the eigenvector of the eigenvalue 0 of $Df(x)$. As stated in Proposition 1.1, every line $\mathcal{E}_x := l(e_0(x), x)$ is transverse to J_0 at generic $x \in J_0$.

Now consider a curve W crossing J_0 transversely at a generic point x_0 in the sense of Proposition 1.1. If W is not tangent to $\mathcal{E}_{x_0}$ then the image $f(W)$ under f is quadratically tangent to J_1 at the point $f(x_0)$. This tangency at $f(x_0)$ is structurally stable, because it corresponds to the projection of a curve that crosses a fold [4]. The exceptional case is given by the line field $\mathcal{E}$. Any curve W that crosses J_0 tangent to $\mathcal{E}$ has its tangent vector mapped to the zero vector by $Df(x_0)$. Then $f(W)$ will have a cusp at $f(x_0)$, unless the two branches of W on either side of J_0 happen to map exactly on top of each other (which is not generic). The "direction" of the cusp, that is, the limiting direction of the induced tangent vector to $f(W)$ as the cusp point $f(x_0)$ is approached, is determined by the curvature of W.

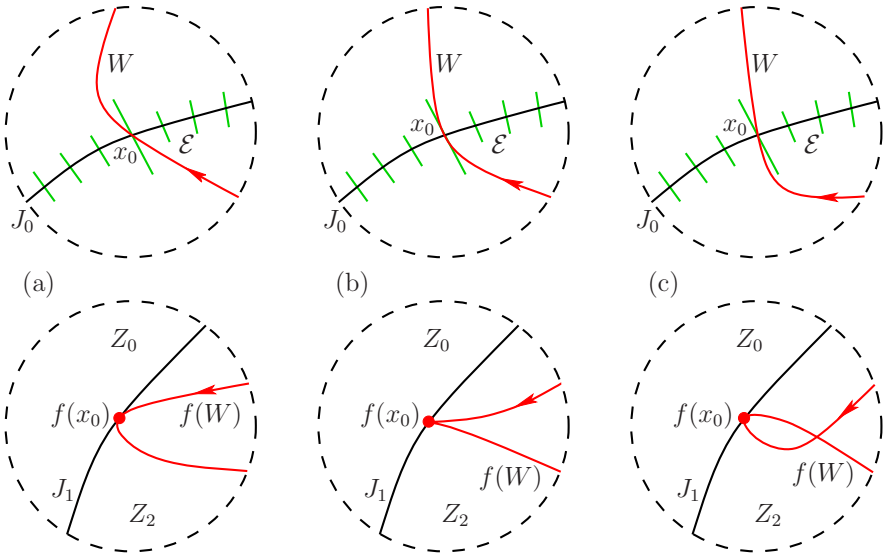


Figure 1. *Sketches of the situation before* (a)*, at* (b)*, and after* (c) *a codimension-one loop creation, where in the top row the invariant curve W (red) becomes tangent at x_0 to the line field $\mathcal{E}$ (green lines). The bottom row shows the respective images of the top row and illustrates the resulting creation of a little loop of $f(W)$ (red curve).*

To illustrate the concept we present in Figure 1 the one-parameter unfolding of the loop-creation bifurcation; the top row shows a neighborhood of a transverse intersection point x_0 of a forward invariant curve W with J_0, and the bottom row a neighborhood around $f(x_0)$ on J_1. Note that both W and $f(W)$ have a direction, indicated by arrows, which needs to be specified to understand the action of the map. Generically, the tangent $dW(x_0)$ of W at x_0 is transverse to the line $\mathcal{E}_{x_0}$, which means that $f(W)$ has a quadratic tangency with J_1; see Figure 1(a) and (c). However, when $l(dW(x_0), x_0) = \mathcal{E}_{x_0}$ then $f(W)$ has a cusp at $f(x_0) \in J_1$; see Figure 1(b). In the unfolding of this codimension-one situation a small loop of $f(W)$ is created (or destroyed) near $f(x_0)$, which explains the name of this bifurcation.

2.1. Characterization of a cusp of J_1. We now consider the basic setting of this paper, namely, an isolated generic cusp point C_1 on J_1 and, thus, an isolated precusp point C_0 on J_0. The situation is sketched in Figure 2(a) and (b) in neighborhoods U and V around C_0 and C_1, respectively. Note that J_1 has a second preimage, denoted $\widehat{J}_0 \subset f^{-1}(J_1)$, in the neighborhood U, which is tangent to J_0 at the precusp point C_0. The three regions below J_0 and in between $\widehat{J}_0$ and J_0 get mapped to the single region with three preimages, which is the region under the cusp point C_1 and bounded by J_1. This situation is referred to in the literature as a map of type $Z_1 < Z_3$. We remark that any map of type $Z_k < Z_{k+2}$ for $k \geq 1$ has the same properties locally near the cusp point, but with $(k-1)$ additional sheets of inverses that play no role in the local unfolding.

Figure 2 also shows the line field $\mathcal{E}$ near J_0. Notice that the line $\mathcal{E}_{C_0}$ is not transverse to J_0. Therefore, the defining property of the cusp-cusp bifurcation, namely, a quadratic

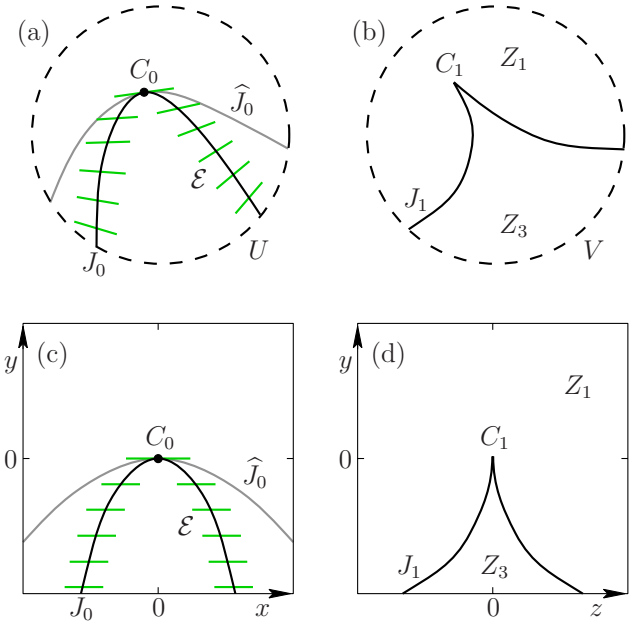


Figure 2. *The line field $\mathcal{E}$ (green lines) at J_0 (black curve) along with the second preimage $\widehat{J}_0$ (gray curve) in a neighborhood U of the precusp point C_0* (a)*, and a neighborhood V of the cusp point C_1 on the image J_1 (black curve) of J_0 and $\widehat{J}_0$* (b)*. Panels* (c) *and* (d) *show the situation for the normal-form map* (2.2)*, where J_0 and $\widehat{J}_0$ are parabolas and the leaves of $\mathcal{E}$ are given by $\{y = const\}$.*

tangency of an invariant curve W with J_0 at C_0, can be interpreted as defining a degenerate loop bifurcation.

2.2. Cusp normal-form map. As is known from singularity theory [4], a generic smooth planar map with a cusp point can locally near the precusp point be brought to the normal form

$$(x, y) \mapsto (u, v) = (axy - bx^3, y)$$

by a smooth change of variables, for any positive nonzero constants a and b. We make a convenient choice; namely, we consider the normal-form map defined as

$$F : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} -x^3 - 3xy \\ y \end{pmatrix}. \tag{2.2}$$

The (x, y)-plane corresponds to a local neighborhood of C_0 and the (z, y)-plane to a local neighborhood of C_1. Both C_0 and C_1 are located at the origin. The exact form of (2.2) was chosen such that the Jacobian

$$DF \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} -3x^2 - 3y & -3x \\ 0 & 1 \end{bmatrix}$$

is singular along the particularly simple critical curve

$$J_0 := \{y = -x^2\}. \tag{2.3}$$

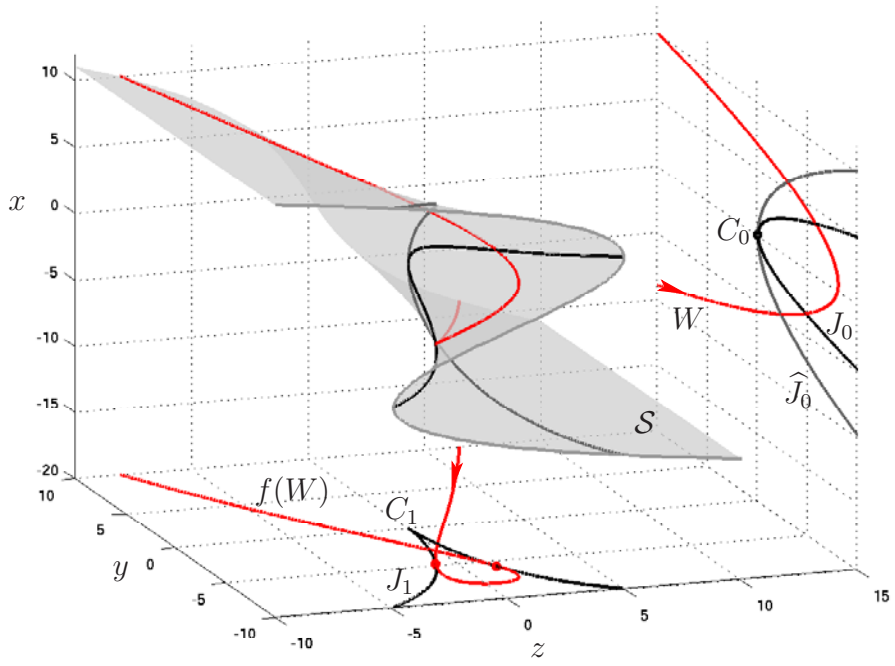


Figure 3. *Visualization in (x, y, z)-space of how the normal-form map (2.2) can be interpreted as a projection via the cusp surface $\mathcal{S}$ (light gray). The curve W (red) in the (x, y)-plane is mapped to the curve $F(W)$ (red) in the (z, y)-plane; the example is phase portrait* 31 *of Figure* 9.

The image of the parabola J_0 is the standard cusp

$$J_1 := \{z = \pm 2(\sqrt{-y})^3 \mid y \leq 0\}. \tag{2.4}$$

A straightforward calculation shows that it has the second preimage

$$\widehat{J}_0 := \{y = -\tfrac{1}{4}x^2\}. \tag{2.5}$$

The line field $\mathcal{E}$ does not depend on x and consists simply of the lines $l((1, 0)^T, x)$.

The situation for the normal-form map F is shown in Figure 2(c) and (d) in neighborhoods around C_0 and C_1, respectively. Note that the coordinate change that transforms a generic map f near a cusp point into its normal form F deforms the curves J_0 and $\widehat{J}_0$ to parabolas and the curve J_1 to a standard cusp. Furthermore, it "straightens out" the line field $\mathcal{E}$ to horizontal lines.

Figure 3 shows how the action of the map F can be interpreted geometrically as a projection via the cusp surface $\mathcal{S}$ in (x, y, z)-space given by $-x^3 - 3xy - z = 0$. Note that $\mathcal{S}$ is a graph over the (x, y)-plane but not over the (z, y)-plane. Any point (x_0, y_0) in the (x, y)-plane corresponds to a unique point on $\mathcal{S}$ under projection in the z-direction. It is then mapped to the unique point $F(x_0, y_0)$ under projection in the x-direction. Conversely, a point (z_0, y_0) in the (z, y)-plane lifts to a single point on $\mathcal{S}$ in the region of unique preimages, to two points

on $\mathcal{S}$ if $(z_0, y_0) \in J_1$, and to three points on $\mathcal{S}$ in the region of three preimages. Notice also the (projection of) the curves J_0 and $\widehat{J}_0$ on $\mathcal{S}$.

Figure 3 also shows how a parabola W in the (x, y)-plane "around" the precusp point C_0 maps under F. The result is a curve $F(W)$ in the (z, y)-plane with a self-intersection and two tangencies with J_1, one at either side of the cusp point C_1. This can be understood by considering the projection of the parabola onto the cusp surface $\mathcal{S}$, and then down to the (z, y)-plane.

3. Normal-form setting of the cusp-cusp bifurcation. At the cusp-cusp bifurcation there is a tangency of an invariant curve W with the curve J_0 at a precusp point C_0. As a consequence, $f(W)$ has a cusp exactly at the cusp point C_1 on J_1. This codimension-two bifurcation can be classified as a global bifurcation, because it involves an invariant curve. Nevertheless, we can consider a normal-form setting in the small neighborhoods U and V of C_0 and C_1, respectively, by replacing f near C_0 with the normal-form map F as given by (2.2). Then the curves J_0 and $\widehat{J}_0$ are the parabolas given by (2.3) and (2.5) and the points C_0 and C_1 are at the origin of the (x, y)- and (z, y)-planes, respectively. Furthermore, the tangency of W with J_0 at C_0 is generically quadratic. As a consequence, in any unfolding the intersections of W with J_0 and $\widehat{J}_0$ in a sufficiently small neighborhood U are determined entirely by the quadratic nature of the curve W. Therefore, from now on we consider the *quadratic normal form for* W in the (x, y)-plane given by

$$W := \{y = \gamma(x - a)^2 + b\}. \tag{3.1}$$

Here the parameters $a, b \in \mathbb{R}$ are the primary unfolding parameters and $\gamma \in \mathbb{R}$ is a higher-order or secondary parameter that determines the "steepness" of W relative to the fixed parabolas J_0 and $\widehat{J}_0$. The invariant curve in the neighborhood V of C_1 is simply given by $F(W)$. Note that the maximum (for $\gamma < 0$) or minimum (for $\gamma > 0$) of W lies at $(x, y) = (a, b)$; the cusp-cusp bifurcation occurs for $a = b = 0$.

The normal-form setting of the cusp-cusp bifurcation is now given as the study of all possible different configurations of $F(W)$ relative to J_1 and C_1 as a function of the unfolding parameters a and b, and for fixed generic values of the higher-order term γ. These different configurations correspond one-to-one to different types of intersections of W with J_0 and $\widehat{J}_0$; see Definition 3.4.

Throughout this paper we use the convention that the orientation of W is from left to right, that is, in the direction of increasing x. The orientation of $F(W)$ is induced by this convention. Notice that the map F reverses the orientation of the invariant curve in Figure 3, which is illustrated by the arrows of W and $F(W)$. We have the following general result.

Proposition 3.1 (orientation of $F(W)$). *For any a and b the image $F(W)$ of an (oriented) parabola W of the form (3.1) is oriented from left to right if $\gamma < -\frac{1}{3}$ and from right to left for $\gamma > -\frac{1}{3}$; see also Figure 4(c)–(e).*

Proof. The curve

$$R_0 := \left\{y = -\frac{1}{3}x^2\right\} \tag{3.2}$$

in the (x, y)-plane is mapped under the normal-form map F in a two-to-one fashion to the line segment $\{z = 0 \text{ and } y \leq 0\}$ in the (z, y)-plane. The region above R_0 is mapped to the

(z, y)-plane in an orientation reversing way, that is, positive x are mapped to negative z. By contrast, the region below R_0 is mapped to the (z, y)-plane in an orientation preserving way. The result follows, since the segments for sufficiently large $|x|$ of a parabola of the form (3.1) both lie in either one of the two regions. ∎

3.1. The cusp-cusp singularity. The nature of the cusp-cusp singularity for $(a, b) = (0, 0)$ depends on the higher-order term γ of (3.1). For $(a, b) = (0, 0)$ and for generic γ the cusp point C_1 is the only intersection point of $F(W)$ and J_1. The exception is that $F(W) \equiv J_1$, which happens when either $W = J_0$ or $W = \widehat{J}_0$. We further distinguish different cases of the cusp-cusp singularity depending on

- the orientation of $F(W)$,
- whether the curve $F(W)$ is below or above J_1, and
- whether or not the cusp of $F(W)$ at C_1 points upward or downward.

Proposition 3.2 (cusp-cusp singularity). *There are five intervals on the γ-line of generic cusp-cusp singularities for $(a, b) = (0, 0)$ in (3.1). The boundary points between these intervals are given by $\gamma = -1$, $\gamma = -\frac{1}{3}$, $\gamma = -\frac{1}{4}$, and $\gamma = 0$.*

Proof. According to Proposition 3.1 the orientation of $F(W)$ changes when W crosses R_0, which gives rise to the boundary point $\gamma = -\frac{1}{3}$. Furthermore, for $W = J_0$ or $W = \widehat{J}_0$ we have that $F(W) = J_1$, which gives rise to the boundary points $\gamma = -1$ and $\gamma = -\frac{1}{4}$, respectively. For W above $\widehat{J}_0$, that is, for $\gamma > -\frac{1}{4}$, the curve $F(W)$ lies above J_1. It follows immediately from (2.2) that the cusp of $F(W)$ is approached from the direction of negative y for any $\gamma < 0$ and from the direction of positive y for $\gamma > 0$. The situation for $\gamma = 0$ is degenerate in that $F(W)$ does not have a cusp at all. ∎

Proposition 3.2 is illustrated in Figure 4, where panels (a), (c), (e), (g), and (i) show the generic cases for $(a, b) = (0, 0)$. Note that panels (a) and (c) do not differ either in position or in orientation in the (z, y)-plane. The difference is that the corresponding curve lies on different sheets of the cusp surface $\mathcal{S}$ (see also Figure 3), and, as we will see in section 5, this leads to different two-parameter unfoldings. Figure 4(b), (d), (f), and (h) show the degenerate situations at the boundary points, namely, the situations when $W = J_0$, $W = R_0$, $W = \widehat{J}_0$, and $W \equiv 0$, respectively.

3.2. Notion of equivalence. In order to speak of typical or generic situations and their bifurcations we now define what we mean by equivalence. Our definition is in the spirit of [33], where a notion of equivalence was defined with respect to certain bifurcation phenomena. Specifically, we formalize the approach from the literature of distinguishing between different configurations of an invariant curve relative to J_1 and C_1. Namely, we consider a curve $f(W)$ in a neighborhood of C_1, which is the image of a curve W in a neighborhood of the associated precusp point C_0. For notational convenience we give the definition in the context of the normal-form setting, which has the advantage of leaving J_0 and J_1 fixed. However, it can be extended in a straightforward manner to define equivalence of two general invariant curves relative to J_1 in two respective neighborhoods of a cusp point and the corresponding precusp point.

Definition 3.3 (generic event). *Consider the image $F(W)$ of the (oriented) parabola W given by (3.1) under the normal-form map F. A* generic event *is*

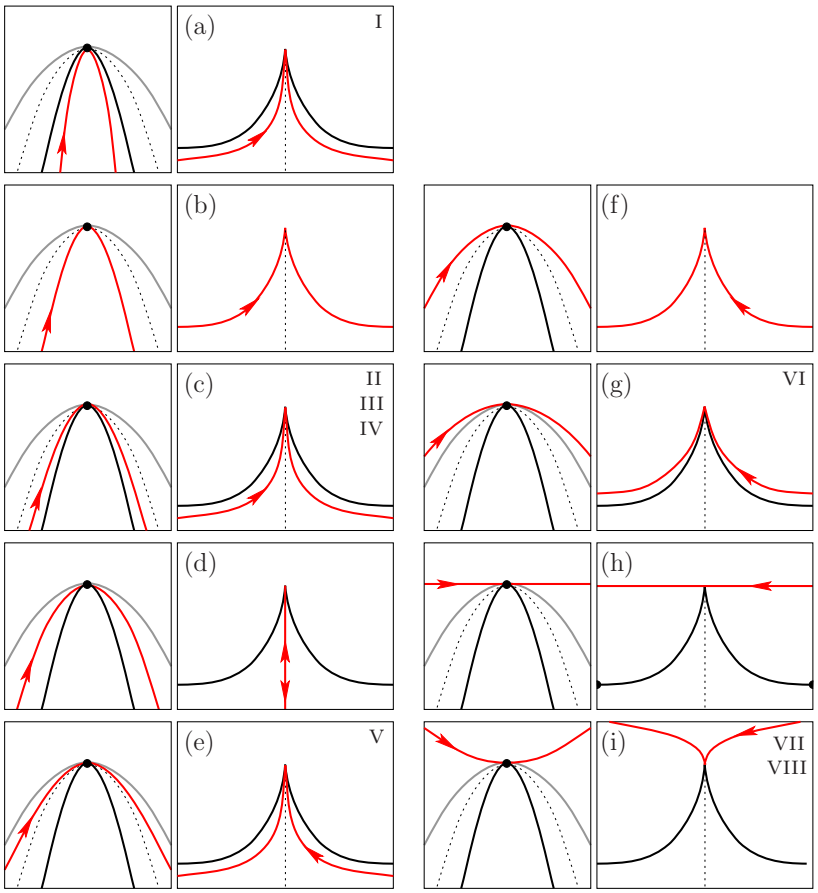


Figure 4. *Sketches of all generic and nongeneric cases of the cusp-cusp singularity; see Proposition 3.2. The left-hand panels show W (red curve) near the precusp point C_0 in the (x, y)-plane, and the right-hand panels show $f(W)$ (red curve) in relation to J_1 in the (z, y)-plane. The curves J_0 and J_1 are black, $\widehat{J}_0$ is gray, and R_0 is dashed and double-covers the dashed straight line; the arrows show the direction of parametrization by x. Shown are $\gamma < -1$ (a), $\gamma = -1$ (b), $-1 < \gamma < -\frac{1}{3}$ (c), $\gamma = -\frac{1}{3}$ (d), $-\frac{1}{3} < \gamma < -\frac{1}{4}$ (e), $\gamma = -\frac{1}{4}$ (f), $-\frac{1}{4} < \gamma < 0$ (g), $\gamma = 0$ (h), and $0 < \gamma$ (i). For the generic cases in panels (a), (c), (e), (g), and (i) the associated two-parameter unfoldings of Figure 8 are indicated.*

- *a quadratic tangency of $F(W)$ with J_1 at a generic point on J_1 (in a neighborhood of which J_1 is locally a smooth curve),*
- *a transverse intersection of $F(W)$ with J_1 at a generic point on J_1, or*
- *an isolated self-intersection of $F(W)$.*

In the first two cases we say that two such generic events are of the same type *if the respective tangency or intersection points*

1. *both lie on the same side of the cusp point C_1 on J_1, that is, both occur for $z < 0$ or both for $z > 0$, and*
2. *they are passed by $F(W)$ (locally) in the same direction, that is, both toward increasing z or both toward decreasing z.*

We are now able to define equivalence based on the notion of generic events. Note that a cusp on W is not a generic event.

Definition 3.4 (equivalence of an invariant curve relative to J_1). *Consider two images $F(W)$ and $F(\widetilde{W})$ of two (oriented) parabolas W and $\widetilde{W}$ given by (3.1) under the normal-form map F. We say that the curves $F(W)$ and $F(\widetilde{W})$ are equivalent with respect to J_1 if they*

1. *have the same induced orientation, and*
2. *both encounter the same types of generic events in the same order.*

The task for the remainder of this paper is to find all equivalence classes of phase portraits according to this notion of equivalence. Specifically, we need to find all generic two-parameter unfoldings in the (a,b)-plane for different choices of the higher-order term γ in (3.1). We start with a straightforward result on the possible number of events.

Proposition 3.5 (number of events). *In the normal-form setting of the cusp-cusp bifurcation, that is, for W given by (3.1) and any a, b, and generic γ, that is, for $\gamma \notin \{-1, -\frac{2}{3}, -\frac{1}{2}, -\frac{1}{3}, -\frac{1}{4}, 0, 1\}$, there are at most*

- *two (quadratic) tangencies of $F(W)$ with J_1,*
- *two transverse intersections of $F(W)$ with J_1, and*
- *one self-intersection of $F(W)$.*

Proof. The curves W, J_0, and $\widehat{J_0}$ are parabolas and graphs over the x-axis. Therefore, for any a, b, and γ, W can have at most two intersections with J_0 or $\widehat{J_0}$, which limits the respective number of tangencies and intersections of $F(W)$ with J_1 to two as well.

Any self-intersection of $F(W)$ must lie in the region under J_1, that is, in the z-interval between the two (if they exist) unique folds of $F(W)$ with respect to z. Furthermore, a self-intersection is due to a pair of points on W with the same y-value. According to (3.1) and (2.2) the curve $F(W)$ covers its y-range in a two-to-one fashion with a single maximum or a single minimum (depending on the sign of γ). Since $F(W)$ has either zero or two folds with respect to z, there can be at most one such pair of points with the same y-value, namely, exactly when the single maximum/minimum of $F(W)$ lies in between the two fold points. ∎

4. Codimension-one bifurcations.

Definition 3.4 gives rise to codimension-one bifurcations that correspond to a transition between equivalence classes of equivalence of the curve $F(W)$ relative to J_1. There are five basic codimension-one bifurcations where the change is local; that is, it occurs in a small neighborhood of the bifurcation point in phase cross parameter space. Furthermore, we find bifurcations at infinity and a transition due to the degeneracy of the parametrization (3.1).

4.1. The basic codimension-one bifurcations.

A great advantage of the normal-form setting is that it is possible to compute the loci of all these bifurcations explicitly.

Proposition 4.1 (basic codimension-one bifurcations). *In the normal-form setting of the cusp-cusp bifurcation there are exactly five codimension-one bifurcations for $(a,b) \neq (0,0)$ in (3.1) that correspond to a local change of events as defined in Definition 3.4. They are illustrated in Figures 5 and 6.*

1. *The* cusp transition*, denoted by C, where the curve W passes through J_0 at the precusp point C_0, which means that $F(W)$ passes exactly through the cusp point C_1 on J_1; see Figure 5(a). The locus of this bifurcation in the (a,b)-plane is the parabola*

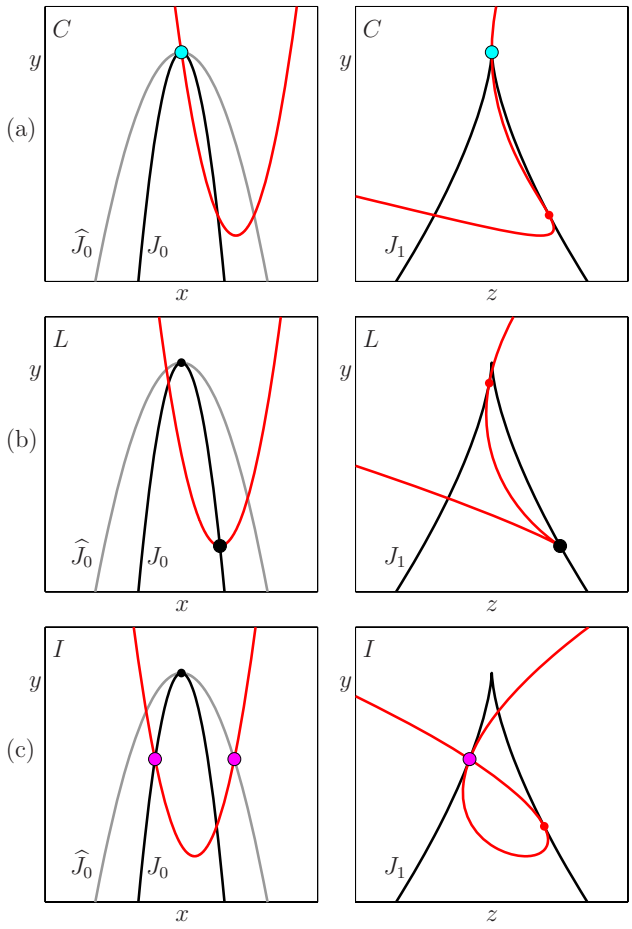


Figure 5. *Arrangement of W and $F(W)$ (red curves) at the cusp-transition bifurcation C* (a)*, the loop-creation bifurcation L* (b)*, and the intersection-at-tangency bifurcation I* (c)*. As indicated, these three bifurcations are assigned colors cyan, black, and magenta, respectively.*

$$b = c_C(\gamma)\, a^2 = -\gamma\, a^2. \tag{4.1}$$

2. *The* loop-creation bifurcation*, denoted by L, where W crosses J_0 tangent to the (horizontal) line field $\mathcal{E}$; see Figure* 5(b)*. The locus of this bifurcation in the (a, b)-plane is the parabola*

$$b = c_L(\gamma)\, a^2 = -a^2. \tag{4.2}$$

3. *The* intersection-at-tangency bifurcation*, denoted by I, where $F(W)$ self-intersects at a tangency point with J_1; see Figure* 5(c)*. The locus of this bifurcation in the (a, b)-plane is the parabola*

$$b = c_I(\gamma)\, a^2 = -(9\gamma + 4)\, a^2. \tag{4.3}$$

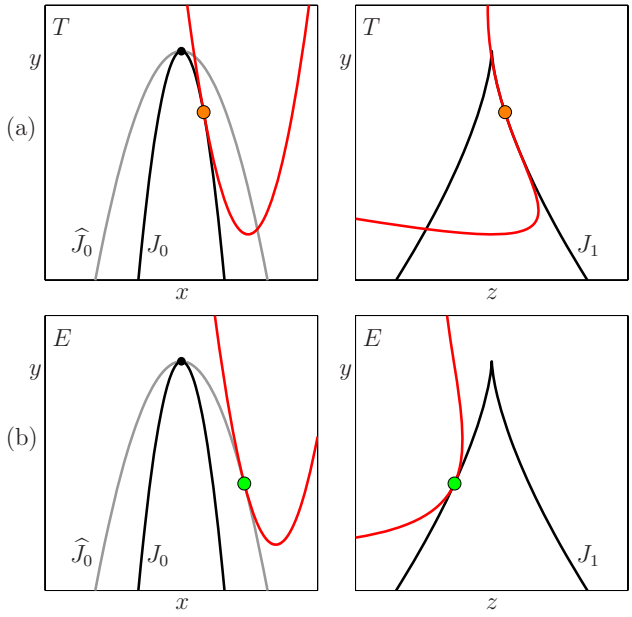


Figure 6. *Arrangement of W and $F(W)$ (red curves) at the tangency-creation bifurcation T* (a)*, and the enter-exit bifurcation E* (b)*. As indicated, these two bifurcations are assigned colors orange and green, respectively.*

4. *The* tangency-creation bifurcation*, denoted by T, where W is tangent to J_0; see Figure* 6(a)*. The locus of this bifurcation in the (a, b)-plane is the parabola*

$$b = c_T(\gamma)\, a^2 = -\frac{\gamma}{1+\gamma}\, a^2. \tag{4.4}$$

5. *The* enter-exit bifurcation*, denoted by E, where W is tangent to $\widehat{J}_0$; see Figure* 6(b)*. The locus of this bifurcation in the (a, b)-plane is the parabola*

$$b = c_E(\gamma)\, a^2 = -\frac{\gamma}{1+4\gamma}\, a^2. \tag{4.5}$$

Proof. Each of these bifurcations is defined by a codimension-one condition on the parabola W as given by (3.1), so that their loci can be computed explicitly.

1. By definition of the cusp-transition bifurcation, W passes through the precusp point C_0 at $(x, y) = (0, 0)$. Substitution into (3.1) gives the locus.

2. Since the line field $\mathcal{E}$ of the normal-form map F consists of horizontal lines, a loop-creation bifurcation occurs when the extremum at $(x, y) = (a, b)$ of W lies on J_0. Substitution into (3.1) gives the locus.

3. A transversal crossing of W and J_0, say, at (x_0, y_0), corresponds to a tangency of $F(W)$ with J_1. Hence, an intersection-at-tangency bifurcation occurs when W intersects $\widehat{J}_0$ at the same y-value, that is, at some point (x_1, y_0). From (2.2), (2.3), and (2.5) we conclude that then $x_1 = -2x_0$, and we determine x_0 from

$$\gamma(x_0 - a)^2 + b = \gamma(-2x_0 - a)^2 + b$$

as $x_0 = -2a$. Hence, in order for (x_0, y_0) to lie on J_0, a and b should be such that

$$y_0 = \gamma(x_0 - a)^2 + b = -x_0^2 \Leftrightarrow y_0 = 9\gamma a^2 + b = -4a^2,$$

which gives the locus.

4. At the tangency-creation bifurcation W is tangent to J_0 at, say, $(x_t, y_t) \in J_0 \cap W$. Since the slopes of W and J_0 at (x_t, y_t) must be equal, we conclude from (2.3) that

$$\begin{cases} y_t &= \gamma(x_t - a)^2 + b &= -x_t^2, \\ & 2\gamma(x_t - a) &= -2x_t, \end{cases}$$
$$\Leftrightarrow \begin{cases} y_t &= \gamma\frac{a^2}{(1+\gamma)^2} + b &= -\frac{a^2\gamma^2}{(1+\gamma)^2}, \\ x_t &= \frac{a\gamma}{1+\gamma}, \end{cases}$$

and this gives the locus.

5. The locus of tangency of W with $\widehat{J}_0$ can be derived exactly as for the tangency with J_0, by using (2.5) instead of (2.3).

Finally, there are only these five codimension-one bifurcations because, generically in the normal form setting, W can only either intersect J_0 and $\widehat{J}_0$ transversely or have a quadratic tangency with J_0 or $\widehat{J}_0$. In particular, higher-codimension singularities, for example, cubic tangencies, are not possible in the normal-form setting. ■

Indeed, the codimension-one bifurcations of Proposition 4.1 give rise to different configurations of $F(W)$ and J_1. The first three bifurcations involve transverse intersection of W and J_0. At the cusp-transition bifurcation, a tangency point of $F(W)$ with J_1 moves from the left to the right of the cusp point C_1. At the loop-creation bifurcation a small loop, that is, a self-intersection, is created; see also Figure 1. At the intersection-at-tangency bifurcation there is an exchange of the order along $F(W)$ between a self-intersection of $F(W)$ and a (generic) tangency between $F(W)$ and J_1. The other two bifurcations correspond to codimension-one tangencies of W with J_0 and $\widehat{J}_0$. At a tangency-creation two transverse intersections of W and J_0 and, hence, two (generic) tangencies between $F(W)$ and J_1, are created. At the enter-exit bifurcation two transverse intersections of $F(W)$ and J_1 are created.

4.2. Representation on the Poincaré disk and bifurcations at infinity. Events as defined in Definition 3.4 can move outside a fixed neighborhood of interest, which changes the equivalence type of $F(W)$ inside this neighborhood by changing the number of events that one encounters. In order to represent the entire phase portrait in a compact region, we now introduce the representation of the (x, y)- and (z, y)-planes on the Poincaré disk. The representation on the Poincaré disk is quite popular in bifurcation theory, because any dynamically relevant objects can be kept track of even if they bifurcate at (the circle representing) infinity; see, for example, [21].

The Poincaré disk is obtained by "closing off" the plane $\mathbb{R}^2$ with a circle that represents the asymptotic directions of curves at infinity. The phase plane is then represented on the unit disk, where points $e^{i\phi}$ on the unit circle correspond to the directional limits ϕ. In the present situation we are concerned with the curve W relative to J_0 and $\widehat{J}_0$, all of which are parabolas and graphs over the horizontal x-axis. Hence, the directional limits of J_0 and $\widehat{J}_0$ for $x \to \pm\infty$ are $-\pi/2$, so that these curves start and end on the Poincaré disk at the point

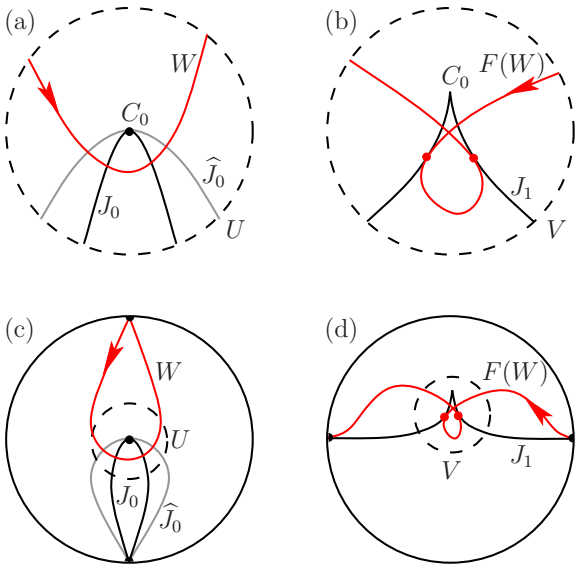


Figure 7. *The local interaction of W (red curve) with J_0 (black curve) and $\widehat{J}_0$ (gray curve) in a neighborhood U* (a) *corresponds to the local interaction of $F(W)$ (red curve) with J_1 (black curve) in a neighborhood V* (b)*. Panels* (c) *and* (d) *show the same situation sketched on the Poincaré disk. The example is phase portrait* 31*; compare with Figure* 3*.*

$-i = e^{-\pi i/2}$. Similarly, the curve W starts and ends at $-i$ for $\gamma < 0$ and at i for $\gamma > 0$. The curve J_1, on the other hand, has the directional limits 0 and π, and the same is true for the image of any parabola as given by (3.1) for any values of a, b, and γ. Therefore, J_1 and $F(W)$ have the limits -1 and $+1$ on the Poincaré disk. The situation is illustrated in Figure 7 for the phase portrait from Figure 3, that is, for $\gamma > 0$. Panel (a) of Figure 7 shows the local picture near J_0 and panel (c) the corresponding picture on the Poincaré disk. The respective images are shown in panels (b) and (d).

We have the following result concerning bifurcations of phase portraits on the Poincaré disk.

Proposition 4.2 (bifurcations at infinity). *In the normal-form setting of the cusp-cusp bifurcation with W given by* (3.1) *there are three bifurcations at infinity on the Poincaré disk that change the equivalence class according to Definition* 3.4*.*

1. *The* tangency at infinity*, where a generic tangency of $F(W)$ with J_1 moves via infinity from one side of the cusp point C_1 to the other. The locus of this bifurcation is given by $\gamma = -1$, which is the vertical asymptote of the coefficient of the tangency creation $c_T(\gamma) = -\frac{\gamma}{1+\gamma}$; see* (4.4) *and, for example, the transition between phase portraits* 2 *and* 9 *in Figure* 9*.*
2. *The* intersection at infinity*, where a transverse intersection of $F(W)$ and J_1 moves via infinity from one side of the cusp point C_1 to the other. The locus of this bifurcation is given by $\gamma = -\frac{1}{4}$, which is the vertical asymptote of the coefficient of the enter-exit bifurcation $c_E(\gamma) = -\frac{\gamma}{1+4\gamma}$; see* (4.5) *and, for example, the transition between phase portraits* 15 *and* 22 *in Figure* 9*.*

3. *The* self-intersection at infinity*, where a self-intersection of $F(W)$ bifurcates at the boundary of the Poincaré disk. This happens by a swap between the begin and end points of $F(W)$ at ± 1, and this also results in a change of direction of $F(W)$. The locus of this bifurcation is given by $\gamma = -\frac{1}{3}$, which corresponds to the crossing of W and R_0; see* (3.2) *and, for example, the transition between phase portraits* 7 *and* 15 *in Figure* 9*.*

Proof. The respective changes (of phase portraits) and their loci follow immediately from the fact that the loci of the tangency bifurcation and the enter-exit bifurcation have asymptotes.

For any fixed values of a and b the limits for $x \to \pm\infty$ of the parabola W lie below the curve R_0 if $\gamma < -\frac{1}{3}$, and above the curve R_0 if $\gamma > -\frac{1}{3}$. This means that the directional limits of $F(W)$ change when γ crosses $-\frac{1}{3}$. Effectively, the two ends of $F(W)$ reconnect differently to the points ± 1 on the boundary of the Poincaré disk. Note that for the degenerate case of $\gamma = -\frac{1}{3}$ they both approach the curve $\{z = 0\}$, that is, the point $-i$ on the boundary of the Poincaré disk. One can view this as an "ambivalent connection" to the points ± 1 via the lower quarters of the circle at infinity. ■

4.3. Inversion near the cusp. As we discuss now, there is one more possibility for transitions between equivalence classes of $F(W)$. Recall that according to Proposition 3.2 the passage through the degenerate case $\gamma = 0$ changes the type of the central singularity for $(a, b) = (0, 0)$. This also has consequences for $F(W)$ if $(a, b) \neq (0, 0)$.

Proposition 4.3 (inversion near cusp). *The parabola W given by* (3.1) *has a maximum for $\gamma < 0$ and a minimum for $\gamma > 0$. As a result, near the cusp point C_1 on J_1 the curve $F(W)$ has a local maximum for $\gamma < 0$ and a local minimum for $\gamma > 0$. We call the transition through the degenerate case $\gamma = 0$ the* inversion near the cusp*, and it results in a change of the phase portrait in the case that $F(W)$ interacts with J_1; see, for example, the transition between phase portraits* 24 *and* 29 *in Figure* 9*.*

Note that the inversion near the cusp is not a bifurcation in a classical sense. Rather it is the vanishing of the quadratic term which changes the direction in which $F(W)$ approaches a neighborhood of the cusp point C_1. This manifests itself quite dramatically as a type of "inversion" in this neighborhood. There are actually only three different transitions featuring an inversion near the cusp; see Figures 8 and 9, and compare panels 24 with 29, 25 with 30, and 26 with 31. We refer to the inversion near the cusp loosely as a bifurcation (albeit an unusual one) of codimension one, because it forms a codimension-one boundary between equivalence classes of $F(W)$ in (a, b, γ)-space.

5. Two-parameter unfoldings in the (a, b)-plane. The five codimension-one bifurcations of section 4.1 are all given by parabolas $b = c_*(\gamma)\, a^2$ through the origin. Hence, the bifurcation diagram in the (a, b)-plane is determined entirely by the ordering of the respective bifurcations, which is given by the sizes of the coefficients $c_*(\gamma)$. In Figure 8 we show all information on the cusp-cusp unfolding (in the normal-form setting) in a convenient and very condensed representation by plotting the graphs of the functions $c_*(\gamma)$. This corresponds geometrically to showing how the slices $\{a = \pm 1\}$ in the (a, b)-plane change with γ. The different bifurcations are shown in different colors as introduced in Figures 5 and 6. In this representation the other bifurcations in sections 4.2 and 4.3 are vertical lines since they do not depend on a and b.

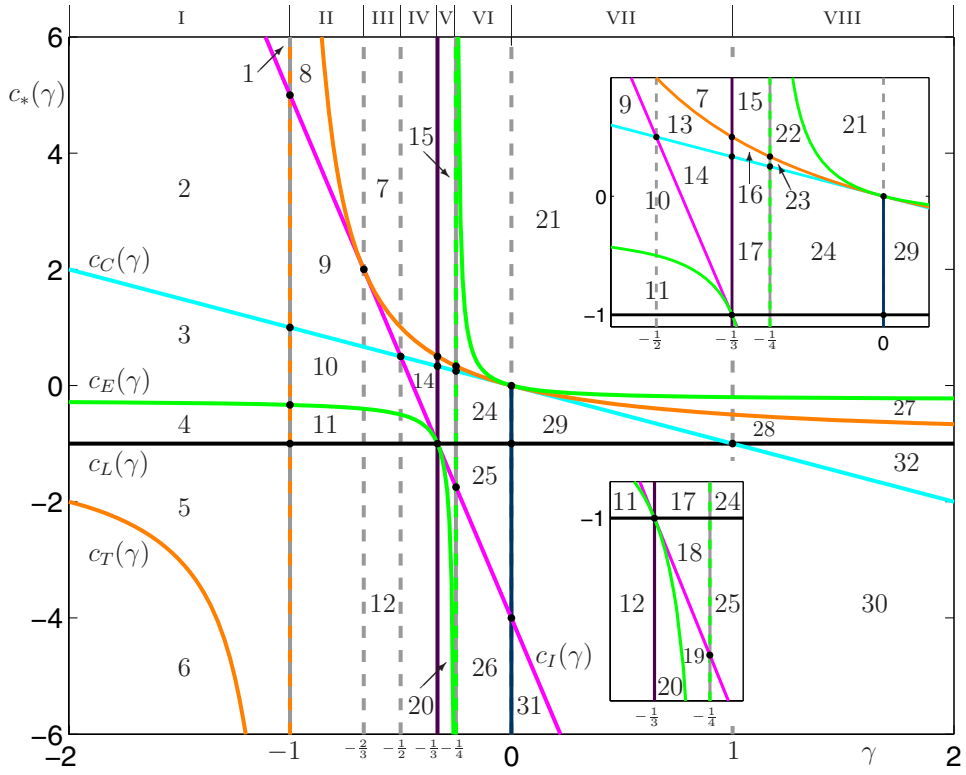


Figure 8. *Division of the $(\gamma, c_*(\gamma))$-plane into regions of different phase portraits for $a > 0$, which are shown in Figure 9; see Proposition 5.1. The γ-axis is divided into regions of different two-parameter unfoldings, cases* I–VIII*, which are shown individually in Figures 10 to 17. The colors of the dividing codimension-one bifurcations from Proposition 4.1 are as introduced in Figures 5 and 6; the dash-colored vertical lines at $\gamma = -1$ and $\gamma = -\frac{1}{4}$ and the dark-purple vertical line at $\gamma = -\frac{1}{3}$ are the respective bifurcations at infinity from Proposition 4.2; the vertical dark blue line is the inversion near the cusp from Proposition 4.3.*

The set of all bifurcation curves divides the $(\gamma, c_*(\gamma))$-plane into 32 regions of equivalent phase portraits; the insets show enlargements of two rather small regions. Representatives of all 32 phase portraits that one finds for $a > 0$ are shown as sketches on the Poincaré disk in Figure 9. Since all bifurcation curves in the (a, b)-plane are parabolas through the origin, the bifurcation diagram is symmetric under the operation $a \mapsto -a$. It follows directly from (2.2) and (3.1) that the corresponding phase portraits for $a < 0$ are obtained by reflecting the curve $F(W)$ on the y-axis on the (z, y)-plane without changing its orientation, so that all events (see Definition 3.3) now occur on the opposite side of C_1 and in the reverse order. We call the thusly obtained phase portrait the *conjugate phase portrait* and denote it by a bar across the number. Note that a phase portrait and its conjugate are generally not equivalent.

The unfoldings represented in Figures 8 and 9 were obtained by a combination of analysis and topological arguments—guided by exploration on the computer. We stress again that in the normal-form setting all bifurcation curves can be found analytically. Furthermore, phase portraits can be explored simply by plotting $F(W)$ relative to J_1 for appropriate values of a, b, and γ. The overall result can be summarized as follows.

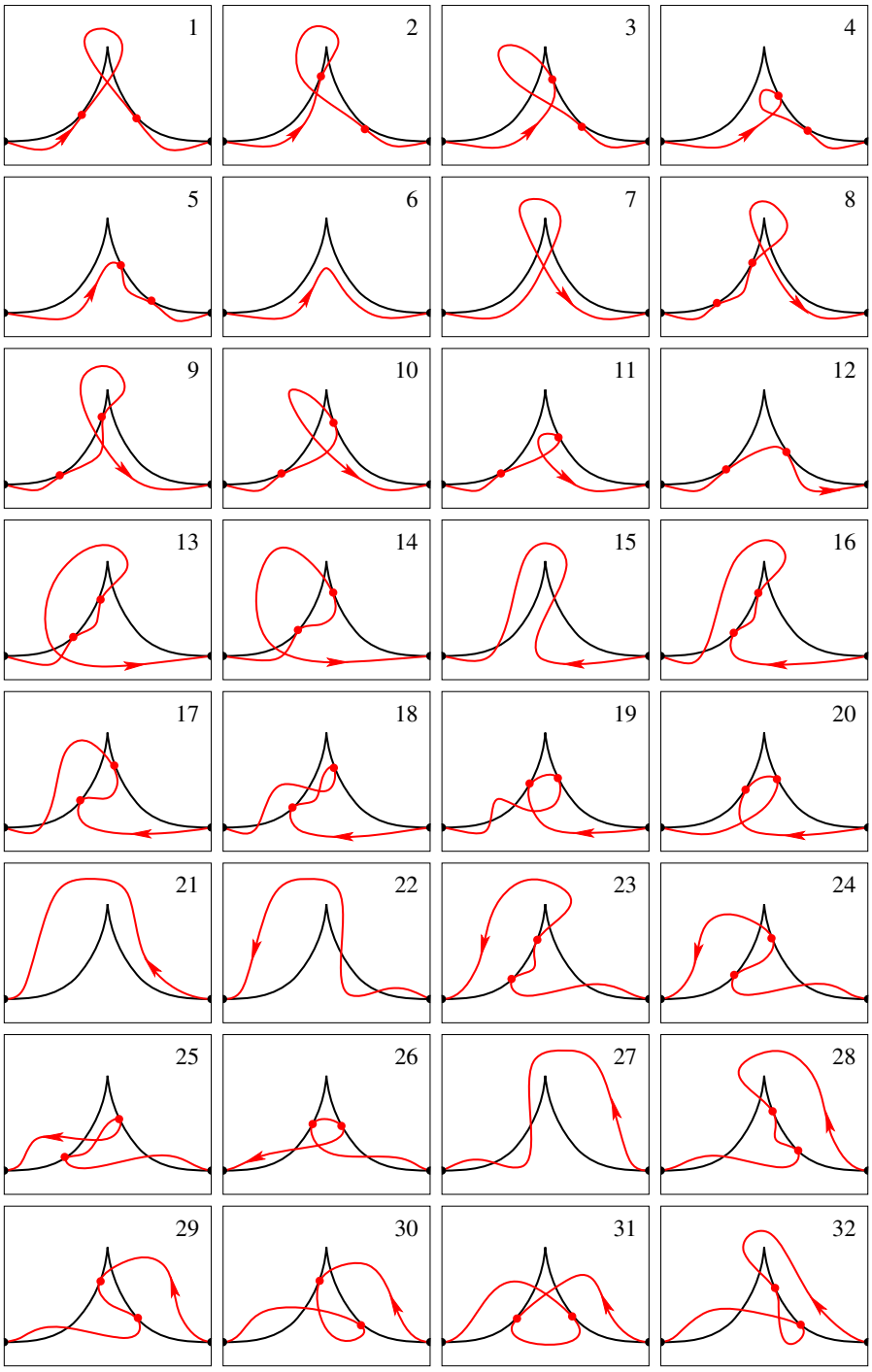


Figure 9. *Sketches on the Poincaré disk of all generic phase portraits for $a > 0$ of $F(W)$ (red curve) in relation to J_1 (black curve); the arrows show the direction of parametrization by x. The respective phase portraits for $a < 0$ are obtained by the operation of conjugation.*

Proposition 5.1 (two-parameter unfoldings). *There are eight two-parameter unfoldings in the (a,b)-plane, denoted* I–VIII*, as shown in Figure 8. The boundary points between the respective intervals of the γ-line are -1, $-\frac{2}{3}$, $-\frac{1}{2}$, $-\frac{1}{3}$, $-\frac{1}{4}$, 0, and 1.*

Proof. According to Propositions 4.2 and 4.3, the points $\gamma = -1$, $\gamma = -\frac{1}{3}$, $\gamma = -\frac{1}{4}$, and $\gamma = 0$ give rise to changes of the equivalence class of phase portraits and, hence, of neighboring two-parameter unfoldings in the (a,b)-plane. It follows from (4.1) and (4.3) that $c_C(\gamma)$ and $c_I(\gamma)$ intersect at $\gamma = -\frac{1}{2}$, where the cusp transition and the intersection-at-tangency bifurcation change their order. Similarly, from (4.1) and (4.2), $c_C(\gamma)$ and $c_L(\gamma)$ intersect at $\gamma = 1$, where the cusp transition and the loop creation change their order. Finally, from (4.3) and (4.4) we conclude that $c_I(\gamma)$ is tangent to $c_T(\gamma)$ at $\gamma = -\frac{2}{3}$. This means that the intersection-at-tangency happens at a double-tangency point on J_1. Therefore, the intersection at $\gamma = -\frac{2}{3}$ is a genuine codimension-two point that leads to a change in the phase portraits. Indeed, for (a,b) in between the tangency-creation and intersection-at-tangency bifurcation curves, the self-intersection of $F(W)$ lies closer to C_1 than the two tangencies of $F(W)$ with J_1 if $\gamma < -\frac{2}{3}$, while it lies furthest away from C_1 if $\gamma > -\frac{2}{3}$; compare phase portraits 8 and 13 in Figures 11 and 12, respectively. ∎

The two-parameter unfoldings I–VIII in the (a,b)-plane are shown individually in Figures 10 to 17. In each figure, the middle panel shows how the colored codimension-two bifurcation curves of Proposition 4.1 divide the (a,b)-plane into regions of different phase portraits. The surrounding panels on the left show the respective configurations of W near C_0 and those on the right that of $F(W)$ near C_1. Both the bifurcation curves and the configurations of W and $F(W)$ were drawn in MATLAB directly from their respective formulas. In this way, one gets an impression of actual shapes and sizes. In particular, we chose a fixed neighborhood of C_0 and its image, which is a fixed neighborhood of C_1. Hence, all panels in Figures 10 to 17 can be directly compared in terms of scale; where necessary, insets provide enlargements of how $F(W)$ lies relative to J_1. A slight disadvantage of choosing a fixed neighborhood for all cases is that an intersection of W with J_0 or $\widehat{J}_0$, and thus the corresponding generic event, may have "moved" outside of a panel. All phase portraits shown in Figures 10 to 17 are for $a > 0$, as was the case for the sketches in Figure 9 (which contain all occurring generic events). Note that crossing the line $a = 0$ above or below all bifurcation curves does not constitute a bifurcation, so that the respective phase portraits must be invariant under conjugation. The fact that this is indeed the case is evidence of the consistency of the unfoldings in Figures 10 to 17.

As mentioned, the phase portraits for $a < 0$ can be obtained by conjugation, but a conjugate phase portrait may be equivalent to a different phase portrait for $a > 0$. The following phase portraits do not result in new phase portraits under conjugation with respect to the original list of 32 for $a > 0$ in Figure 9.

- invariant under conjugation: 1, 6, 7, 12, 15, 20, 21, 26, and 31.
- identities under conjugation: $\overline{22} = 27$, $\overline{23} = 28$, and $\overline{24} = 29$.

As a result, there are 17 new classes of phase portraits that one obtains under conjugation from the original list for $a > 0$, namely,

- $\overline{2}$, $\overline{3}$, $\overline{4}$, $\overline{5}$, $\overline{8}$, $\overline{9}$, $\overline{10}$, $\overline{11}$, $\overline{13}$, $\overline{14}$, $\overline{16}$, $\overline{17}$, $\overline{18}$, $\overline{19}$, $\overline{25}$, $\overline{30}$, and $\overline{32}$.

This means that there is a total of 51 different phase portraits according to Definition 3.4. If one chooses to consider conjugate phase portraits to be the same, that is, one does not

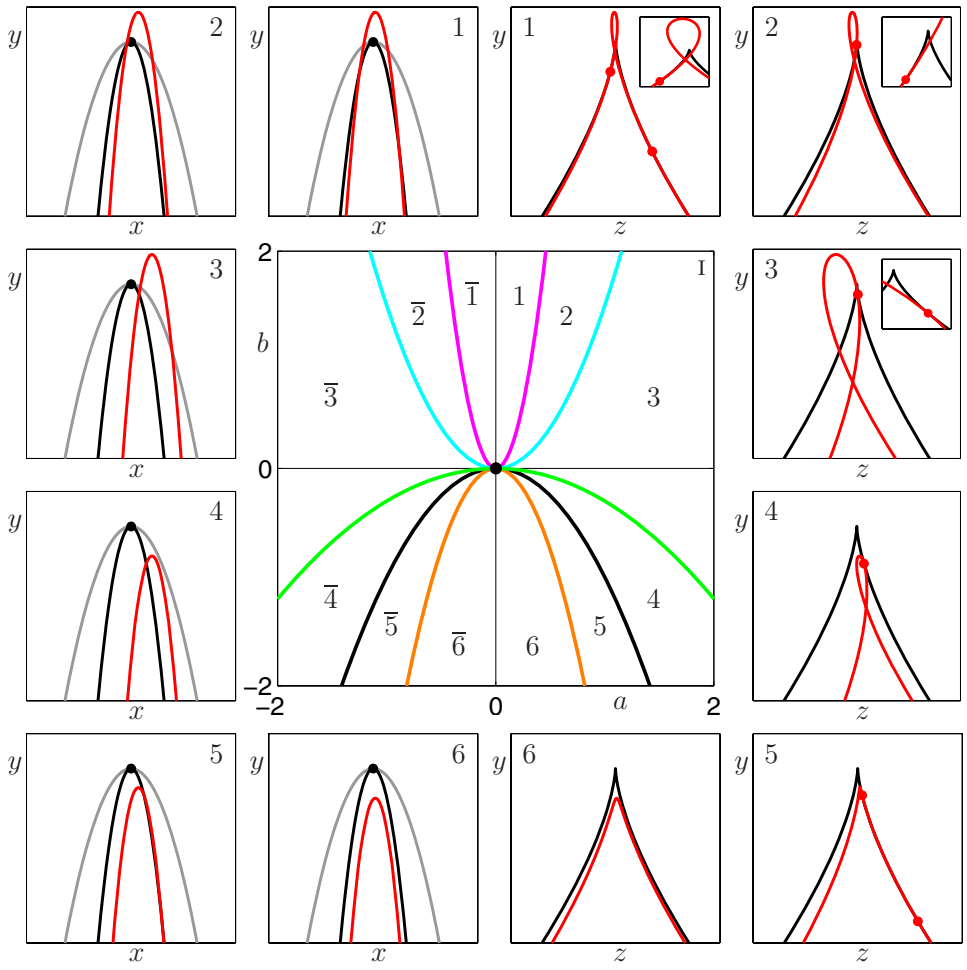


Figure 10. *Case* I: $-\infty < \gamma < -1$. *The two-parameter unfolding in the* (a, b)*-plane with the bifurcation curves* C *(cyan),* L *(black),* I *(magenta),* T *(orange), and* E *(green); compare with Figures 5 and 6. Representative phase portraits near* C_0 *are presented counterclockwise from the second image at the top, and near* C_1 *clockwise from the third image at the top; shown are* W *and* $F(W)$ *(red curves),* J_0 *and* J_1 *(black curves), and* $\widehat{J}_0$ *(gray curve). The data is for* $\gamma = -1.5$*, with phase portraits for* $(a, b) = (0.2, 1.7)$*,* $(a, b) = (0.7, 1.7)$*,* $(a, b) = (2.0, 1.7)$*,* $(a, b) = (2.0, -1.7)$*,* $(a, b) = (0.7, -1.1)$*, and* $(a, b) = (0.2, -1.7)$*, respectively.*

distinguish between the two parts of J_1 (either side of the cusp point) on which the intersections and tangencies occur, then there is a total of 29 phase portraits.

5.1. Relevance of normal-form unfoldings. The map F is the normal form for any endomorphism f in a sufficiently small neighborhood of a generic cusp point C_1 on J_1. Furthermore, at the cusp-cusp bifurcation the tangency of the smooth invariant curve W with J_0 at the precusp point is quadratic. Therefore, we have the following.

Corollary 5.2. *Consider a family of planar endomorphisms* f_λ *with a cusp-cusp bifurcation at the parameter point* $(\lambda_1^*, \lambda_2^*)$*, that is, a quadratic tangency of an invariant curve* W *to a precusp point* C_0 *on* J_0*. Then there exist sufficiently small neighborhoods* U *of* C_0 *and* V *of*

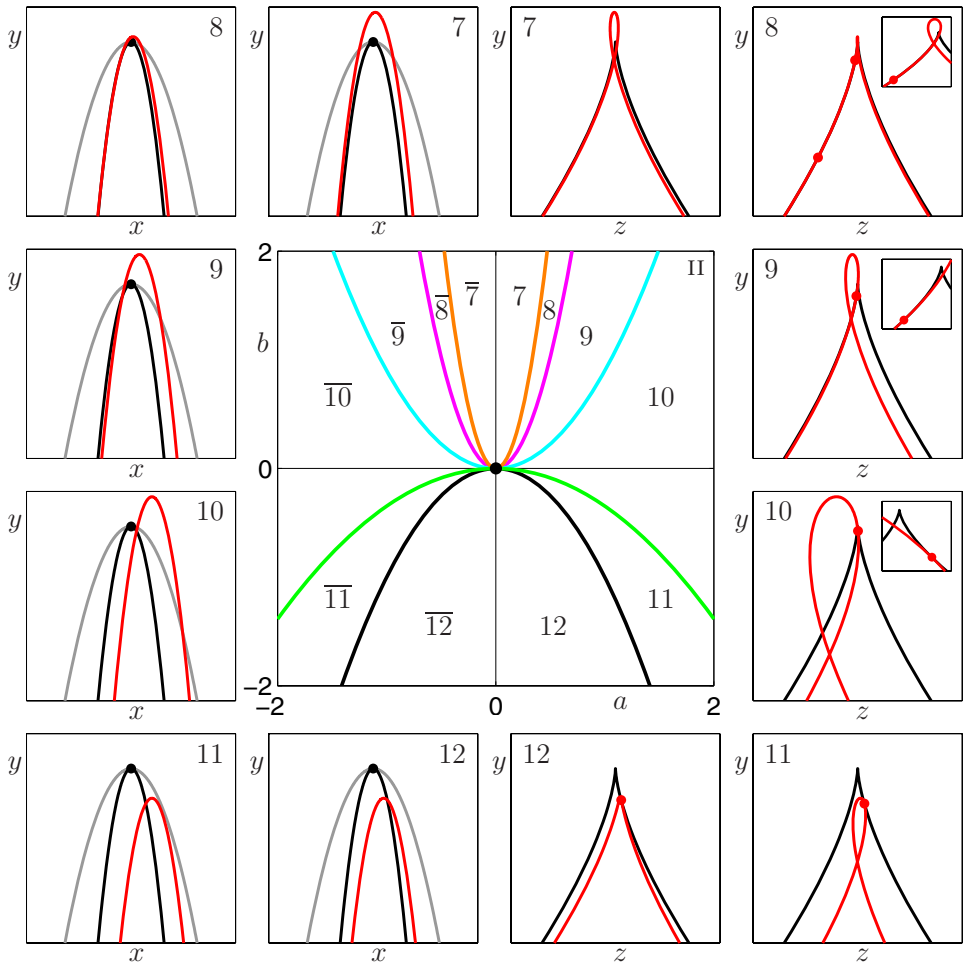


Figure 11. *Case* II: $-1 < \gamma < -\frac{2}{3}$. *The two-parameter unfolding in the* (a, b)*-plane with the bifurcation curves* C *(cyan),* L *(black),* I *(magenta),* T *(orange), and* E *(green); compare with Figures 5 and 6. Representative phase portraits near* C_0 *are presented counterclockwise from the second image at the top, and near* C_1 *clockwise from the third image at the top; shown are* W *and* $F(W)$ *(red curves),* J_0 *and* J_1 *(black curves), and* $\widehat{J}_0$ *(gray curve). The data is for* $\gamma = -0.9$*, with phase portraits for* $(a, b) = (0.2, 1.7)$*,* $(a, b) = (0.2, 0.3)$*,* $(a, b) = (0.8, 1.7)$*,* $(a, b) = (2.0, 1.7)$*,* $(a, b) = (2.0, -1.7)$*, and* $(a, b) = (1.0, -1.7)$*, respectively.*

the cusp point $C_1 = f(C_0)$ *such that the following hold:*

1. *Generically, one of the cases* I–VIII *of two-parameter unfoldings occurs as the bifurcation diagram of* f *in* V*. Which case occurs depends on the exact position of the two-jet of the curve* W *at the cusp-cusp bifurcation relative to the normal-form coordinate systems generated by the two-jets of* J_0 *and* $\widehat{J}_0$*.*
2. *For each of the cases* I–VIII *the parameters* λ_1 *and* λ_2 *generically unfold the cusp-cusp bifurcation in a sufficiently small neighborhood* Λ *of* $(\lambda_1^*, \lambda_2^*)$*. In particular, by varying* λ_1 *and* λ_2 *all respective classes of phase portraits can be found in* V*.*

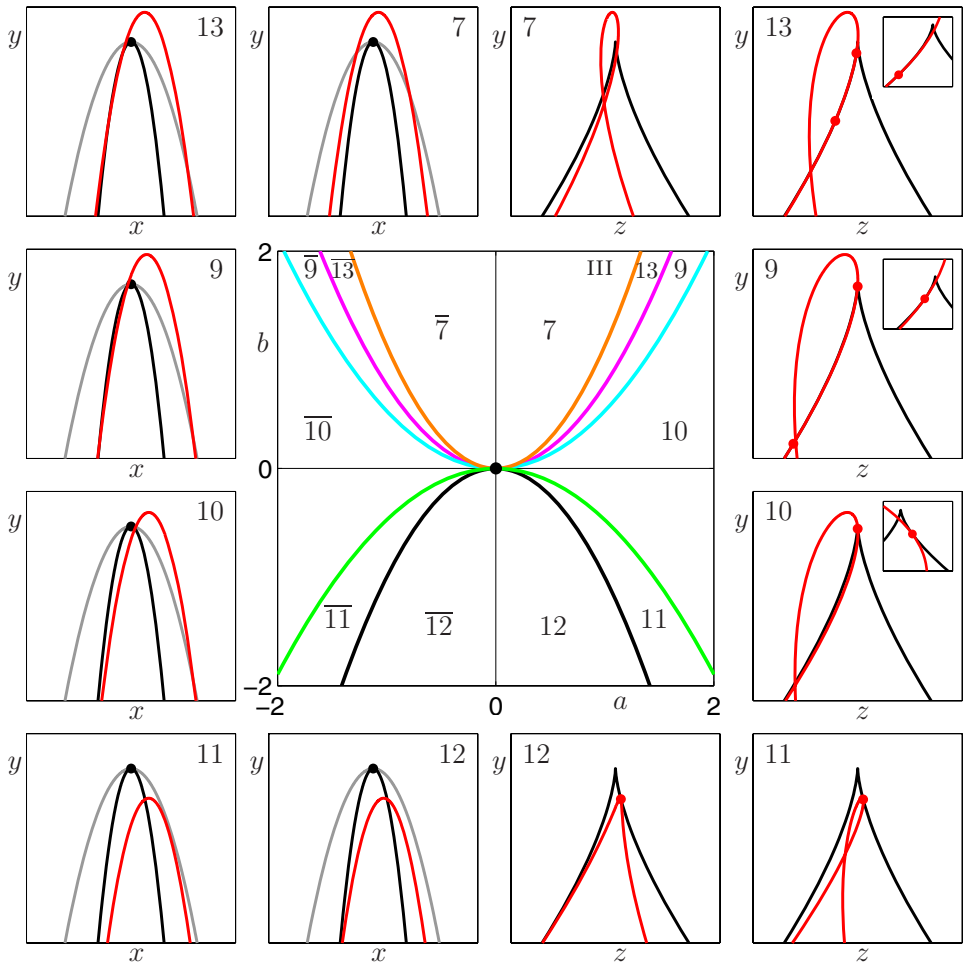


Figure 12. *Case* III: $-\frac{2}{3} < \gamma < -\frac{1}{2}$. *The two-parameter unfolding in the* (a, b)*-plane with the bifurcation curves* C *(cyan),* L *(black),* I *(magenta),* T *(orange), and* E *(green); compare with Figures 5 and 6. Representative phase portraits near* C_0 *are presented counterclockwise from the second image at the top, and near* C_1 *clockwise from the third image at the top; shown are* W *and* $F(W)$ *(red curves),* J_0 *and* J_1 *(black curves), and* $\widehat{J}_0$ *(gray curve). The data is for* $\gamma = -0.53$*, with phase portraits for* $(a, b) = (0.5, 1.7)$*,* $(a, b) = (1.3, 1.7)$*,* $(a, b) = (1.5, 1.7)$*,* $(a, b) = (1.7, 0.8)$*,* $(a, b) = (1.7, -1.7)$*, and* $(a, b) = (1.0, -1.7)$*, respectively.*

Proof. Consider the coefficient of the two-jet of W at C_0 relative to the normal-form coordinate systems generated by the two-jets of J_0 and $\widehat{J}_0$. If this coefficient lies in one of the open regions representing cases I–VIII, then this is an open condition. Since the two-jet of W depends continuously on (λ_1, λ_2), its coefficient does not leave the respective open region for $(\lambda_1^*, \lambda_2^*)$ as long as Λ is a sufficiently small neighborhood. Furthermore, generically the map from (λ_1, λ_2) to (a, b) is a local diffeomorphism. Hence, generically, (λ_1, λ_2) unfolds the cusp-cusp bifurcation. ■

Proposition 5.2 is in the spirit of local bifurcation theory: it states that the generic open regions corresponding to cases I–VIII are stable under perturbation of the purely quadratic

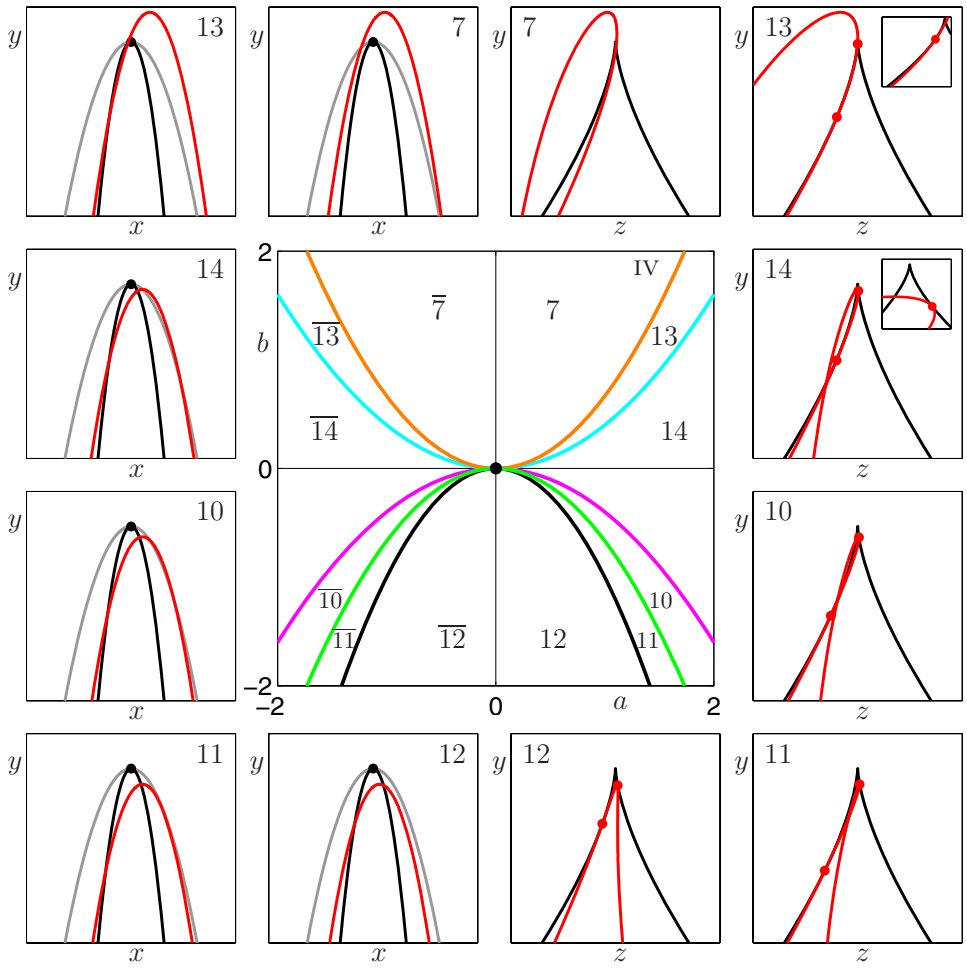


Figure 13. *Case* IV: $-\frac{1}{2} < \gamma < -\frac{1}{3}$. *The two-parameter unfolding in the* (a, b)*-plane with the bifurcation curves* C *(cyan),* L *(black),* I *(magenta),* T *(orange), and* E *(green); compare with Figures 5 and 6. Representative phase portraits near* C_0 *are presented counterclockwise from the second image at the top, and near* C_1 *clockwise from the third image at the top; shown are* W *and* $F(W)$ *(red curves),* J_0 *and* J_1 *(black curves), and* $\widehat{J}_0$ *(gray curve). The data is for* $c = -0.4$*, with phase portraits for* $(a, b) = (1.1, 1.7)$*,* $(a, b) = (1.8, 1.7)$*,* $(a, b) = (1.1, -0.3)$*,* $(a, b) = (1.1, -0.6)$*,* $(a, b) = (1.1, -0.9)$*, and* $(a, b) = (0.6, -0.9)$*, respectively.*

character of the normal form, so that they must be expected in practical situations. Note also that, no matter which case of unfolding occurs, the five codimension-one bifurcations of Proposition 4.1 are, generically, part of the bifurcation diagram of f in V. When λ_1 and λ_2 are changed outside Λ then an intersection or (generic) tangency of $f(W)$ with J_1 may leave the neighborhood V. This corresponds to bifurcations at infinity; see Proposition 4.2.

The open regions of the normal-form setting are stable under perturbation, but this may or may not be true for the boundaries between them; see Figure 8. First, we expect the tangency between the curves $c_I(\gamma)$ and $c_T(\gamma)$ at $\gamma = -\frac{2}{3}$ to be preserved under perturbations, because it corresponds to a genuine codimension-two bifurcation, namely, a simultaneous tangency-

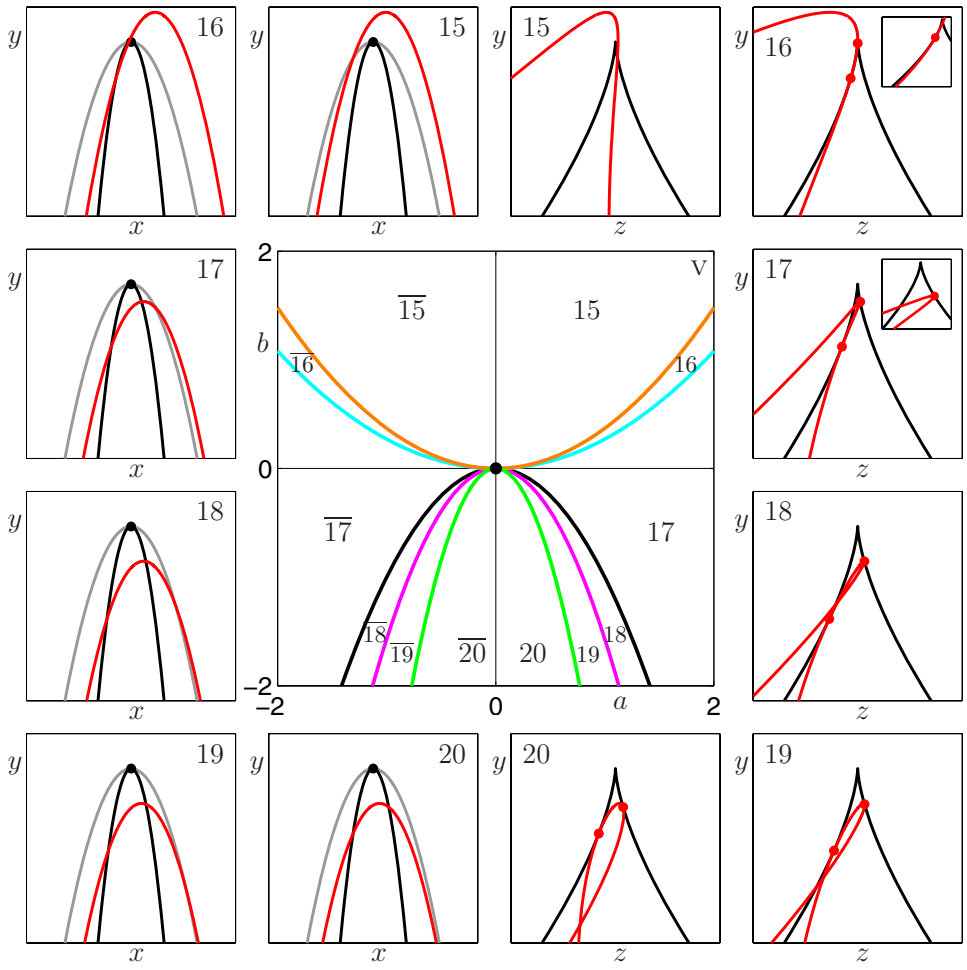


Figure 14. *Case* V: $-\frac{1}{3} < \gamma < -\frac{1}{4}$. *The two-parameter unfolding in the* (a, b)*-plane with the bifurcation curves* C *(cyan),* L *(black),* I *(magenta),* T *(orange), and* E *(green); compare with Figures 5 and 6. Representative phase portraits near* C_0 *are presented counterclockwise from the second image at the top, and near* C_1 *clockwise from the third image at the top; shown are* W *and* $F(W)$ *(red curves),* J_0 *and* J_1 *(black curves), and* $\widehat{J}_0$ *(gray curve). The data is for* $\gamma = -0.27$*, with phase portraits for* $(a, b) = (1.2, 1.7)$*,* $(a, b) = (2.3, 1.7)$*,* $(a, b) = (1.2, -1.0)$*,* $(a, b) = (1.2, -2.0)$*,* $(a, b) = (1.0, -2.0)$*, and* $(a, b) = (0.6, -2.0)$*, respectively.*

creation and intersection-at-tangency bifurcation. On the other hand, the tangency between the curves $c_E(\gamma)$, $c_T(\gamma)$, and $c_C(\gamma)$ at $\gamma = 0$ should be destroyed by generic perturbations. However, as follows from the definition of the bifurcations involved, the order of the bifurcations cannot change. Similarly, we expect that the tangency between the curves $c_E(\gamma)$ and $c_I(\gamma)$ at $\gamma = -\frac{1}{3}$ exactly at the intersection with $c_L(\gamma)$ will be destroyed, while the order of the curves $c_E(\gamma)$ and $c_I(\gamma)$ will be maintained. Overall, new regions may be created that feature phase portraits that do not occur in the normal form. However, because the invariant curve W is assumed to have a quadratic tangency (which is a generic assumption), the perturbation of the boundary curves in Figure 8 is small and goes to zero as the cusp-cusp point $(\lambda_1^*, \lambda_2^*)$ is

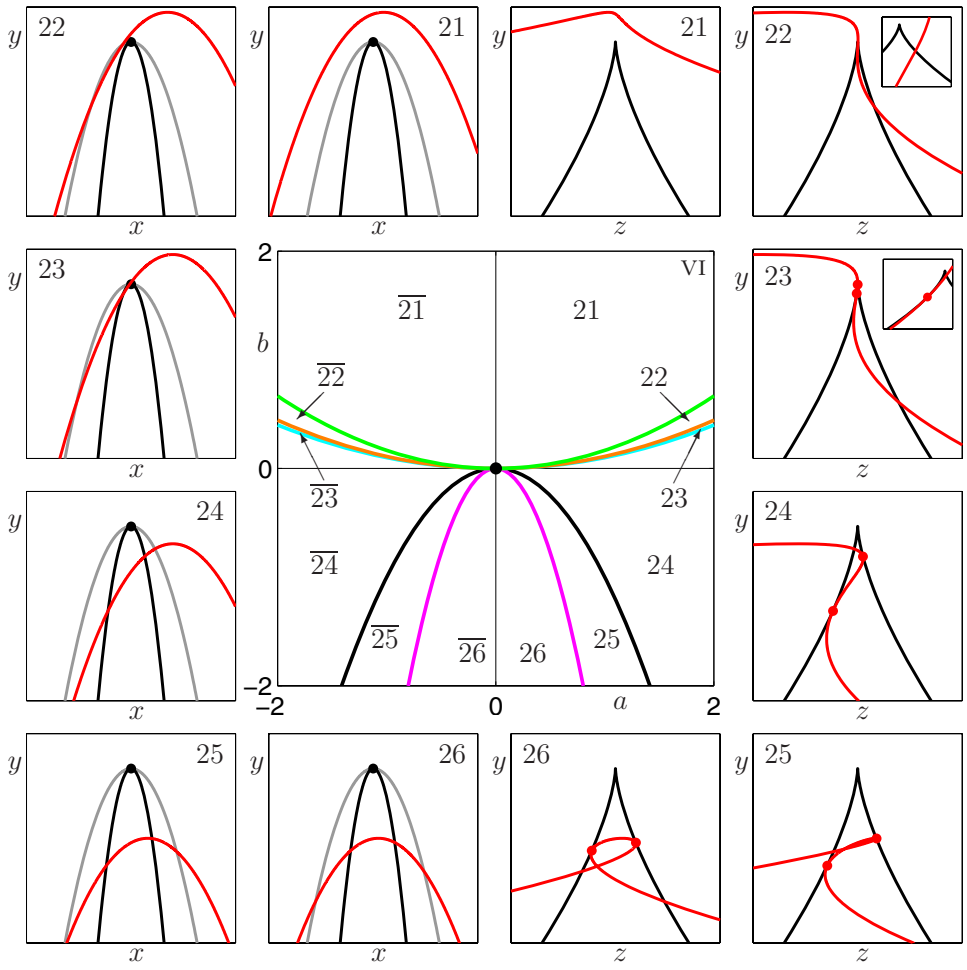


Figure 15. *Case* VI: $-\frac{1}{4} < \gamma < 0$. *The two-parameter unfolding in the* (a, b)*-plane with the bifurcation curves* C *(cyan),* L *(black),* I *(magenta),* T *(orange), and* E *(green); compare with Figures* 5 *and* 6. *Representative phase portraits near* C_0 *are presented counterclockwise from the second image at the top, and near* C_1 *clockwise from the third image at the top; shown are* W *and* $F(W)$ *(red curves),* J_0 *and* J_1 *(black curves), and* $\widehat{J}_0$ *(gray curve). The data is for* $\gamma = -0.1$*, with phase portraits for* $(a, b) = (1.0, 1.7)$*,* $(a, b) = (3.5, 1.7)$*,* $(a, b) = (4.0, 1.7)$*,* $(a, b) = (4.0, -1.0)$*,* $(a, b) = (1.6, -4.0)$*, and* $(a, b) = (0.5, -4.0)$*, respectively.*

approached. In other words, sufficiently close to $(\lambda_1^*, \lambda_2^*)$, one should expect to see "only" the two-parameter unfoldings I–VIII presented.

In any practical application of Proposition 5.2 it will be interesting to see how big the neighborhoods Λ and V are. Indeed, the larger they can be chosen, the easier it is to identify the different types of phase portraits that must occur near the cusp-cusp bifurcation. In practice, it is possible to deduce which case one is dealing with by carefully finding the respective phase portraits in V, as we will see from the example in the next section.

6. Cusp-cusp bifurcation in an adaptive control system. As a concrete example we consider the one-dimensional first-order approximation of a linear, time-invariant, single-input,

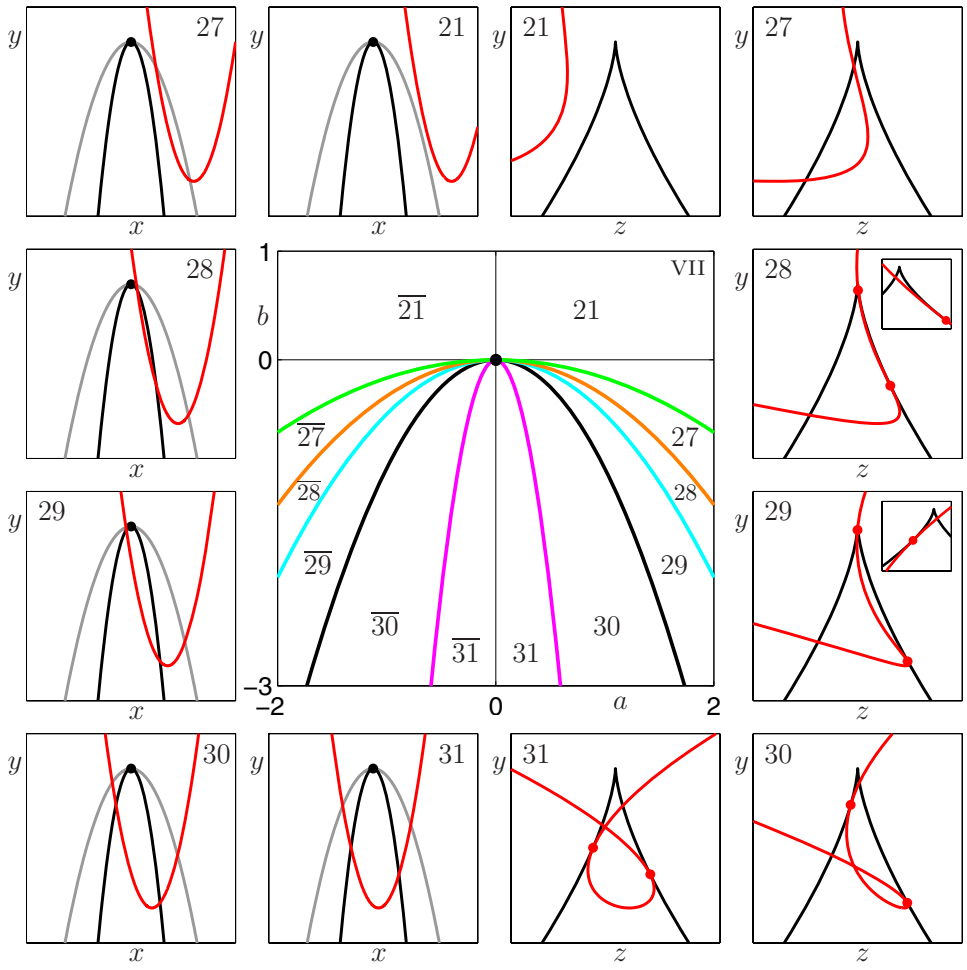


Figure 16. *Case* VII: $0 < \gamma < 1.0$. *The two-parameter unfolding in the* (a, b)*-plane with the bifurcation curves* C *(cyan),* L *(black),* I *(magenta),* T *(orange), and* E *(green); compare with Figures 5 and 6. Representative phase portraits near* C_0 *are presented counterclockwise from the second image at the top, and near* C_1 *clockwise from the third image at the top; shown are* W *and* $F(W)$ *(red curves),* J_0 *and* J_1 *(black curves), and* $\widehat{J}_0$ *(gray curve). The data is for* $\gamma = 0.5$*, with phase portraits for* $b = -8.0$ *and* $a = 7.5$*,* $a = 6.0$*,* $a = 4.5$*,* $a = 3.5$*,* $a = 2.0$*, and* $a = 0.5$*, respectively.*

single-output process with a unit time delay introduced in [9]. In this example, the plant contains an unknown parameter that must be estimated using input-output data available from previous time intervals. The objective is to design a controller that will track a constant nonzero reference signal. In nondimensional form, this leads to the planar endomorphism

$$g : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -xy + \eta \\ \beta y + \frac{px(-yx+\eta-1)}{c+x^2} \end{pmatrix}, \tag{6.1}$$

where x is the output signal—which should equal 1 if the controller achieves its objective—and y corresponds to the estimation of the unknown parameter in the system. The parameter η

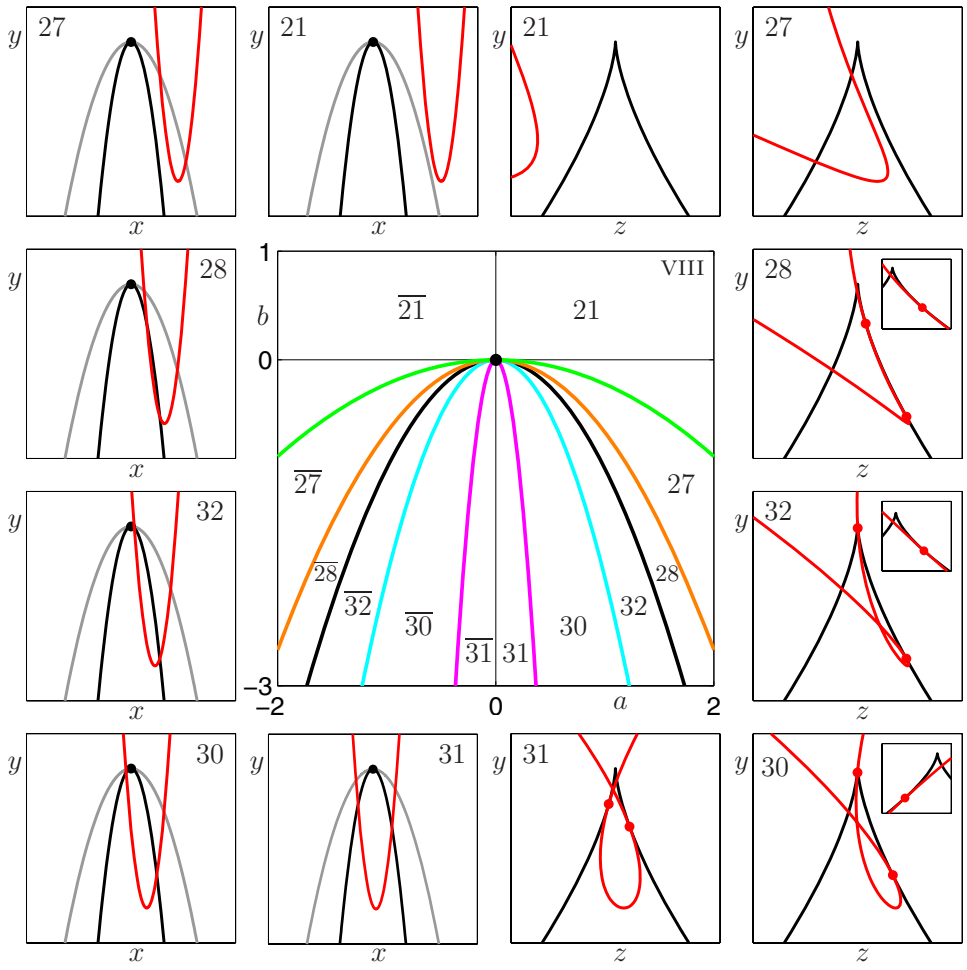


Figure 17. *Case* VIII: $1.0 < \gamma < \infty$. *The two-parameter unfolding in the* (a, b)*-plane with the bifurcation curves* C *(cyan),* L *(black),* I *(magenta),* T *(orange), and* E *(green); compare with Figures 5 and 6. Representative phase portraits near* C_0 *are presented counterclockwise from the second image at the top, and near* C_1 *clockwise from the third image at the top; shown are* W *and* $F(W)$ *(red curves),* J_0 *and* J_1 *(black curves), and* $\widehat{J}_0$ *(gray curve). The data is for* $\gamma = 2.0$*, with phase portraits for* $b = -8.0$ *and* $a = 6.5$*,* $a = 4.5$*,* $a = 3.2$*,* $a = 2.3$*,* $a = 1.5$*, and* $a = 0.3$*, respectively.*

corresponds to a constant disturbance of the system, and p is directly related to the adaptation gain. The parameters β and c determine the effectiveness of the feedback controller; $\beta = 1$ was used in [9]. The noninvertibility of the map is due to the time delay, or sampling time, when constructing the feedback.

The three-parameter family with $\beta = 1$ was also studied in the tutorial paper [10], where it was shown that a saddle exists with an unstable manifold that interacts with a cusp on J_1. The phase portraits in [10] were a direct motivation for us to investigate the codimension-two bifurcation organizing this type of behavior, namely, the cusp-cusp bifurcation. Our study of the possible unfoldings in section 4 indicates that the family (6.1) for $\beta = 1$ displays the

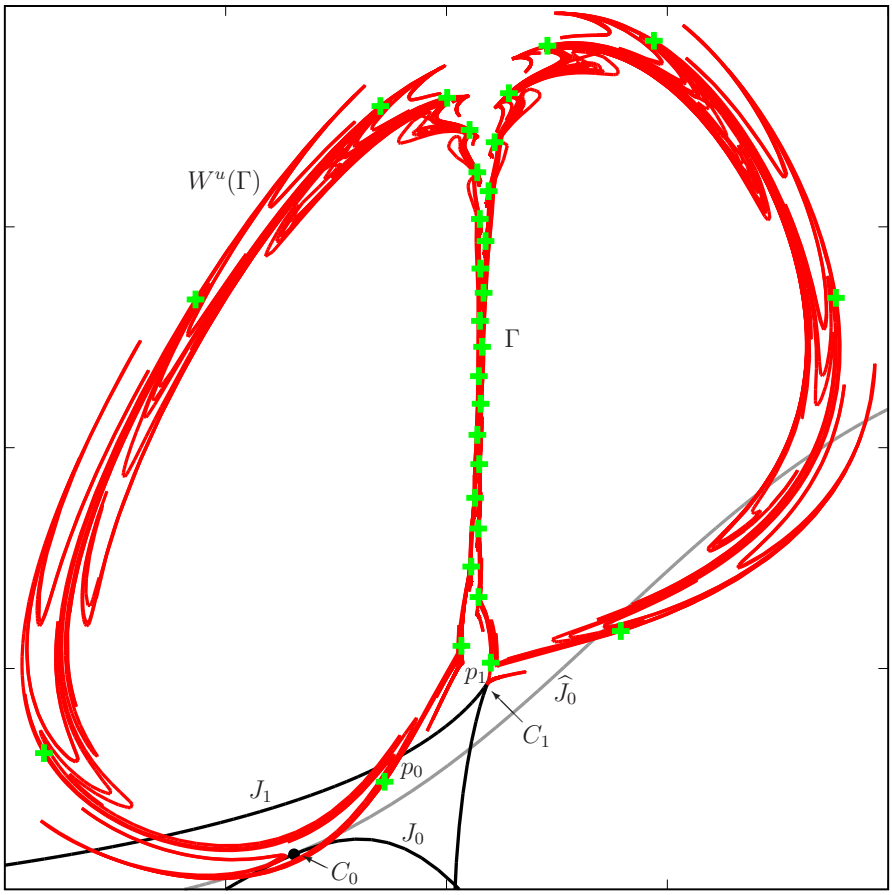


Figure 18. *Global phase portrait of the map g of (6.1) for $p = 0.81$, $c = 1.2$, $\eta = -0.153$, and $\beta = 1.0005$. Shown is how the unstable manifold $W^u(\Gamma)$ (red curve) of the saddle period-30 orbit Γ (green crosses) interacts with the curves J_0 (black curve), $\widehat{J}_0$ (gray curve), and J_1 (black cusped curve). The situation is close to a cusp-cusp bifurcation, as is evidenced by how $W^u(\Gamma)$ approaches the cusp point C_1.*

arrangements of an invariant curve near a cusp point C_1 that can be found in unfolding VII. The respective phase portraits are part of a scenario that involves the unstable manifold of a period-30 orbit inside a resonance tongue in some appropriate parameter plane. Unfortunately, while this unstable manifold has a quadratic tangency with J_0 very close to the precusp point C_0, it is not possible to vary the parameters such that the tangency occurs exactly at C_0 before leaving the narrow period-30 resonance tongue.

In order to find a cusp-cusp bifurcation for the family (6.1) it is necessary to vary β. Specifically, we consider the (η, β)-plane for fixed $p = 0.81$ and $c = 1.2$. By starting at $(\eta, \beta) = (-0.14, 1.0)$, we found a cusp-cusp point at $(\eta, \beta) \approx (-0.14439095, 1.00181158)$. Figure 18 shows the general arrangement of the period-30 orbit Γ (green crosses) and its unstable manifold $W^u(\Gamma)$ (red curves) for $p = 0.81$, $c = 1.2$, $\eta = -0.153$, and $\beta = 1.0005$. Successive images on the orbit Γ alternate from left to right, move up the middle of the figure, and then separate around the outsides. There are 15 points on each of the two "circles"

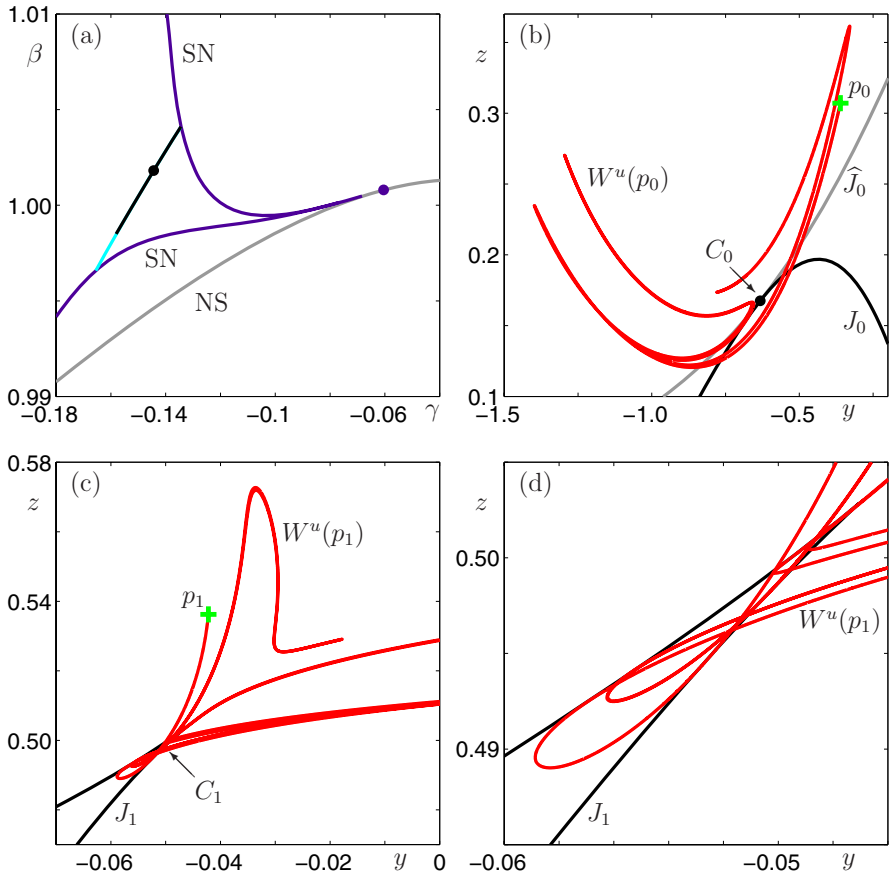


Figure 19. *Panel* (a) *shows the bifurcation diagram in the* (η, β)*-plane of* (6.1) *for fixed* $p = 0.81$ *and* $c = 1.2$*. A period-30 resonance tongue emanates from a Neimark–Sacker bifurcation NS (gray curve) and is bounded by saddle-node curves SN (purple). The curves C of cusp transition (cyan) and L of loop creation (black) inside the tongue are tangent at a cusp-cusp bifurcation point (black dot). Panels* (b)–(d) *are for* $(\eta, \beta) = (-0.153, 1.0005)$ *and show enlargements of Figure* 18*. The period-30 saddle* p_0 *(green cross) and its unstable manifold* $W^u(p_0)$ *(red curve) are shown in panel* (b) *near* C_0*. They map to* p_1 *(green cross) and* $W^u(p_1)$ *(red curve) under g, which are shown near* C_1 *in panels* (c) *and the enlargement of* (d)*.*

surrounding a period-2 orbit (not shown). Also shown are the curves J_0, $\widehat{J}_0$, and J_1; see also Figure 21 for the global arrangements of these curves. On the scale of Figure 18 it appears that the system is very close to a quadratic tangency of $W^u(\Gamma)$ at the precusp point C_0 on J_0 and a resulting cusp-cusp singularity at C_1 on J_1. From Proposition 3.2 and Figure 4 we conclude, with the direction of cuspidal appearance of $W^u(\Gamma)$ near C_1, that we must expect an unfolding of either case VII or case VIII.

As part of our subsequent analysis of (6.1) we computed the bifurcation diagram of Figure 19(a) in the (η, β)-plane for fixed $p = 0.81$ and $c = 1.2$; see Appendix A for details of the numerical algorithms. Figure 19(a) shows a curve *NS* of Neimark–Sacker bifurcations along which a torus bifurcates. The resonance tongue of the period-30 orbit emanates from a resonance point on *NS* and is bounded by two saddle-node bifurcations of periodic orbits,

denoted *SN*. Shown in black is the curve L of loop-creation, which is very close to the cyan curve C of cusp-transition. The black dot is the codimension-two cusp-cusp point, where these two codimension-one curves are tangent. To obtain the phase portraits relating to the cusp-cusp bifurcation, one needs to identify a relevant part of the unstable manifold $W^u(\Gamma)$. These are the lower left branch of the unstable manifold $W^u(p_0)$ of the point $p_0 \in \Gamma$, and its image, the lower left branch of $W^u(p_1)$ of the point $p_1 = g(p_0) \in \Gamma$. Portions of these two local manifolds are shown in relation to J_0 and J_1 in Figure 19(b) and (c), respectively, for the case $(\eta, \beta) = (-0.153, 1.0005)$ from Figure 18.

Note from Figure 19(b) that $W^u(p_0)$ makes a number of close passes "around" C_0. This explains that Figure 19(d) shows several invariant curves locally near C_1 that all form loops as in phase portrait 31, namely, one loop for each "pass" of $W^u(p_0)$ around C_0; When parameters are changed, it appears possible to create further cusp-cusp points, namely, at tangencies of different segments of $W^u(p_0)$ with J_0 at C_0.

In Figure 20 we show the numerically obtained unfolding of the "first" such cusp-cusp bifurcation point, where $W^u(p_0)$ has a quadratic tangency with J_0 and $\widehat{J}_0$ at the precusp point C_0 when it first enters a neighborhood of C_0; see panel (a). The cusp-cusp bifurcation was determined numerically as $(\eta, \beta) \approx (-0.14439095, 1.00181158)$. The central panel shows the bifurcation diagram around this point in the rescaled (η_r, β_r)-plane, where η_r is the distance from the cusp point in the η-direction and β_r is defined as the distance in the β-direction from the curve L. In this representation the curve L appears as the horizontal η_r-axis and the cusp-cusp singularity is at the origin, and this allows for a good comparison with the two-parameter unfoldings of the normal-form setting. Notice the largely differing scales of the axes, which corresponds to the fact that the curves L (black) and C (cyan) are very close to each other; compare with Figure 19(a). The remaining curves I, T, and E are sketched in Figure 20, because they have not been found directly. Rather, they were determined indirectly by a careful examination of the phase portraits associated with the (η_r, β_r)-plane. Indeed their existence can be deduced from the existence of the respective phase portraits of $W^u(p_1)$ relative to J_1 near the cusp point C_1, which are shown in the surrounding panels of Figure 20. Overall we conclude that the unfolding of the cusp-cusp bifurcation in Figure 20 is topologically the same as case VII in Figure 16.

We finish by showing in Figure 21 a global view of the period-30 saddle orbit Γ of g with is unstable manifold $W^u(\Gamma)$ as well as the part of its stable set $W^s(\Gamma)$ that is connected to Γ. Notice how the complicated structure of homoclinic and heteroclinic orbits interacts with J_1, including the cusp point C_1. A parameter-dependent study of this structure can lead to dramatic changes in the dynamics; see [36]. The interplay between features of diffeomorphisms and the folding nature due to noninvertibility is responsible for the interest in the break-up of invariant curves of endomorphisms; see, for example, [10, 11, 24, 27, 22]. It should be clear from Figure 21 that many interesting questions remain for future research.

7. Conclusions. In this paper we have identified and analyzed a codimension-two bifurcation—the cusp-cusp bifurcation—that arises in generic planar endomorphisms f. This bifurcation is due to the feature of folding of the phase plane along the singular curve J_0. Its image J_1 generically has isolated cusp points, which are associated with precusp points on J_0 where the eigenvector of the eigenvalue zero of the Jacobian Df is tangent to the curve J_0. The cusp-cusp bifurcation occurs when an invariant curve W is tangent to J_0 exactly at a

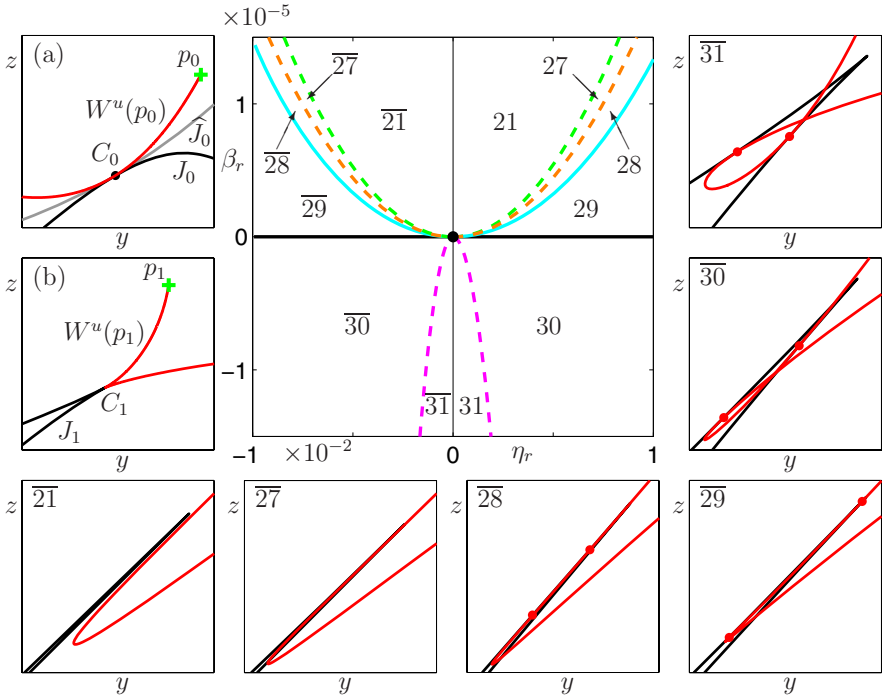


Figure 20. *The main panel is the bifurcation diagram in the rescaled* (η_r, β_r)*-plane for fixed* $p = 0.81$ *and* $c = 1.2$*, where* η_r *is the distance from the cusp point and* β_r *the* β*-distance from the curve* L *of loop creation. In this representation the cusp-cusp bifurcation point* $(\eta, \beta) \approx (-0.14439095, 1.00181158)$ *is at the origin. The solid curves* C *of cusp transition (cyan) and* L *of loop creation (black) were continued numerically, while the dashes curves* I *(magenta),* T *(orange) and* E *(green) were found by a systematic exploration of the* (η_r, β_r)*-plane. In combination with the surrounding representative phase portraits (for* $\eta_r < 0$*), the bifurcation diagram can be identified as an unfolding of case* VII*. Panel* (a) *shows a global overview for* $\eta = -0.153$ *and* $\beta = 0.999727$ *of* $W^u(p_0)$ *near* C_0*, and panel* (b) *of* $W^u(p_1)$ *near* C_1*. At this scale panels* (a) *and* (b) *are quite representative for all regions of the bifurcation diagram. The exact structure of the interaction of* $W^u(p_1)$ *with* J_1 *comes to light only in the enlargements shown in the other panels, which are for* $\eta = -0.153$ *and* $\beta = 0.999727$ *(case* $\overline{21}$*),* $\beta = 0.999729$ *(case* $\overline{27}$*),* $\beta = 0.999732$ *(case* $\overline{28}$*),* $\beta = 0.999738$ *(case* $\overline{29}$*),* $\beta = 0.9998$ *(case* $\overline{30}$*), and* $\beta = 1.0005$ *(case* $\overline{31}$*).*

precusp point C_0. This forces the images $f(W)$ of the invariant curve to have a cusp exactly at the cusp point C_1 on J_1.

Under perturbation, the cusp point C_1 on J_1 persists, but the image $f(W)$ may have many different configurations relative to C_1 and J_1. We chose to define an equivalence relation between two phase portraits in terms of ordered "generic events" along the image of the invariant curve $f(W)$ (see Definitions 3.3 and 3.4). In particular, the codimension-one bifurcations, of which there are five, can be defined rigorously. Our definition makes equivalences easy to verify in specific examples and applications. In particular, it agrees with the distinctions between the different phase portraits and the descriptions of codimension-one bifurcations that have been reported in the literature. Specifically, in the adaptive control example we presented it was straightforward to establish the equivalence of the observed phase images with the respective normal-form unfolding. The technical challenge here was to achieve numerically the

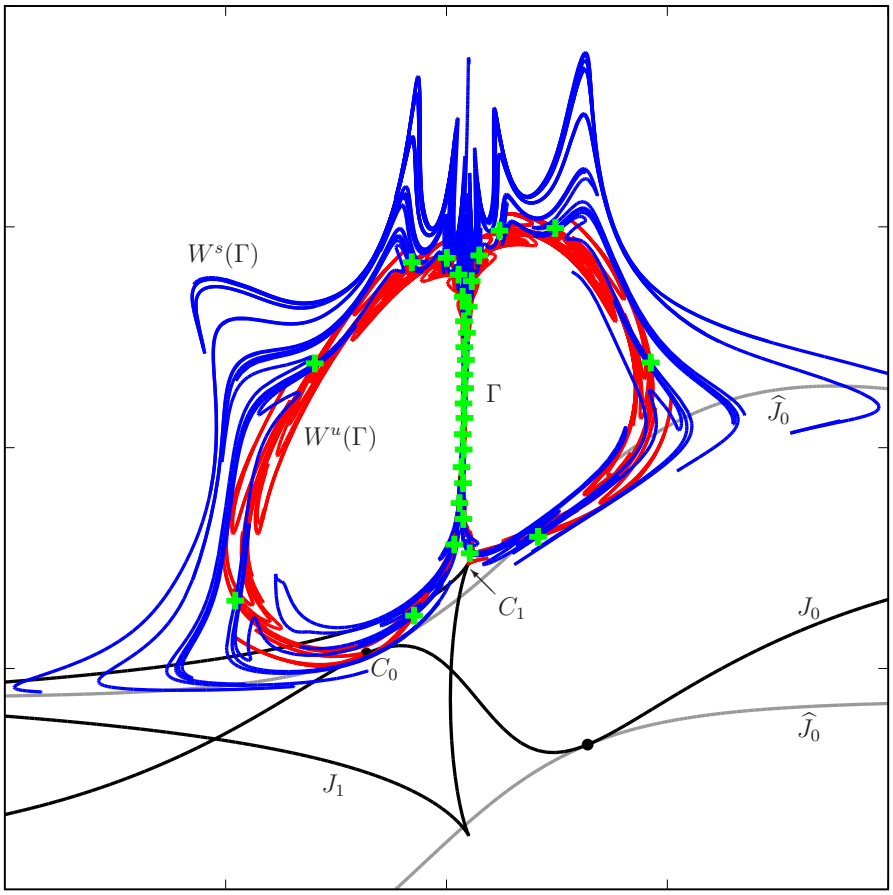


Figure 21. *Global phase portrait of the map g of (6.1) for $p = 0.81$, $c = 1.2$, $\eta = -0.153$, and $\beta = 1.0005$, showing the unstable manifold $W^u(\Gamma)$ (red curve) and also the stable set $W^s(\Gamma)$ (blue curves) of the period-30 orbit Γ (green crosses); compare with Figure 18.*

required resolution of J_1 and the unstable manifold segments.

All possible phase portraits in a local neighborhood of C_1 are organized into eight different classes of two-parameter unfoldings. Which case occurs depends on the relative positions of W, J_0, and $\widehat{J}_0$ at the bifurcation point. We presented these unfoldings and the associated 32 classes of phase portraits (with an additional 17 that may be obtained by conjugation). Our approach has been to represent the quadratic tangency of W with J_0 at C_0 by a parabola of a given steepness, which is a suitable approximation in a small enough neighborhood of the cusp-cusp point. In combination with the normal form of a map with a cusp singularity this allowed us to introduce a normal-form setting of the codimension-two bifurcation. The unfolding parameters in this setting are the horizontal and vertical positions of the maximum or minimum of the parabola (a and b in (3.1)), where J_0 and $\widehat{J}_0$ are fixed parabolas and C_0 is at the origin. The relative position of the parabola W is represented by the quadratic coefficient γ. Depending on the value of γ, we found the eight cases I–VIII. The main results in the paper are represented and illustrated by the division of the γ-line in Figure 8, the list of possible

phase portraits in Figure 9, and the images of the eight unfoldings in Figures 10 through 17. How these unfoldings manifest themselves in a practical example was demonstrated with a planar, noninvertible model from adaptive control.

Within the theoretical and numerical framework presented here, it should be possible to locate and continue the respective bifurcation curves in a variety of examples, even noninvertible maps without a cusp point on J_1. This means that one may encounter additional codimension-two bifurcations as organizing centers. As we have already briefly mentioned, new bifurcations—of codimension two and three—will also be associated with the boundaries between the different unfoldings presented here. We believe that new bifurcations of planar endomorphisms (for example, homoclinic bifurcations [36]) could be investigated using the same singularity tools utilized in this paper.

Finally, we mention endomorphisms with phase spaces of dimension larger than two, of which there are hardly any examples in the literature. The reason seems to be that the way phase space is folded is much more complicated and particularly difficult to visualize. We believe that the bifurcation analysis of such noninvertible dynamical systems can only be tackled effectively with tools from singularity theory in the spirit of the study presented here. This presents an interesting and serious challenge for future research.

Appendix A. Numerical methods.

The unstable manifolds in Figures 18 and 21 (red curves) were computed with the method from [20]. The stable sets shown in Figure 21 (blue curves) were computed with the Search Circle algorithm introduced in [6], which does not require the inverse of the map. In fact, shown are the primary manifolds of the 30 periodic points, that is, the unique pieces of the stable set (of the 30th iterate) that contains the respective periodic point; see [6] for details. The bifurcation curves for the adaptive control example Figure 19(a) were computed with the software package TBC [32]. Standard continuation techniques were used to compute the Neimark–Sacker and saddle-node bifurcation curves.

In this section we explain how to continue the curves of loop-creation bifurcation and cusp-transition bifurcation, and to detect the codimension-two cusp-cusp point. To this end we first formulate abstract bifurcation conditions and then discuss their implementation. Conditions for other bifurcations can be formulated and implemented similarly within this framework.

A.1. Abstract bifurcation conditions.

For the endomorphism f we consider the Jacobian determinant function $J(x) = \det(Df(x))$. The critical curve J_0 is implicitly defined by $J_0 \equiv \{x \in \mathbb{R}^2 : J(x) = 0\}$; see (1.2). The tangent to J_0 at a point $x \in J_0$ is then given by the perpendicular to the gradient, $(\nabla J(\mathbf{x}))^{\perp}$. Thus, the defining condition for C_0 to be a precusp point on J_0 is that

$$Df(C_0) \cdot (\nabla J(C_0))^{\perp} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A.1}$$

To specify the interaction of J_0 with the unstable manifold we assume that $p \in \mathbb{R}^2$ is a saddle point of f (or of an appropriate iterate of f) and we let

$$\begin{aligned} \boldsymbol{\alpha} : [0, \infty) &\to \mathbb{R}^2, \\ t &\mapsto \boldsymbol{\alpha}(t) \end{aligned}$$

be a parametrization of a branch of the unstable manifold $W^u(p)$, where $\boldsymbol{\alpha}(0) = p$. Furthermore, we assume that $W^u(p)$ intersects the critical curve J_0 at a point $\boldsymbol{\alpha}(t^*)$.

We can now formulate the *loop-creation condition* as

$$Df(\boldsymbol{\alpha}(t^*)) \cdot \boldsymbol{\alpha}'(t^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{A.2}$$

which ensures that $W^u(p)$ crosses J_0 tangent to the zero eigenvector, that is, the line field $\mathcal{E}$; see also Proposition 4.1.

Similarly, we formulate the *cusp-transition condition* as

$$Df(\boldsymbol{\alpha}(t^*)) \cdot (\nabla J(\boldsymbol{\alpha}(t^*)))^{\perp} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{A.3}$$

which ensures that the crossing point is exactly C_0.

At the codimension-two cusp-cusp point itself, both $\boldsymbol{\alpha}'(t^*)$ and $(\nabla J(\boldsymbol{\alpha}(t^*)))^{\perp}$ are zero eigenvectors; that is, both (A.2) and (A.3) must hold. However, requiring the latter as a condition for a cusp-cusp point turns out to be overdetermined. All we really need is either (A.2) or (A.3), along with ensuring the *parallel-vectors condition*

$$\boldsymbol{\alpha}'(t^*) \cdot \nabla J(\boldsymbol{\alpha}(t^*)) = 0. \tag{A.4}$$

A.2. Implementation. The defining conditions for all three noninvertible bifurcations described above, the loop-creation, the cusp-transition, and the cusp-cusp point, are solved numerically by Newton's method. We use the variables

$$\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k, \mu_1, \mu_2, \tag{A.5}$$

where the phase variables $\mathbf{x}_i \in \mathbb{R}^2$ and the parameters $\boldsymbol{\mu} = (\mu_1, \mu_2) \in \mathbb{R}^2$. In total these are $2k + 4$ scalar variables.

Each system of equations that we construct must, therefore, satisfy $2k+4$ scalar equations. We first list the relevant equations.

1. The condition that $\mathbf{x}_0$ is a period-q point gives the two scalar equations:

$$f^q_{\boldsymbol{\mu}}(\mathbf{x}_0) - \mathbf{x}_0 = \mathbf{0}. \tag{A.6}$$

2. We require that $\mathbf{x}_1$ be in the unstable eigenspace, which is the common approximation for the condition that $\mathbf{x}_1 \in W^u_{\mathrm{loc}}(\mathbf{x}_0)$; see [25, 35] for proofs of why this also holds for noninvertible maps. This gives the single scalar equation

$$\frac{\mathbf{x}_1 - \mathbf{x}_0}{\|\mathbf{x}_1 - \mathbf{x}_0\|} \cdot \mathbf{v}^{\perp} = 0, \tag{A.7}$$

   where $\mathbf{v}^{\perp}$ is a unit vector perpendicular to the unstable eigenvector $\mathbf{v}$ of $\mathbf{x}_0$.
3. The condition that $\mathbf{x}_1, \ldots, \mathbf{x}_k$ is an orbit for f^q_{μ} (which lies on $W^u(\mathbf{x}_0)$ if $\mathbf{x}_1$ does) gives rise to the $2k - 2$ scalar equations

$$f^q_{\boldsymbol{\mu}}(\mathbf{x}_i) - \mathbf{x}_{i+1} = \mathbf{0}, \quad i = 1, \ldots, k-1. \tag{A.8}$$

   Note that q needs to be replaced here with $2q$ if the unstable eigenvalue is negative.

4. The condition $\mathbf{x}_k = C_0$ gives rise to the two scalar equations

$$Df^q_{\boldsymbol{\mu}}(\mathbf{x}_k) \cdot (\nabla(J_\mu(\mathbf{x}_k)))^\perp = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A.9}$$

The partial derivatives in the gradient are computed using a forward difference quotient.

5. The condition that $W^u_{\rm loc}(\mathbf{x}_0)$ undergoes a loop-creation associated with $\mathbf{x}_k$ gives rise to the two scalar equations

$$Df^q_{\boldsymbol{\mu}}(\mathbf{x}_k) \cdot \mathbf{t}_k = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{A.10}$$

Here, the vector $\mathbf{t}_k = Df^{q(k-1)}(\mathbf{x}_1) \cdot (\mathbf{x}_1 - \mathbf{x}_0)/\|Df^{q(k-1)}(\mathbf{x}_1) \cdot (\mathbf{x}_1 - \mathbf{x}_0)\|$ is an approximation to a unit tangent vector of the unstable manifold at $\mathbf{x}_k$. This equation is intended to be used with (A.7) so that $\mathbf{x}_1 - \mathbf{x}_0$ is an eigenvector at $\mathbf{x}_0$ and thus an approximation to a tangent vector to the unstable manifold at $\mathbf{x}_1$ *if* $\mathbf{x}_1$ is close to $\mathbf{x}_0$.

6. The condition that the tangent to $W^u_{\rm loc}(\mathbf{x}_0)$ and the gradient of the Jacobian function are perpendicular gives rise to the single scalar equation

$$(\nabla(J_{\boldsymbol{\mu}}(\mathbf{x}_k))) \cdot \mathbf{t}_k = 0. \tag{A.11}$$

7. We also require the standard pseudo-arclength continuation condition that convergence from the initial guess for Newton's method is perpendicular to the pseudo tangent $\mathbf{T}$ to the bifurcation curve being computed. This gives rise to the single scalar equation

$$\mathbf{T} \cdot ((\mathbf{x}_k, \mu_1, \mu_2) - (\mathbf{x}_k, \mu_1, \mu_2)_0) = 0. \tag{A.12}$$

Here we approximate the tangent vector $\mathbf{T}$ by the direction through the last two computed points. In our case, we project the tangent vector to the four-dimensional product of phase and parameter space, where $\mathbf{x}_k$ is considered to be the phase space "representative" for the orbit; the zero subscript of the second tuple indicates that this is the initial guess provided to Newton's method.

For each of the three bifurcations from section A.1 we require that the $2k+1$ equations (A.6), (A.7), and (A.8) are satisfied. Furthermore, these bifurcations are defined by additional requirements as follows:

- The codimension-one cusp-transition bifurcation is determined by (A.9) and (A.12),
- the codimension-one loop-creation bifurcation is determined by (A.10) and (A.12), and
- the codimension-two cusp-cusp point is determined by (A.11) along with either (A.9) or (A.10).

Acknowledgment. We thank Yannis Kevrekidis for helpful discussions.

REFERENCES

[1] R. A. Adomaitis and I. G. Kevrekidis, *Noninvertibility and structure of basins of attraction in a model adaptive control system*, J. Nonlinear Sci., 1 (1991), pp. 95–105.

[2] A. Agliari, *Global bifurcations in the basins of attraction in noninvertible maps and economic applications*, Nonlinear Anal., 47 (2000), pp. 5241–5252.
[3] A. Agliari, L. Gardini, and C. Mira, *On the fractal structure of basin boundaries in two-dimensional noninvertible maps*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 13 (2003), pp. 1767–1785.
[4] V. I. Arnol′d, *Catastrophe Theory*, 3rd ed., Springer-Verlag, Berlin, 1992.
[5] H. W. Broer, M. Golubitsky, and G. Vegter, *The geometry of resonance tongues: A singularity theory approach*, Nonlinearity, 16 (2003), pp. 1511–1538.
[6] J. P. England, B. Krauskopf, and H. M. Osinga, *Computing one-dimensional stable manifolds of planar maps without the inverse*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 161–190.
[7] J. P. England, B. Krauskopf, and H. M. Osinga, *Bifurcations of stable sets in noninvertible planar maps*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (2005), pp. 891–904.
[8] C. E. Frouzakis, R. A. Adomaitis, and I. G. Kevrekidis, *An experimental and computational study of subcriticality, hysteresis and global dynamics for a model adaptive control system*, Comp. Chem. Engng., 120 (1996), pp. 1029–1034.
[9] C. E. Frouzakis, R. A. Adomaitis, I. G. Kevrekidis, M. P. Golden, and B. E. Ydstie, *The structure of basin boundaries in a simple adaptive control system*, in Chaotic Dynamics: Theory and Practice, T. Bountis, ed., Plenum Press, New York, 1992, pp. 195–210.
[10] C. E. Frouzakis, L. Gardini, I. G. Kevrekidis, G. Millerioux, and C. Mira, *On some properties of invariant sets of two-dimensional noninvertible maps*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 7 (1997), pp. 1167–1194.
[11] C. E. Frouzakis, I. G. Kevrekidis, and B. B. Peckham, *A route to computational chaos revisited: Noninvertibility and the breakup of an invariant circle*, Phys. D, 177 (2003), pp. 101–121.
[12] M. Golubitsky and D. G. Shaeffer, *Singularities and Groups in Bifurcation Theory*, Vol. 1, Springer-Verlag, New York, 1985.
[13] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York, Berlin, 1983.
[14] I. Gumowski and C. Mira, *Dynamique Chaotique*, Cepadues Éditions, Toulouse, France, 1980.
[15] I. Gumowski and C. Mira, *Recurrences and Discrete Dynamic Systems*, Springer-Verlag, New York, 1980.
[16] M. Hénon, *A two-dimensional mapping with a strange attractor*, Comm. Math. Phys., 50 (1976), pp. 69–77.
[17] K. Ikeda, *Multiple-valued stationary state and its instability of the transmitted light from a ring cavity system*, Optics Communications, 30 (1979), pp. 257–261.
[18] K. Josić and E. Sander, *The structure of synchronization sets for noninvertible systems*, Chaos, 14 (2004), pp. 249–262.
[19] H. Kitajima, H. Kawakami, and C. Mira, *A method to calculate basin bifurcation sets for a two-dimensional noninvertible map*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 10 (2000), pp. 2001–2014.
[20] B. Krauskopf and H. M. Osinga, *Growing 1D and quasi-2D unstable manifolds of maps*, J. Comput. Phys., 146 (1998), pp. 406–419.
[21] B. Krauskopf and C. Rousseau, *Codimension-three unfoldings of reflectionally symmetric planar vector fields*, Nonlinearity, 10 (1997), pp. 1115–1150.
[22] E. N. Lorenz, *Computational chaos—a prelude to computational instability*, Phys. D, 35 (1989), pp. 299–317.
[23] V. Maistrenko, Yu. Maistrenko, and I. Sushko, *Noninvertible two-dimensional maps arising in radiophysics*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 4 (1996), pp. 383–400.
[24] V. Maistrenko, Yu. Maistrenko, and E. Mosekilde, *Torus breakdown in noninvertible maps*, Phys. Rev. E (3), 67 (2003), 046215.
[25] R. P. McGehee and E. Sander, *A new proof of the stable manifold theorem*, Z. Angew. Math. Phys., 47 (1996), pp. 497–513.
[26] C. Mira, J.-P. Carcassés, G. Millérioux, and L. Gardini, *Plane foliation of two-dimensional noninvertible maps*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 6 (1996), pp. 1439–1462.
[27] C. Mira, L. Gardini, A. Barugola, and J.-C. Cathala, *Chaotic Dynamics in Two-Dimensional Noninvertible Maps*, World Sci. Ser. Nonlinear Sci. Ser. A Monogr. Treatises 20, World Scientific, Singapore, 1996.

[28] C.-H. Nien, *The dynamics of planar quadratic maps with nonempty bounded critical set*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 8 (1998), pp. 95–105.
[29] C.-H. Nien and F. J. Wicklin, *An algorithm for the computation of preimages in noninvertible mappings*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 8 (1998), pp. 415–422.
[30] J. Palis and W. de Melo, *Geometric Theory of Dynamical Systems*, Springer-Verlag, New York, Berlin, 1982.
[31] B. B. Peckham, *Real perturbation of complex analytic families: Points to regions*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 8 (1998), pp. 73–93.
[32] B. B. Peckham, *To Be Continued . . . , Continuation and Phase Space Software for Discrete Dynamical Systems*, http://www.d.umn.edu/~bpeckham/tbc_home.html (1986–present).
[33] B. B. Peckham and J. Montaldi, *Real continuation from the complex quadratic family: Fixed-point bifurcation sets*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 10 (2000), pp. 391–414.
[34] R. Rico-Martinez, R. A. Adomaitis, and I. G. Kevrekidis, *Noninvertibility in neural networks*, Comp. Chem. Engng., 24 (2000), pp. 2417–2433.
[35] E. Sander, *Hyperbolic sets for noninvertible maps and relations*, Discrete Contin. Dynam. Systems, 5 (1999), pp. 339–357.
[36] E. Sander, *Homoclinic tangles for noninvertible maps*, Nonlinear Anal., 41 (2000), pp. 259–276.
[37] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Vol. 1, 2nd ed., Publish or Perish, Houston, TX, 1979.
[38] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, Perseus Books, Reading, MA, Cambridge, MA, 1994.
[39] H. Whitney, *On singularities of mappings of Euclidean spaces.* I. *Mappings of the plane into the plane*, Ann. of Math. (2), 62 (1955), pp. 374–410.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 2, pp. 441–456

Mathematical Modeling of the GnRH Pulse and Surge Generator*

Frédérique Clément† and Jean-Pierre Françoise‡

Abstract. We propose a mathematical model allowing for the alternating pulse and surge pattern of gonadotropin releasing hormone (GnRH) secretion. The model is based on the coupling between two systems running on different time scales. The faster system corresponds to the average activity of GnRH neurons, while the slower one corresponds to the average activity of regulatory neurons. The analysis of the slow/fast dynamics exhibited within and between both systems allows for explaining the sequence of different secretion patterns (slow oscillations, fast oscillations, and periodical surge) of GnRH secretion as a hysteresis loop. Specifications on the model parameter values are derived from physiological knowledge in terms of amplitude, frequency, and plateau length of oscillations. The behavior of the model is finally illustrated by numerical simulations reproducing natural ovarian cycles and either direct or indirect actions of ovarian steroids on GnRH secretion.

Key words. coupled oscillators, hysteresis, fast-slow dynamics, amplitude and frequency control, ovulation, neuroendocrinology, GnRH pulsatility, GnRH surge

AMS subject classifications. 34c15, 34c23, 34c26, 92B05

DOI. 10.1137/060673825

1. Introduction.

1.1. Endocrine background.

The reproductive function involves tightly and finely controlled processes. The reproductive axis, usually called the gonadotrope axis, includes the hypothalamus, within the central nervous system, the pituitary gland, and the gonads (ovaries in females and testes in males). Specific hypothalamic neurons secrete the gonadotropin releasing hormone (GnRH) in a pulsatile manner. The pulsatile GnRH secretion pattern ensues from the synchronization of the secretory activity of individual GnRH neurons. The release of GnRH into the pituitary portal blood induces the secretion of the luteinizing hormone (LH) and follicle stimulating hormone (FSH) by the pituitary gland. The changes in the frequency of GnRH pulses (between 1 pulse per hour and 1 pulse every 6 hours in the course of an ovarian cycle) has a fundamental role in the differential control of the secretion of both gonadotropins: the secretion of LH is enhanced by higher frequencies, while that of FSH is enhanced by lower frequencies [13]. On the gonadic level, FSH and LH sustain germ cell production and hormone secretion. In turn, hormones secreted by the gonads modulate the secretion of GnRH, LH, and FSH within entangled feedback loops.

In females, the frequency of GnRH pulses is subject to the control exerted by ovarian

*Received by the editors October 31, 2006; accepted for publication (in revised form) by B. Ermentrout March 11, 2007; published electronically June 15, 2007.
http://www.siam.org/journals/siads/6-2/67382.html
†Unité de recherche, INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt BP 105, 78153 Le Chesnay Cedex, France (frederique.clement@inria.fr).
‡Laboratoire Jacques-Louis Lions, Université P.-M. Curie, Paris 6, UMR 7598 CNRS, BP 187, 75252 Paris Cedex 05, France (jpf@ccr.jussieu.fr).

steroids—estradiol and progesterone. Progesterone slows down GnRH pulsatility [12], which leads to a slower frequency during the luteal phase (when progesterone levels are high) compared to the follicular phase [7]. On the contrary, estradiol speeds up GnRH pulse frequency, but at the expense of a decrease in pulse amplitude [6], so that the whole feedback effect on rate secretion is rather inhibitory (negative estradiol feedback).

The GnRH secretion pattern alters dramatically once per ovarian cycle, resulting in the GnRH surge characterized by massive continuous release of GnRH in response to increasing levels of estradiol [10] (positive estradiol feedback). The GnRH surge triggers the LH surge, which is responsible for ovulation, leading to the release of fertilizable oocytes.

The estradiol signal is conveyed to GnRH neurons by regulatory neurons (also designed as interneurons) either directly or after other neuronal relays. Transmission from interneurons calls to many different neurotransmitters (see review in [8, 16]). The balance between stimulatory and inhibitory signals emanating from interneurons controls the behavior of the GnRH network [8].

Recently, an especially interesting type of regulatory neurons has been discovered. Kisspeptin neurons act on GnRH neurons via the G-protein coupled receptor GPR54 [5]. They are very good candidates for relaying both positive and negative estradiol feedback, since they react to estradiol in opposite ways according to the anatomic area of the hypothalamus where they lie.

1.2. Model objectives. We aim at formulating a phenomenological, data-driven model of GnRH secretion. This paper focuses on the coupling between the GnRH neuron network and the regulatory interneuron network. Each network is represented by the behavior of a single average neuron. The key point in the design of the model consists in entering as reasonable input 2 coupled systems (a slow one with a faster one) to generate a definite sequence of events in the model output: GnRH secretion. This coupling yields a 3 time-scaled model, which is able to capture not only the cyclic transition from a pulsatile to a surge secretion pattern of GnRH but also the increase in the pulsatility frequency between the luteal and follicular phases. It also separates a specific dynamical state corresponding to pulsatility resumption after the surge. Besides, parametrization of the model is subject to physiological specifications expressed as constraints on the GnRH output and allows us to reproduce the direct (on the GnRH network) or indirect (via the regulatory network) effects of ovarian steroid hormones (estradiol and progesterone) on GnRH secretion.

In summary, we aim at reproducing a synthetic mathematical representation of the GnRH secretion pattern, fitting available observations—in agreement with schematic "hand-made" representations such as that proposed by Herbison (top of Figure 4 in [8]).

The paper is organized as follows. In section 2, we introduce the equations of the model. We comment on the numerical simulations in section 3 to motivate the analysis of the bifurcations in a model-derived 2-scaled system in section 4. We finally address the question of amplitude and frequency control in section 5.

2. Model design and analysis.

2.1. Coupling oscillatory neuronal dynamics. We consider the following four-dimensional (4D) dynamical system:

$$\epsilon\delta\dot{x} = -y + f(x), \tag{2.1a}$$

$$\epsilon\dot{y} = a_0 x + a_1 y + a_2 + cX, \tag{2.1b}$$

$$\epsilon\gamma\dot{X} = -Y + F(X), \tag{2.1c}$$

$$\dot{Y} = b_0 X + b_1 Y + b_2, \tag{2.1d}$$

$$z(t) = \chi_{\{y(t)>y_s\}}.$$

Equations (2.1a) and (2.1b) correspond to a fast system representing an average GnRH neuron, while (2.1c) and (2.1d) correspond to the slower system representing an average regulatory neuron. The x, X variables represent the neuron electrical activities (action potential), while the y, Y variables relate to ionic and secretory dynamics. The fast variables are assumed to have two stable stationary points separated by a saddle. Their bistability is accounted for by the cubic functions $f(x)$ and $F(x)$. The intrinsic dynamics of the slow variables follows a growth law of very small velocity ($a_1 \ll 1$). In each system, the fast and slow variables act reciprocally on each other. The coupling between both systems is mediated through the unilateral influence of the slow regulatory neurons onto the fast GnRH ones (cX term in (2.1b)). The coupling term aggregates the global balance between inhibitory and stimulatory neuronal inputs onto the GnRH neurons. The global system exhibits 3 time scales given by $\epsilon\delta$, ϵ, and 1. Constant γ is close to 1.

In many cell types, the reaching of threshold intracellular calcium concentrations is known to trigger secretion. As far as GnRH secretion is concerned, much evidence for the inducing effect of calcium has also been gathered (see discussion in [17] for details). Hence, we associate GnRH secretion to a threshold ionic activity $y = y_s$ and finally keep as representative GnRH signal $z(t) = \chi_{\{y(t)>y_s\}}$.

2.2. Mechanisms underlying the pulse to surge transition.

2.2.1. Bifurcations in the fast GnRH system. System (2.1a)–(2.1d) can be analyzed in the general setting of fast-slow dynamics. The slow variable X then enters the fast system (2.1a)–(2.1b) as a parameter. Bifurcations may thus arise as the X parameter varies. The fast system (after changing time t into ϵt) also exhibits fast-slow dynamics features due to the δ time scale:

$$\delta\dot{x} = -y + f(x), \tag{2.2a}$$

$$\dot{y} = a_0 x + a_1 y + a_2 + cX, \tag{2.2b}$$

$$\text{with } f(x) = \nu_0 x^3 + \nu_1 x^2 + \nu_2 x. \tag{2.2c}$$

The fast nullcline $\dot{x} = 0$ has a cubic shape and the stationary points are obtained as intersections of this cubic nullcline with the other nullcline $\dot{y} = 0$, a straight line which moves in the plane depending on the values of the slow variable X. Without loss of generality, we assume that $\lim_{x\to-\infty} f(x) = +\infty$. We also assume that the parameters (a_0, a_1) are chosen in such a way that there is one single stationary point whatever the value of X may be (see Figure 1). Intersection points of the cubic nullcline with the line nullcline are solutions of the equation

$$f(x) + \frac{a_0}{a_1}x + \frac{(cX + a_2)}{a_1} = 0. \tag{2.3}$$

Let $x_0 = s(X)$ denote the unique solution to (2.3). We now derive from classical arguments the nature of the stationary point $(s(X), f(s(X)))$. The eigenvalues of the linearized system are solutions of the characteristic polynomial

$$\lambda^2 - \left[\frac{f'(x_0)}{\delta} + a_1\right]\lambda + \frac{1}{\delta}[a_0 + a_1 f'(x_0)] = 0.$$

If the following conditions are fulfilled, two roots are complex conjugates with a negative real part:

$$\begin{cases} \dfrac{f'(x_0)}{\delta} + a_1 < 0, \\ \dfrac{a_0}{\delta} + \dfrac{a_1}{\delta} f'(x_0) > 0, \\ \left[\dfrac{f'(x_0)}{\delta} - a_1\right]^2 - 4\dfrac{a_0}{\delta} < 0. \end{cases}$$

Let x_0^- and x_0^+ denote the two roots of the equation

$$\frac{f'(x)}{\delta} + a_1 = 0.$$

Then, if $s(X) < x_0^-$, the stationary point is a stable focus. If X varies in such a way that $x_0 = s(X)$ crosses the value x_0^-, the system (2.2a)–(2.2b) undergoes a Hopf bifurcation, and the stationary point becomes unstable as a stable limit cycle appears. This stable limit cycle disappears when $x_0 > x_0^+$ and the stationary point becomes stable again. This reasoning is illustrated in Figure 1.

2.2.2. Limit cycle of the slow system. The dynamics of the X variable is such that, if the straight-line nullcline sweeps the (x, y) phase plane (from left to right and right to left) periodically, the Hopf bifurcations occur themselves periodically.

The qualitative analysis of the slow dynamics is analogous to that of the fast dynamics. We assume that the parameters (b_0, b_1, b_2) are fixed so that there is one single stationary point $(X_0, F(X_0))$ which is an unstable focus (allowing the slow system to display a stable limit cycle):

$$\begin{cases} \dfrac{F'(X_0)}{\gamma} + b_1 > 0, \\ \dfrac{b_0}{\gamma} + \dfrac{b_1}{\gamma} F'(X_0) > 0, \\ \left[\dfrac{F'(X_0)}{\gamma} - b_1\right]^2 - 4\dfrac{b_0}{\gamma} < 0, \end{cases}$$

where $F(x) = \mu_0 X^3 + \mu_1 X^2 + \mu_2 X$. The oscillation associated to this limit cycle is of relaxation type (as in the van der Pol system). The dynamics along the limit cycle displays slow and fast parts alternatively (see Figure 2).

2.2.3. Dynamics of the coupled system. The different phases of the limit cycle exhibited by the slow regulatory system (see Figure 2) drive the global behavior of the fast GnRH system (see Figure 3):

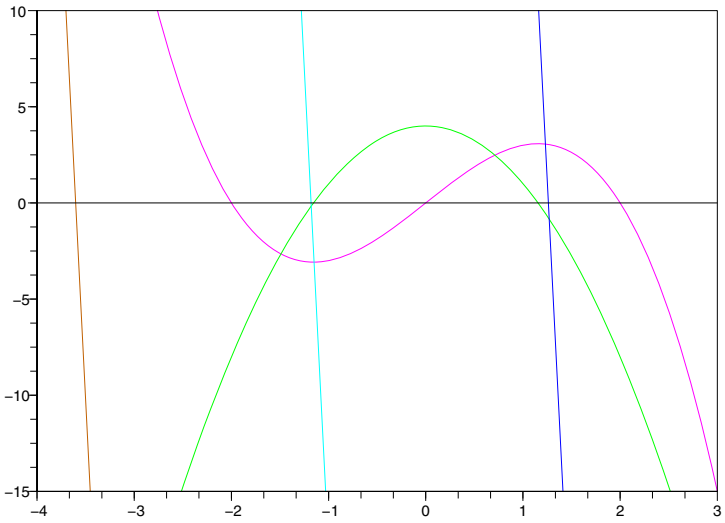


Figure 1. *Nullclines of the fast system. The pink solid line corresponds to the* $\dot{x} = 0$ *nullcline and the cyan solid line to the* $\dot{y} = 0$ *nullcline for* $X = 0$ *(no coupling). The green solid line corresponds to the curve* $y = f'(x)$*. The* (x_0^-, x_0^+) *roots of* $f'(x) = -a_1\delta \approx 0$ *roughly correspond to the intersection points of the green line with the black horizontal line. As the value of* X *varies, the cyan nullcline sweeps the* x*-axis within or outside the* $[x_0^-, x_0^+]$ *interval, between extreme positions delimited by the brown and blue straight lines. Clicking on the above image displays the accompanying animation* (67382_01.mov [3.03MB]).

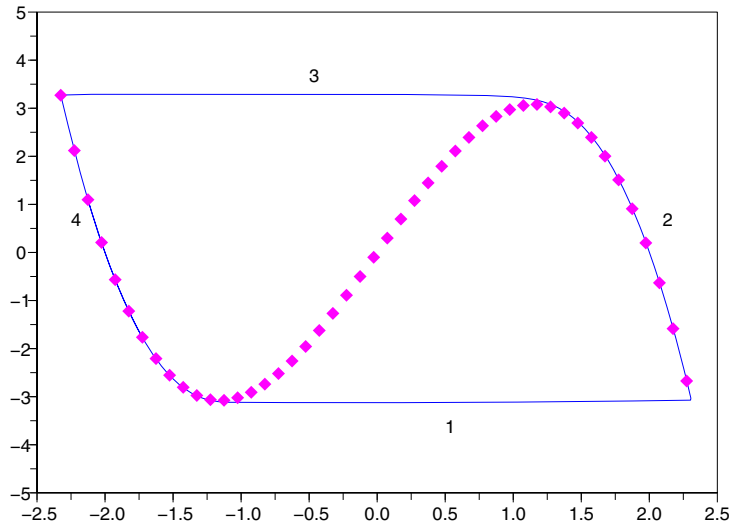


Figure 2. *Slow* (X, Y) *limit cycle. The blue line corresponds to the limit cycle in the* (X, Y) *plane. The slow parts of the cycle correspond to branches of the cubic nullcline represented by pink diamonds (phases* 2 *and* 4*), and the fast parts correspond to the jumps from one branch of the cubic to the other (phases* 1 *and* 3*).*

Table 1
Numerical values of the model parameters.

ε	1/40	δ	1/80	γ	1
a_0	1	a_1	0.01	a_2	1.18
b_0	1	b_1	0.04	b_2	1
ν_0	-1	ν_1	0	ν_2	4
μ_0	-1	μ_1	0	μ_2	4
c	1.05	y_s	2.8		

1. the first phase of the cycle, where X increases abruptly, corresponds to the ascending part of the surge;
2. the second phase of the cycle, where X decreases slowly, corresponds to the duration of the surge;
3. the third phase of the cycle, where X decreases abruptly, corresponds to the decreasing part of the surge;
4. the fourth and most lasting phase of the cycle, where X increases slowly, encompasses two different GnRH secretion patterns:
   (a) as long as $X \geq x_0^+$, the GnRH level is almost constant, since the fast system admits a stable steady state and the sweeping dynamics of the straight line $\dot{y} = 0$ lies in a slow phase; hence this phase explains the existence of a plateau after the surge (we provide as additional material a video illustrating that point);
   (b) when $X < x_0^+$ the pulsatility of GnRH is recovered and the pulse frequency increases with X.

3. Numerical simulations. The numerical values of the model parameters can be constrained by physiological specifications regarding the features (frequency, amplitude, and plateau length) of the GnRH secretory patterns [7]. The GnRH output should be characterized by

- the pulse duration to pulse frequency ratio,
- the pulse amplitude to surge amplitude ratio,
- the surge frequency to pulse frequency ratio, and
- the surge duration to whole-ovarian-cycle duration.

A set of parameters subject to such constraints is displayed in Table 1, and the corresponding model outputs are illustrated in Figures 3, 4, and 5.

The main qualitative features captured by the model consist in

1. the cyclic transition from a pulse to a surge secretion pattern, which occurs after a short transitory period and seems not to be subject to initial conditions;
2. a delay before resumption of pulsatility; and
3. the increase in pulse frequency from the luteal (postsurge) phase to the follicular (presurge) phase (see Figure 6).

If the first two properties (pulse to surge alternating and pulsatility resumption delay) were expected from the study derived in section 2, the third one (frequency increase) remains unexplained at this point, even if it is consistent with endocrinological data. We now explain the mechanisms underlying those properties from the bifurcation analysis of a 2-scaled system derived from the model.

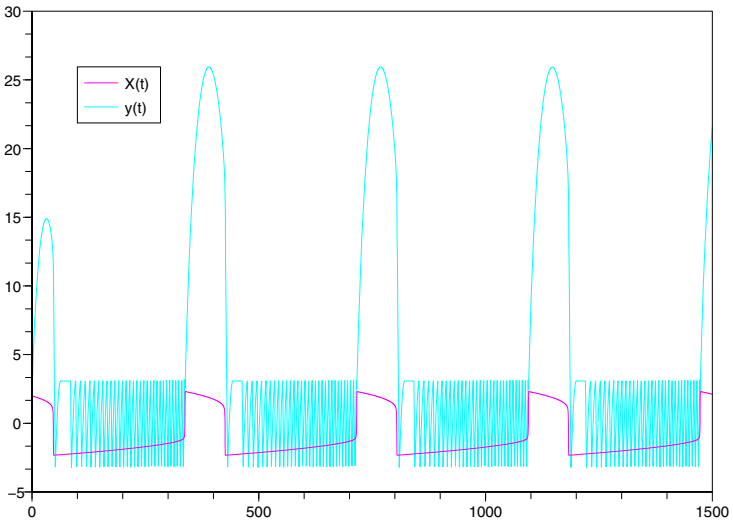


Figure 3. *Outputs from the coupled systems. Pink line:* $X(t)$*. Cyan line:* $y(t)$*.*

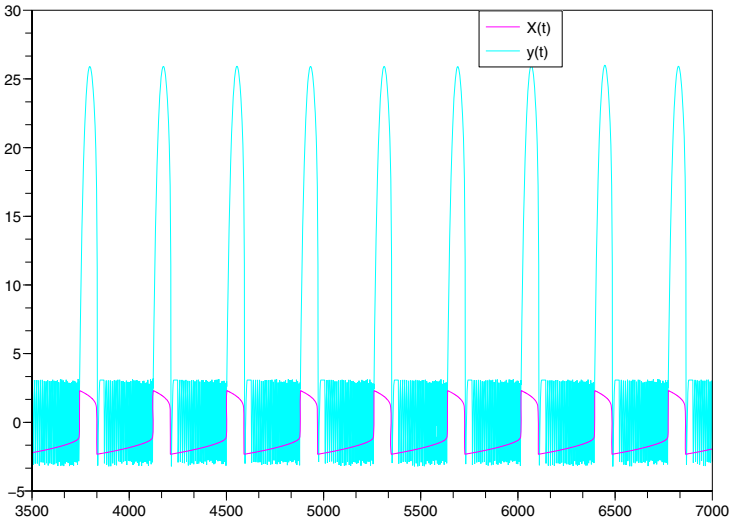


Figure 4. *Long-term outputs from the coupled systems. Pink line:* $X(t)$*. Cyan line:* $y(t)$*.*

4. Bifurcation analysis of a model-derived 2-scaled system. Our approach is adapted from the "geometrical dissection" [1] that has been successfully applied to several models in Computational Neurosciences, especially those dealing with bursting oscillations.

In classical slow/fast systems, the slow variable is "frozen" and intervenes as a parameter in the study of the bifurcations of the fast system. In a similar way, we consider the fast three-dimensional (3D) system with two time scales,

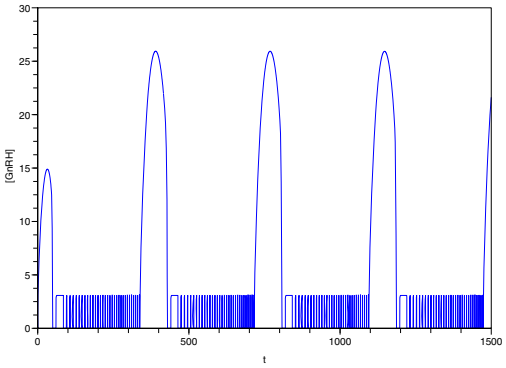


Figure 5. *GnRH secretion pattern. The secretory activity of GnRH neurons occurs for threshold ionic activity* $z(t) = \chi_{\{y(t)>y_s\}}$.

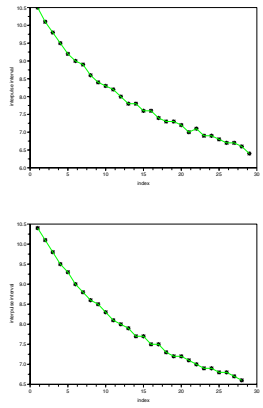


Figure 6. *Decrease in interpulse intervals from the luteal to the follicular phase (top: from time* 100 *to time* 337*; bottom: from time* 475 *to time* 710*).*

$$\delta \dot{x} = -y + f(x), \tag{4.1a}$$

$$\dot{y} = a_0 x + a_1 y + a_2 + cX, \tag{4.1b}$$

$$\gamma \dot{X} = -Y + F(X), \tag{4.1c}$$

where Y acts as a varying parameter. Below, we restrict our study to the case where the other parameter values are fixed and close to those of Table 1. This fast system breaks into an independent, one-dimensional (1D) system (4.1c), and a two-dimensional (2D) system (4.1a)–(4.1b) forced by the 1D system.

Depending on the Y value, $Y = F(X)$ may have one, two, or three roots. Accordingly, (4.1c) displays either one of the two possible attracting points (denoted, respectively, by $X_-(Y)$ and $X_+(Y)$) or both of them separated by a repulsive point (denoted by $X_0(Y)$). A saddle-node bifurcation occurs for the values of Y corresponding to the ordinates of the local extrema of the cubic function: $(X, Y) = \pm(2/\sqrt{3}, 16/(3\sqrt{3}))$. The whole analysis of Y-driven bifurcations in the 3D system (4.1a)–(4.1c) is summarized in Table 2. Assuming $a_1 \approx 0$, we introduce $x_-(Y)$, $x_0(Y)$, and $x_+(Y)$ as the x values associated, respectively, with $X_-(Y)$, $X_0(Y)$, and $X_+(Y)$, from $x_i = -(a_2 + cX_i)/a_0$, $i = -, 0, +$.

- For $-\infty < Y < -16/(3\sqrt{3})$, equation (4.1c) admits the attractive node $X_+(Y)$ as single stationary point, and $x_+(Y) > x_0^+ \approx 1.15$ (cf. legend of Figure 1). It ensues from the study of subsection 2.2.1 that the 2D system (4.1a)–(4.1b) exhibits a stable focus, so that the 3D system (4.1a)–(4.1c) displays one single stable stationary point.
- When $Y = -16/(3\sqrt{3})$, $X_+(Y) = 4\sqrt{3}$, $x_+(Y) \approx -3.60$, equation (4.1c) undergoes a saddle-node bifurcation. The $X_0(Y)$ saddle and the $X_-(Y)$ attractive node appear and dissociate from the coincident point $X_0(Y) = X_-(Y) = -2/\sqrt{3}$, associated with $x_0(Y) = x_-(Y) = -0.032$. Both $x_0(Y)$ and $x_-(Y)$ belong to the $\left[x_0^-; x_0^+\right]$ interval (cf. legend of Figure 1), for which the 2D system displays an unstable focus and a

Table 2

Bifurcations in the 3D system (4.1a)–(4.1c) according to the Y value. 1D corresponds to (4.1c), 2D to system (4.1a)–(4.1b), and 3D to system (4.1a)–(4.1c). $X_-(Y)$, $X_0(Y)$, and $X_+(Y)$ are the three possible roots of $F(X) = Y$, and $x_-(Y)$, $x_0(Y)$, and $x_+(Y)$ are the corresponding x values, with $x_i = -(a_2 + cX_i)/a_0$, $i = -, 0, +$. The grey superimposed oriented line describes the hysteresis loop in the original 4D system (2.1a)–(2.1d). Numbers in bold font correspond to exact values, and numbers in standard font correspond to approximate values. The • symbol stands for saddle-node bifurcation, the ⋆ symbol for Hopf bifurcation, and the ⊙ symbol for saddle-node bifurcation of periodic orbits. The circled numbers refer to the sequence of phases listed in subsection 2.2.3.

Y	↗	$-\frac{16}{3\sqrt{3}}$	↗	-0.096	↗	2.1	↗	$\frac{16}{3\sqrt{3}}$	↗
$X_+(Y)$	↘	$\frac{4}{\sqrt{3}}$	↘	2.01	↘	1.65	↘	$\frac{2}{\sqrt{3}}$	
$x_+(Y)$	↗	-3.6	↗	-3.29	↗	-2.91	↗	-2.39	
1D	attractive node		attractive node				attractive node	•	
2D	stable focus		stable focus	②			stable focus		
3D	stable		stable				stable		
	stationary point		stationary point				stationary point		
$X_0(Y)$		$-\frac{2}{\sqrt{3}}$	↗	-0.024	↗	0.57	↗	$\frac{2}{\sqrt{3}}$	
$x_0(Y)$		-0.032	↘	$-\frac{2}{\sqrt{3}}$	↘	-1.78	↘	-2.39	
1D		•	repulsive node		repulsive node		repulsive node	•	
2D			unstable focus	⋆	stable focus		stable focus		
			stable limit cycle						
3D	①	⊙	hyperbolic periodic orbit		hyperbolic		hyperbolic		③
			+unstable stationary point		stationary point		stationary point		
$X_-(Y)$		$-\frac{2}{\sqrt{3}}$	↘	-1.99	↘	-2.22	↘	$-\frac{4}{\sqrt{3}}$	↘
$x_-(Y)$		-0.032	↗	0.91	↗	$\frac{2}{\sqrt{3}}$	↗	1.24	↗
1D		•	attractive node		attractive node		attractive node		attractive node
2D			unstable focus	④	unstable focus	⋆	stable focus		stable focus
			stable limit cycle		stable limit cycle				
3D		⊙	attractive periodic orbit		attractive periodic orbit		stable		stable
			+ hyperbolic stationary point		+ hyperbolic stationary point		stationary point		stationary point

stable limit cycle. The 3D system thus exhibits simultaneously an attractive stationary point associated to $X_+(Y)$, an unstable stationary point together with a hyperbolic periodic orbit associated to $X_0(Y)$, and a hyperbolic stationary point together with an attractive periodic orbit associated to $X_-(Y)$. Hence $Y = -16/(3\sqrt{3})$ is also a bifurcation point for the 3D system, which triggers both a saddle-node bifurcation of periodic orbits and a saddle-node bifurcation of stationary points.

- As Y keeps increasing, $x_0(Y)$ reaches the value of the abscissa of the $f(x)$ cubic local minimum, where $x_0(Y) = -2/\sqrt{3} \equiv x_0^-$, $X_0(Y) \approx -0.024$, and $Y \approx -0.096$. The 2D system undergoes a Hopf bifurcation. In the 3D system, the hyperbolic periodic orbit and the unstable stationary point coalesce into a hyperbolic stationary point.
- As Y increases further, $x_-(Y)$ reaches the value of the abscissa of the $f(x)$ cubic local maximum, where $x_-(Y) = 2/\sqrt{3} \equiv x_0^+$, $X_-(Y) \approx -2.22$, and $Y \approx 2.1$. The 2D system undergoes another Hopf bifurcation. In the 3D system, the attractive periodic orbit and the hyperbolic stationary point coalesce into an attractive stationary point.
- Finally, when $Y = -16/(3\sqrt{3})$, equation (4.1c) undergoes a saddle-node bifurcation again. The $X_0(Y)$ saddle and the $X_+(Y)$ attractive node disappear beyond the coincident point $X_0(Y) = X_+(Y) = 2/\sqrt{3}$, associated with $x_0(Y) = x_-(Y) = -2.39$.

We now go back to the 4D system (2.1a)–(2.1d) to unravel the hysteresis loop underlying the sequence of phases listed in subsection 2.2.3.

- In phase 1, Y remains almost constant and system (4.1a)–(4.1c) displays the single attractive point associated to $X_+(Y)$, corresponding to the ascending part of the surge.
- As $X_+(Y)$ decreases slowly, the solution of the 4D system remains close to the attractive point, corresponding to the duration of the surge (phase 2), until this node disappears through a saddle-node bifurcation. Then the solution switches to the other attractive node $X_-(Y)$, corresponding to the decreasing part of the surge (phase 3).
- In phase 4, as $X_-(Y)$ increases slowly, the solution remains close to the attractive point associated to $X_-(Y)$, corresponding to the plateau. Eventually, this attractive point disappears into an attractive periodic orbit via a Hopf bifurcation, initiating the pulsatile phase.
- As phase 1 starts again, X speeds up and the pulse frequency increases. At some point the attractive periodic orbit disappears into a saddle-node bifurcation of periodic orbits. The solution of the 4D system then jumps back to the single attractive node associated to $X_+(Y)$ and recovers the ascending part of the surge.

5. Identification of parameter targets for steroid control.

In this section we derive some tools that will be useful in controlling the amplitude and frequency of oscillations in either the slow regulatory system or the GnRH system and hence in mimicking either indirect or direct effects of steroid feedback.

5.1. Amplitude and frequency of the slow regulatory oscillations.

5.1.1. Amplitude of the slow limit cycle. Let us first derive a parametric expression of the $F(X)$ cubic as a function of its extrema. Denote by $-\alpha$ and β the roots of $F'(X) = 0$. Without loss of generality, we assume $\alpha > 0$ and $\beta > 0$ and choose the cubic function $F(X)$ so that it passes through the origin and takes minus unity as coefficient for the higher (cubic) power:

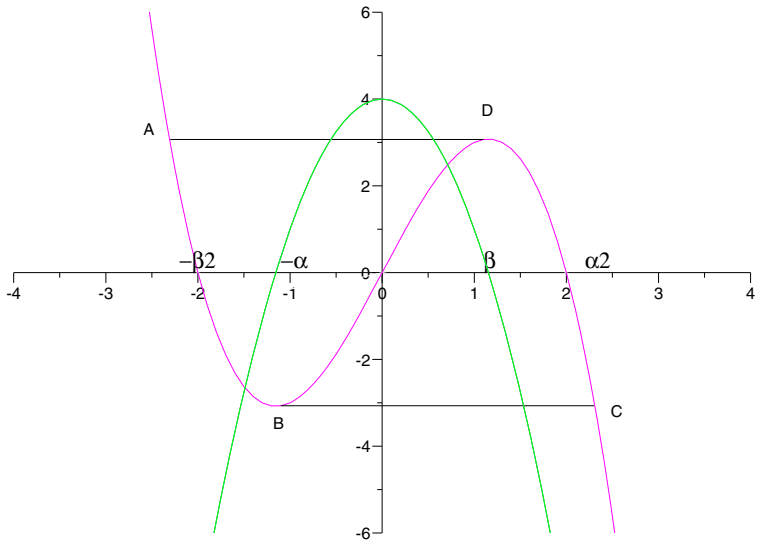


Figure 7. *Amplitude and period of the slow limit cycle (X, Y). The period of the slow (X, Y) limit cycle is computed from the time taken to go along the path* $\mathbf{A}(-\beta_2, F(-\beta_2))$—$\mathbf{B}(-\alpha, F(-\alpha))$—$\mathbf{C}(\alpha_2, F(\alpha_2))$—$\mathbf{D}(\beta, F(\beta))$, *with* $F'(\alpha) = F'(\beta) = 0$, $F(\alpha_2) = F(\alpha)$, $F(\beta_2) = F(\beta)$. *The global amplitude of* X *corresponds to* $\alpha_2 + \beta_2$, *and the surge-related amplitude corresponds to* $\alpha_2 + \alpha = \beta + \beta_2$.

$$
\begin{aligned}
F(X) &= -X^3 + \mu_1 X^2 + \mu_2 X \quad \text{with } \mu_1 > 0,\ \mu_2 > 0, \text{ and } \mu_0 = -1,\\
F'(X) &= -3X^2 + 2\mu_1 X + \mu_2\\
&= -3(X+\alpha)(X-\beta)\\
&= -3\left[X^2 + X(\alpha-\beta) - \alpha\beta\right],\\
F(X) &= -X^3 - \frac{3}{2}(\alpha-\beta)X^2 + 3\alpha\beta X.
\end{aligned}
$$

Hence α and β are the positive roots of

$$
\begin{cases}
3\alpha\beta = \mu_2,\\
3\beta^2 - 2\mu_1\beta - \mu_2 = 0
\end{cases}
$$

so that

$$
\alpha = \frac{\mu_2}{\mu_1 + \sqrt{\mu_1^2 + 3\mu_2}} \qquad \text{and} \qquad \beta = \frac{\mu_1 + \sqrt{\mu_1^2 + 3\mu_2}}{3}.
$$

To compute the amplitude of the X variable, we now seek the points of the cubic verifying either $F(X) = F(-\alpha)$ or $F(X) = F(\beta)$, and $F'(X) \neq 0$ (see Figure 7):

$$
F(X) = F(-\alpha) = -\frac{1}{2}\alpha^2(\alpha + 3\beta) \Leftrightarrow (X-\alpha)^2\left[X - \frac{1}{2}(\alpha + 3\beta)\right] = 0,
$$

$$F(X) = F(\beta) = \frac{1}{2}\beta^2(\beta + 3\alpha) \Leftrightarrow (X - \beta)^2 \left[X + \frac{1}{2}(\beta + 3\alpha)\right] = 0.$$

Define

$$\alpha_2 \equiv \frac{1}{2}(\alpha + 3\beta) \quad \text{and} \quad \beta_2 \equiv \frac{1}{2}(\beta + 3\alpha).$$

Then $F(\alpha_2) = F(-\alpha)$, $F(-\beta_2) = F(\beta)$, and the global amplitude of X is thus given by $\alpha_2 + \beta_2 = 2(\alpha + \beta)$, while the fast jump amplitude (related to the surge amplitude in the GnRH system) is given by $\alpha_2 + \alpha = \beta + \beta_2 = \frac{3}{2}(\alpha + \beta)$. In the symmetric case where $\alpha = \beta$, those amplitudes respectively reduce to 4α and 3α.

5.1.2. Period of the slow limit cycle. The frequency of the regulatory oscillations drives the frequency of the GnRH surge. The period of the slow cycle (i.e., the time taken to move along the whole cycle) in X indeed corresponds to the period of the GnRH surge.

Let us denote by T the period of the (X, Y) limit cycle. The time T_X taken to move along the slow part (i.e., along the branches of the cubic function F) of the cycle can be computed as

$$\begin{aligned} T_X &= \int_{F(-\beta_2)}^{F(-\alpha)} \frac{dY}{b_0X + b_1Y + b_2} + \int_{F(\alpha_2)}^{F(\beta)} \frac{dY}{b_0X + b_1Y + b_2} \\ &= \int_{-\beta_2}^{-\alpha} \frac{F'(X)dX}{b_0X + b_1F(X) + b_2} + \int_{\alpha_2}^{\beta} \frac{F'(X)dX}{b_0X + b_1F(X) + b_2}. \end{aligned}$$

It is worth noticing that, within the T_X duration, the time taken to climb up the descending (right) branch of the cubic (integration from $F(\alpha_2)$ to $F(\beta)$) corresponds to the duration of the surge.

The time T_Y taken to move along the quasi-horizontal fast parts (corresponding to the jumps from one branch of the cubic function to another) of the (X, Y) limit cycle can be computed as

$$T_Y = \int_{-\alpha}^{\alpha_2} \frac{\gamma dX}{F(X) - Y} + \int_{\beta}^{-\beta_2} \frac{\gamma dX}{F(X) - Y}.$$

Alternatively, T_Y can be estimated as $\approx 2\epsilon\gamma T_X$, according to the ratio of time constants in (2.1c) and (2.1d).

The whole (X, Y) limit cycle period can be computed finally from $T = T_X + T_Y$.

5.1.3. Changes in amplitude and frequency according to a prescribed ratio. From the previous analysis, we are able to assess numerically the frequency and amplitude of the slow (X, Y) limit cycle for any arbitrary parameter set. We now make a further remark.

Given the system

$$\begin{aligned} \gamma\dot{X} &= F(X) - Y, \\ \dot{y} &= b_0X + b_1Y + b_2, \end{aligned}$$

one can find $\{\bar{\gamma}, \bar{f}, \bar{b_0}, \bar{b_1}, \bar{b_2}\}$ to transform it into another system,

$$\bar{\gamma}\dot{X} = \bar{F}(X) - Y, \tag{5.1a}$$

$$\dot{Y} = \bar{b_0}X + \bar{b_1}Y + \bar{b_2}, \tag{5.1b}$$

oscillating with prescribed λ_1 frequency and λ_2 amplitude ratios.

This change of variables leads to

$$\begin{aligned}
\bar{\gamma} &= \lambda_1 \gamma, \\
\bar{b_0} &= \lambda_1 b_0, \\
\bar{b_1} &= \lambda_1 b_1, \\
\bar{b_2} &= b_2 \frac{\lambda_1}{\lambda_2}, \\
\bar{\mu_0} &= \lambda_2^2 \mu_0, \\
\bar{\mu_1} &= \lambda_2 \mu_1.
\end{aligned}$$

In other words, the new system reads

$$\begin{aligned}
\gamma \lambda_1 \dot{X} &= -Y + \lambda_2^2 \mu_0 x^3 + \lambda_2 \mu_1 x^2 + \mu_2 x, \\
\dot{Y} &= \lambda_1 \left(b_0 X + b_1 Y + \frac{b_2}{\lambda_2} \right).
\end{aligned}$$

The transformation does not modify the number of intersection points between the $\dot{x} = 0$ and $\dot{y} = 0$ nullclines, so that the qualitative behavior of the system is preserved.

5.2. Targeting parameters for steroid effect. The model allows us to distinguish between two possible pathways for steroid feedback on GnRH secretion, either directly, by acting on the parameters of the faster system, or indirectly, by acting on those of the slower one. The former way is dedicated to the control of the frequency and amplitude of GnRH pulses, while the latter is dedicated to the control of the onset time and size of the GnRH surge.

5.2.1. Steroid-like direct effects. Targeting adequate model parameters in the fast system in an acute way allows us to transiently alter the amplitude and frequency of GnRH pulses.

Figure 8 illustrates the effect of a parameter-targeting bolus ($\lambda_1 = 2$, $\lambda_2 = 1.5$) from time 525 to 575 leading to a frequency-increased, amplitude-decreased pulsatile regime. Such an output can be compared with experimental results gathered in [6].[1] These effects consist in a reduction of GnRH pulse amplitude and a stimulatory action on GnRH pulse frequency, as accounted for by the model results.

Figure 9 illustrates the effect of a parameter-targeting bolus ($\lambda_1 = 0.5$, $\lambda_2 = 0.5$) from time 525 to 575 leading to a frequency-decreased, amplitude-increased pulsatile regime. Such an output can be compared by experimental results gathered in [12].[2] These effects consist in a stimulatory action on GnRH pulse amplitude and a decrease in GnRH pulse frequency, as accounted for again by the model results.

Care should be taken in the mechanistic interpretation of those simulations. Direct effect in the models should be interpreted as acute effects rather than direct steroid effects on GnRH

[1] Where estradiol effects on GnRH secretory characteristics are summarized in Figure 3 of [6] and GnRH portal time series are displayed in Figure 2 of [6].

[2] Where progesterone effects on GnRH secretory characteristics are summarized in Figure 4 of [12].

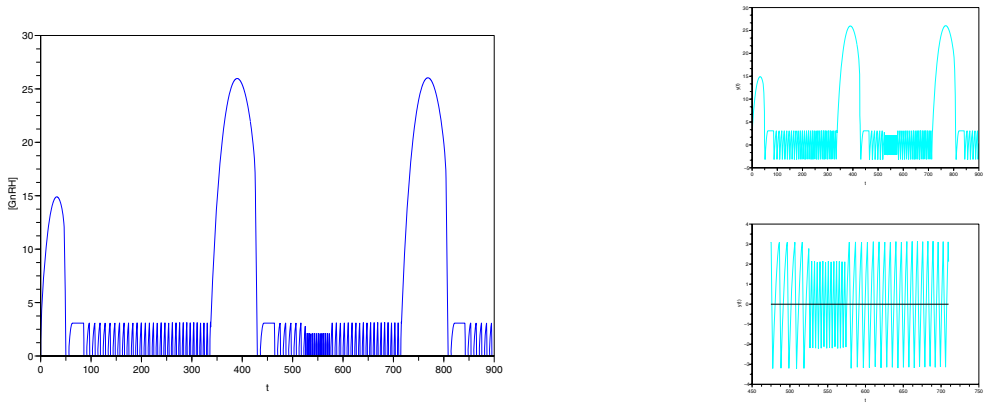


Figure 8. *Direct estradiol-like effect on portal GnRH output (left) and y output (top right and enlarged on bottom right). The parameter values were acutely altered in a bolus way from time* 525 *to* 575.

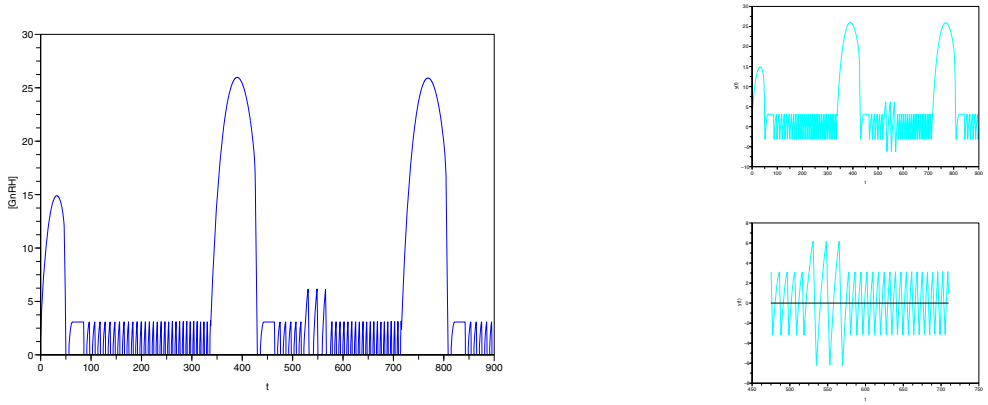


Figure 9. *Direct progesterone-like effect on portal GnRH output (left) and y output (top right and enlarged on bottom right). The parameter values were acutely altered in a bolus way from time* 525 *to* 575.

neurons, whether they are nuclear-initiated, through steroid receptors,[3] or membrane-initiated (a recently discovered signaling pathway [14]). But at the least, the simulations correspond to physiological short-term effect implying at most a few neuronal delays on short time scales, in contrast to the indirect, long-term effects described below.

5.2.2. Steroid-like indirect effects. Targeting adequate parameters in the slow system in a chronic way allows us to reproduce the known effects of progesterone on the surge amplitude and delay for surge onset. Figure 10 illustrates the effect of decreasing the amplitude and increasing the frequency of the oscillations in the X regulatory variable. This leads to a decrease in the delay between two consecutive GnRH surges as well as in the surge amplitude. Such combined effects mimic those that have been observed in an experimental study of the long term effect of progesterone priming [4], which compared the GnRH surge after exposure

[3]GnRH neurons are endowed with type β estradiol receptors [9] but do not own progesterone receptors [15].

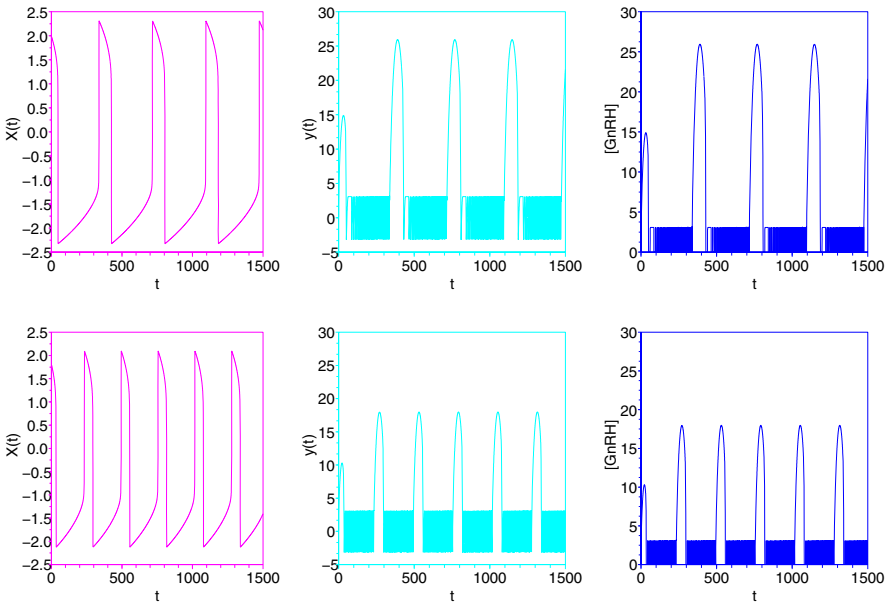


Figure 10. *Effect of progesterone-like priming on surge onset and amplitude. Top panel: Reference situation. The bottom panel situation corresponds to the absence of progesterone priming, with shortened surge onset and decreased surge amplitude. Pink line: $X(t)$. Cyan line: $y(t)$. Blue line: GnRH output.*

or not to progesterone. In the absence of progesterone priming, the size of the GnRH surge was decreased and its onset shortened.

6. Conclusion. We addressed in this paper the question of how the GnRH generator switches from a pulsatile to a surge secretion mode. Such a question arises on the physiological scale when considering the behavior of average GnRH neurons whose synchronization is taken for granted. Zeeman, Weckesser, and Gokhman [18] tackled the question of the GnRH-induced LH surge on the pituitary gland level, considering the GnRH self-priming on gonadotroph cells as a resonance phenomenon. The question of synchronization has been addressed in a recent paper [11].

On the hypothalamic level, only the variability in the frequency of GnRH pulses (rather than its control) has been, up to now, the focus of mathematical models based on nonlinear dynamics [2, 3]. Our modeling approach is comparable to these previous ones in the sense that it also considers the effect of the average activity of one group of neurons on the activity of another group. But the way by which this effect is introduced differs. They used as external inputs an impulsion train, whereas we assume that both groups can be represented by the same type of equations (of FitzHugh–Nagumo type) but with different time scales. Following a 3 time-scaled approach, we have managed to account not only for the alternating pulse and surge pattern of GnRH secretion but also for the frequency increase in the pulsatile regime. We have also unraveled the possible existence of a pause before pulsatility resumption after the surge, which could be investigated from an experimental viewpoint. Hence the capacity of our model to display complex features interpretable against experimental evidence

suggests that such a modeling approach may be a useful complement to experimental studies of neuroendocrine systems.

REFERENCES

[1] A. Borisyuk and J. Rinzel, *Understanding neuronal dynamics by geometrical dissection of minimal models*, in Methods and Models in Neurophysics, Les Houches Summer School 80, C. Chow, B. Gutkin, D. Hansel, and C. Meunier, eds., Elsevier, New York, 2004, pp. 19–72.

[2] D. Brown, J. P. A. Foweraker, and R. W. Marrs, *Dynamic equilibria and oscillations of a periodically stimulated excitable system*, Chaos Solitons Fractals, 5 (1995), pp. 359–369.

[3] D. Brown, A. E. Herbison, J. E. Robinson, R. W. Marrs, and G. Leng, *Modelling the luteinizing hormone-releasing hormone pulse generator*, Neuroscience, 63 (1994), pp. 869–879.

[4] A. Caraty and D. C. Skinner, *Progesterone priming is essential for the full expression of the positive feedback effect of estradiol in inducing the preovulatory gonadotropin-releasing hormone surge in the ewe*, Endocrinology, 140 (1999), pp. 165–170.

[5] H. M. Dungan, D. K. Clifton, and R. A. Steiner, *Minireview: Kisspeptin neurons as central processors in the regulation of gonadotropin-releasing hormone secretion*, Endocrinology, 147 (2006), pp. 1154–1158.

[6] N. P. Evans, G. E. Dahl, B. H. Glover, and F. J. Karsch, *Central regulation of pulsatile gonadotropin-releasing hormone (GnRH) secretion by estradiol during the period leading up to the preovulatory GnRH surge in the ewe*, Endocrinology, 134 (1994), pp. 1806–1811.

[7] N. P. Evans, T. A. Richter, D. C. Skinner, and J. E. Robinson, *Neuroendocrine mechanisms underlying the effects of progesterone on the oestradiol-induced GnRH/LH surge*, Reprod. Suppl., 59 (2002), pp. 57–66.

[8] A. E. Herbison, *Multimodal influence of estrogen upon gonadotropin-releasing hormone neurons*, Endocr. Rev., 19 (1998), pp. 302–330.

[9] E. Hrabovszky, A. Steinhauser, K. Barabas, P. J. Shughrue, S. L. Petersen, I. Merchenthaler, and Z. Liposits, *Estrogen receptor-beta immunoreactivity in luteinizing hormone-releasing hormone neurons of the rat brain*, Endocrinology, 142 (2001), pp. 3261–3264.

[10] F. J. Karsch, J. M. Bowen, A. Caraty, N. P. Evans, and S. M. Moenter, *Gonadotropin-releasing hormone requirements for ovulation*, Biol. Reprod., 56 (1997), pp. 303–309.

[11] A. Khadra and Y. X. Li, *A model for the pulsatile secretion of gonadotropin-releasing hormone from synchronized hypothalamic neurons*, Biophys. J., 91 (2006), pp. 74–83.

[12] S. M. Moenter, A. Caraty, A. Locatelli, and F. J. Karsch, *Pattern of gonadotropin-releasing hormone (GnRH) secretion leading up to ovulation in the ewe: Existence of a preovulatory GnRH surge*, Neuroendocrinology, 129 (1991), pp. 1175–1182.

[13] A. S. Mc Neilly, J. L. Crawford, C. Taragnat, L. Nicol, and J. R. Mc Neilly, *The differential secretion of LH and FSH/ regulation through genes, feedback and packaging*, Reprod. Suppl., 61 (2003), pp. 463–476.

[14] O. K. Ronnekleiv and M. J. Kelly, *Diversity of ovarian steroid signaling in the hypothalamus*, Front. Neuroendocrinol., 26 (2005), pp. 65–84.

[15] D. C. Skinner, A. Caraty, and R. Allingham, *Unmasking the progesterone receptor in the preoptic area and hypothalamus of the ewe: No colocalization with gonadotropin-releasing neurons*, Endocrinology, 142 (2001), pp. 573–579.

[16] M. J. Smith and L. Jennes, *Neural signals that regulate GnRH neurones directly during the oestrous cycle*, Reproduction, 122 (2001), pp. 1–10.

[17] Ei. Terasawa, W. K. Schanhofer, K. L. Keen, and L. Luchansky, *Intracellular Ca2+ oscillations in luteinizing hormone-releasing hormone neurons derived from the embryonic olfactory placode of the rhesus monkey*, J. Neurosci., 19 (1999), pp. 5898–5909.

[18] M. L. Zeeman, W. Weckesser, and D. Gokhman, *Resonance in the menstrual cycle: A new model of the LH surge*, Reprod. Biomed. Online, 7 (2003), pp. 295–300.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 2, pp. 457–474

Spurious Lyapunov Exponents Computed from Data*

Joshua A. Tempkin† and James A. Yorke‡

Abstract. Lyapunov exponents can be difficult to determine from experimental data. In particular, when using embedding theory to build chaotic attractors in a reconstruction space, extra "spurious" Lyapunov exponents can arise that are not Lyapunov exponents of the original system. By studying the local linearization matrices that are key to a popular method for computing Lyapunov exponents, we determine explicit formulas for the spurious exponents in certain cases. Notably, when a two-dimensional system with Lyapunov exponents λ and μ is reconstructed in a five-dimensional space by a generic embedding, the reconstructed system has exponents λ, μ, 2λ, $\lambda+\mu$, and 2μ. In particular, when $\mu < 0 < \lambda$, we have exponents $2\mu < \mu < \lambda + \mu < \lambda < 2\lambda$. The importance of this is that, for a generic embedding, the true exponents λ and μ are neither the largest nor smallest values nor even in the middle of this list. We view this result as a caution to those who wish to separate the spurious exponents from the true, especially since we rarely know a priori how many exponents are spurious.

Key words. Lyapunov exponents, embedding, time-series analysis, attractor reconstruction

AMS subject classifications. 34D08

DOI. 10.1137/040619211

1. Introduction.

In the analysis of observed physical systems, it has become standard practice to study an auxiliary system reconstructed from a time series of measured data. Successful reconstruction of the original system's attractor is the basis of the method for Lyapunov exponent calculation proposed by Eckmann and Ruelle [1] (see also [2]) and by Sano and Sawada [3]. However, since the reconstructed attractor often lies in a larger-dimensional space than the original system, the calculations produce too many exponents. This leads to an important question: how do we distinguish the true Lyapunov exponents of the underlying system from the extra "spurious" ones present for the reconstructed system? For the present paper, we shall assume that we have ideal conditions, namely, that we know the dimension of the underlying dynamics, that the Lyapunov exponents of the underlying dynamics exist, that there is no noise, that we have infinitely long data trajectories, and that numerical methods converge to the right objects. In other words, we attempt to determine the values that the Eckmann–Ruelle procedure should ideally produce if all practical issues could be satisfactorily addressed. In this paper, we answer the above question for one specific case: when a two-dimensional dynamical system is reconstructed in five-dimensional space. A two-dimensional system (such as the Hénon map) has two Lyapunov exponents, denoted λ and μ. The Eckmann–Ruelle procedure yields five exponents since the reconstruction space is $\mathbb{R}^5$.

*Received by the editors November 18, 2004; accepted for publication (in revised form) by T. Kaper November 22, 2006; published electronically June 29, 2007. This research was supported under NSF grant DMS 0104087.
http://www.siam.org/journals/siads/6-2/61921.html
†Department of Defense, 9800 Savage Road, Suite 6419, Fort Meade, MD 20755-6419 (tempkin@glue.umd.edu).
‡Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742 (yorke2@ipst.umd.edu).

We show that for a generic embedding (in the ideal case) these five exponents are λ, μ, 2λ, $\lambda + \mu$, and 2μ.

At this point, we make several observations. First, given five numbers, if one were told that they must be λ, μ, 2λ, $\lambda + \mu$, and 2μ for some λ and μ, it would be straightforward to determine the corresponding values of λ and μ. More importantly, if $\mu < 0 < \lambda$, then the computed exponents will be ordered

$$2\mu < \mu < \lambda + \mu < \lambda < 2\lambda,$$

and both the largest and smallest exponents will be spurious. This observation should serve as a caution to the users of these methods: even under ideal circumstances, size criteria will not let you select the true exponents, especially when you do not know how many exponents are spurious. Others have also shown that the largest exponent can be spurious, for example Dechert and Gençay [4], who gave an explicit example of an embedding from $\mathbb{R}^2$ into $\mathbb{R}^3$ where the spurious exponent could be arranged to take any value and in particular could be larger than the largest true exponent. It follows from our results that generically the largest exponent of the reconstructed dynamics will be spurious (when one of the exponents is positive), and that this property is independent of the reconstruction.

Our arguments extend naturally to other cases as well [5]. For example, when a one-dimensional system (such as the logistic map) with Lyapunov exponent λ is reconstructed in m-dimensional space, the Eckmann–Ruelle procedure will produce the m exponents $\lambda, 2\lambda, \ldots, m\lambda$. Observations in [2, pp. 4976–4978] allude to such a formula. In addition, since time-delay embeddings are a special case of measurement function (consisting of time-delayed versions of a scalar measurement [6, 7]), our results hold true for these popular embeddings as well. These results were announced in [8] and are presented here in more detail. Numerical studies are reported in [9], where noise is added.

The *"Eckmann–Ruelle" algorithm* presented in [2] for computing Lyapunov exponents has three major steps. First, one reconstructs the attractor in some m-dimensional Euclidean space of measurements. (The attractor reconstruction process will be described in more detail in section 2.) Along the reconstructed attractor, there is a time-τ map F, which takes the m-vector P_t at time t to the m-vector $P_{t+\tau}$ at time $t + \tau$. This map F represents the reconstructed dynamics, and $F(P_t)$ is defined to be $P_{t+\tau}$. In the second step, one computes a local linearization matrix for F at each point P of the reconstructed attractor by finding the $m \times m$ matrix M (depending on P) that satisfies as closely as possible

$$P_{t+\tau} - F(P) \approx M(P_t - P) \qquad \text{for all } t \text{ for which } P_t \text{ is close to } P.$$

We call this best local linearization $M = M(P)$ at P the "Eckmann–Ruelle linearization" at P. In the last step (which will be discussed in more detail in section 3), one computes the Lyapunov exponents of F from a matrix product of these local linearizations. Write $M_i = M(P_{t+i\tau})$ for the best local linearization matrix (i.e., the Eckmann–Ruelle linearization) at the point $P_{t+i\tau}$ in the trajectory. Then, the Lyapunov exponents (including the spurious ones) are the values achieved by

$$\lim_{n\to\infty} \frac{1}{n} \ln \|M_{n-1} M_{n-2} \cdots M_1 M_0 v\|$$

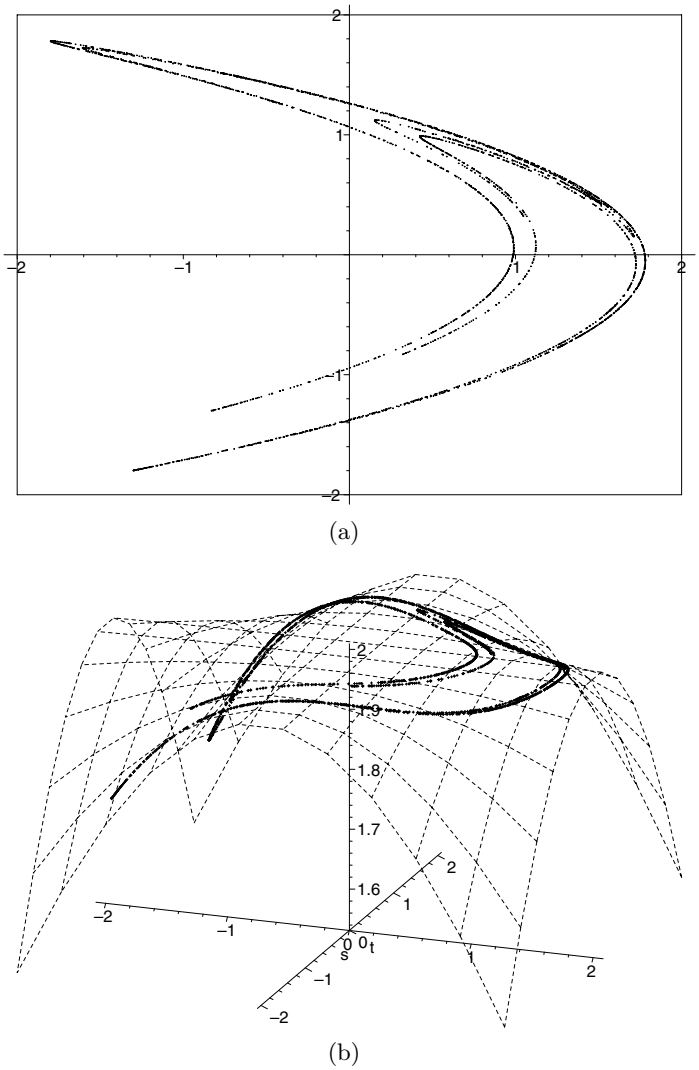


Figure 1. *A simple function π from $\mathbb{R}^2$ to $\mathbb{R}^3$ reconstructs the true attractor of the underlying dynamics (a Hénon attractor shown in* (a)*) as a set in $\mathbb{R}^3$ (shown in* (b)*). The surface indicated by dashed lines is the image of the plane. The reconstructed attractor lies within this lower-dimensional surface.*

for (unit) vectors $v \in \mathbb{R}^m$. Almost all vectors v will yield the largest exponent, but a hyperplane in $\mathbb{R}^m$ will contain vectors yielding the other exponents.

This procedure will produce one exponent for each dimension of the reconstruction space. To understand the numbers produced by the algorithm, it is crucial to determine the Eckmann–Ruelle linearizations. At first glance, one might think these linearizations should be derivative matrices, DF_P, but this is not the case (unless the underlying phase space and the reconstruction space have the same dimension). Suppose, for example, that an attractor reconstructed in m-dimensional Euclidean space lies within a lower-dimensional surface in $\mathbb{R}^m$; see Figure 1. The reconstructed dynamics are well defined on this surface, but they are not defined off of it, and so the classical $m \times m$ derivative matrix DF_P will not exist. Therefore,

the local linearizations cannot be derivative matrices.

In this paper, we study the local linearization matrices. For our main case, a reconstruction from $\mathbb{R}^2$ into $\mathbb{R}^5$, we will show that these matrices have several important properties. First and foremost, the linearizations of the reconstructed dynamics are surprisingly good in the following sense (see Theorem 2.2 in section 2). For any matrix L, we can determine how good a linearization L is by finding the largest integer k such that

$$F(P_t) - F(P) - L\,(P_t - P) = O\left(\|P_t - P\|^k\right) \quad \text{as } P_t \to P. \tag{1}$$

(We write $g(x) = O(\|x\|^k)$ as $x \to 0$ to mean that there exists a constant C such that $\|g(x)\| \leq C\|x\|^k$ for all x in some neighborhood of 0.) For most matrices L in (1), we expect $k = 1$. For the traditional derivative $L = DF_P$, we expect $k = 2$ (when it exists). However, the Eckmann–Ruelle linearization $L = M(P)$ often has k larger than this. In the case of a two-dimensional system reconstructed in five dimensions, we show $k = 3$. In the case of a one-dimensional system reconstructed in m dimensions, we find $k = m + 1$ [5, 8]. The second important property of these linearizations is that there are natural coordinate systems with respect to which the linearization matrices have a special upper-triangular matrix representation. This upper-triangular form allows us to easily compute the Lyapunov exponents. Moreover, for generic embeddings, these exponents will be completely independent of the specific embedding used in the reconstruction, depending only on the dynamics of the original system. This key property allows us to derive explicit formulas for the Lyapunov exponents produced by the Eckmann–Ruelle procedure in the low noise limit. We show how to obtain these formulas in section 3.

In practice, one implements the Eckmann–Ruelle algorithm by computing local linearization matrices using reconstructed trajectories of the dynamical system. Mera and Morán [10] find conditions ensuring the convergence of this algorithm. We assume that the local linearization matrices computed using reconstructed trajectories converge to the Eckmann–Ruelle linearizations. Convergence is assumed to facilitate the study of the output of the Eckmann–Ruelle algorithm. See also Sauer and Yorke [9] for the case where noise is present.

In [4], Dechert and Gençay showed that the true Lyapunov exponents should be a subset of the computed exponents. Several other methods for determining which exponents are spurious have already been put forward. Parlitz [11] proposed that recomputing the exponents using the reversed time series would make the true exponents switch sign. However, Kantz and Schreiber report [12, p. 183] that Parlitz's results have been difficult to verify. In our main case, all five exponents would change sign when time is reversed, including the spurious exponents. Bryant, Brown, and Abarbanel suggested that the local "thickness" of the data set could help identify spurious exponents [13, p. 2792]. Other authors have proposed removing the extra dimensions present in the reconstruction by projecting the reconstructed dynamics to the attractor's tangent space (see [14, pp. 336–339] and [15, pp. 2156–2157]). Mera and Morán [10, 16] discuss the convergence of this projection method. Ott and Yorke [17, section 7] prove that the spurious exponents may be identified simply by examining the tangent space at a single point of the reconstructed trajectory. Each true exponent corresponds to a Lyapunov vector in the tangent space. Furthermore, "tangent space" is defined there even for cases where the attractor is not a manifold. Whether this approach is practical or numerically

feasible is not known, but we intend to implement these ideas in a future project.

Our paper is structured in the following manner. In section 2, we determine the linear map that provides the best local linearization to the reconstructed dynamics. In section 3, we give the matrix representation alluded to above and derive formulas for the Lyapunov exponents produced by the algorithm. Section 4 presents numerical studies illustrating the theoretical results.

2. Local linearizations.

2.1. Attractor reconstruction. The attractor reconstruction process begins by choosing a number m and observing the present state p of the underlying system with that number m of independent measurements $\pi_i(p)$, $i = 1, \ldots, m$. For each point p in the phase space, there is an m-dimensional vector $\pi(p) = (\pi_1(p), \ldots, \pi_m(p))$. This produces a *measurement function* π that associates points in $\mathbb{R}^m$ with points in the phase space of the underlying dynamical system. Under certain generic conditions [18], the original dynamical phase space attractor A will be topologically equivalent to its "reconstructed" image $\pi(A)$ in m-dimensional Euclidean space. These generic conditions typically force the reconstruction space to have larger dimension than the underlying attractor. In the situations that interest us, the generic measurement function can be taken to be a diffeomorphism from the underlying phase space into m-dimensional Euclidean space $\mathbb{R}^m$. An example of a nongeneric map from $\mathbb{R}^2$ to $\mathbb{R}^5$ is where $\mathbb{R}^2$ maps into a linear subspace of dimension less than 5. Our results do not hold for these atypical cases. In practice, of course, an experimenter does not know the measurement function that arises from the attractor reconstruction process. This is why we focus on generic measurement functions π.

The set of measurement vectors $\pi(p)$ in the reconstructed attractor $\pi(A)$ can be studied for geometrical and dynamical properties. Looking to the reconstructed attractor for dynamical properties of the original attractor was suggested in 1985 by Eckmann and Ruelle and coworkers [1, 2] and also by Sano and Sawada [3]. It is usually necessary that the number m of measurements be chosen large enough that there is a one-to-one correspondence between points of the original attractor and points of the reconstructed attractor [18]. This requirement often forces the reconstruction space to have larger dimension than the original system. In these cases, Lyapunov exponent calculations in the reconstruction space produce m real numbers, not all of which can be Lyapunov exponents of the original dynamical system. For example, when a two-dimensional dynamical system is reconstructed in five dimensions, typical methods compute five "exponents" in the reconstruction space. At most two of these can be Lyapunov exponents of the original system; the other numbers are "spurious" exponents. Our goal is to distinguish the spurious exponents from the true exponents.

2.2. Delay coordinates. As electronics and computer power increase, experimenters find it easier to make several independent measurements of the system being observed. However, Takens [6] built a theory based on his observation that for generic systems it is possible to create multiple measurements from one, as follows. Assume that $\pi_1(p)$ is a (scalar) measurement and that $X_\tau(p)$ denotes the point where the trajectory now at p was at time τ ago. We can then define $\pi_2(p) = \pi_1(X_\tau(p))$ and $\pi_{r+1}(p) = \pi_1(X_{r\tau}(p))$ for integers $r > 0$. The vector $\pi(p) = (\pi_1(p), \ldots, \pi_m(p))$ is an m-dimensional measurement function in the above sense. If $p(t)$ denotes a trajectory and $h(t) = \pi_1(p(t))$ is the recorded measurement, then $\pi(p(t))$ may

be written $(h(t), h(t-\tau), \ldots, h(t-[m-1]\tau)$, which is referred to as a "delay coordinate embedding." See [6] for restrictions on the set of periodic orbits when using this approach. In the present paper, the reader interested in the single measurement case can make the translation using the π_{r+1} function above. It is, however, preferable to make more measurements when they can be made with equal accuracy. Furthermore, [18] emphasizes that to minimize the effects of noise, it is best either to choose τ small and m large or to replace $p(t)$ with the average of p over an interval $[t-\tau, t]$. Eckmann and Ruelle and coworkers emphasize this delay coordinate approach, though their ideas are equally applicable to the more general case of independent measurements.

2.3. Notation. We adopt the convention that lower-case letters refer to the underlying system (Figure 1(a)), while upper-case letters refer to the reconstructed system (Figure 1(b)). For example, $P = \pi(p)$ is the point in the reconstructed phase space $\mathbb{R}^m$ corresponding to the point p in the underlying phase space. We represent the dynamical flow on the underlying phase space by the time-τ map f for some τ. The measurement function π maps the underlying phase space into $\mathbb{R}^m$ for some m. In $\mathbb{R}^m$, there is an induced time-τ flow map $F = \pi f \pi^{-1}$ that maps each $\pi(p)$ to $\pi(f(p))$, and we refer to F as the *reconstructed dynamics*. Assuming that π is a one-to-one correspondence on the underlying phase space (i.e., that $p \neq q$ implies $\pi(p) \neq \pi(q)$), the map F is well defined on the reconstructed phase space. Finally, we say that a function is C^k if it has k continuous derivatives.

2.4. Local linearizations. A simple example where $f(p) = 2p \pmod{2\pi}$ on the interval $[0, 2\pi]$ allows us to describe the general problem. We reconstruct this interval in $\mathbb{R}^2$ via the measurement function $\pi(p) = (\cos(p), \sin(p))$. The observed dynamics $F = \pi f \pi^{-1}$ maps points on the unit circle. What are the (computed) Lyapunov exponents of this example? Most algorithms will compute two numbers, at most one of which can be correct.

This question is central to our discussion: for any point $P = \pi(p)$ on the unit circle $\pi([0, 2\pi])$, which 2×2 matrix provides the best local linearization of F around P (in the sense of (1))? Since F is not defined off the unit circle, the traditional derivative of F does not exist. On the other hand, the linearizations of F are 2×2 matrices, while the derivatives of the original system f are 1×1. It follows that no linearization of F can be a derivative matrix. What, then, is the best local linearization of F near P (if one even exists)? In this section, we show how to determine this best local linearization.

We examine the special case when a two-dimensional dynamical system f on the unit square $[0, 1] \times [0, 1]$ is reconstructed into five-dimensional space. For convenience, we shall denote this unit square in $\mathbb{R}^2$ by *Sqr*. We make the following specific assumptions about f, π, and a point $P = \pi(p)$ in $\mathbb{R}^5$:

(A1) *f maps Sqr into itself, and f is C^3.*

(A2) *π is a C^3 diffeomorphism of Sqr into $\mathbb{R}^5$. In particular, for any point $P \in \pi(Sqr)$, the set $\pi^{-1}(P)$ contains exactly one point.*

(A3) *The first- and second-order partial derivative vectors for π at $p = (x, y)$, namely,*

$$\pi_x(p) := \frac{\partial \pi}{\partial x}(p), \qquad \pi_y(p) := \frac{\partial \pi}{\partial y}(p),$$

$$\pi_{xx}(p) := \frac{\partial^2 \pi}{\partial x^2}(p), \qquad \pi_{xy}(p) := \frac{\partial^2 \pi}{\partial x \partial y}(p), \qquad \pi_{yy}(p) := \frac{\partial^2 \pi}{\partial^2 y}(p),$$

are linearly independent in $\mathbb{R}^5$.

Property (A3) is generic in the space of C^3 functions from Sqr into $\mathbb{R}^5$. Since π is a diffeomorphism, there is a constant $C_\pi > 1$ such that if $P_i = \pi(p_i)$ with $p_i \in Sqr$, $i = 1, 2$, then

$$\frac{1}{C_\pi}\|P_1 - P_2\| \le \|p_1 - p_2\| \le C_\pi \|P_1 - P_2\|. \tag{2}$$

This inequality will be useful in translating statements back and forth between the underlying space and the reconstruction space.

We begin our analysis by examining the Taylor expansion of π about a point $P = \pi(p)$ in $\mathbb{R}^5$ for which (A3) holds. Any point $P+\Delta P$ in $\pi(Sqr)$ near P can be written with $h = (h_1, h_2)$ in $\mathbb{R}^2$:

$$\begin{aligned} P + \Delta P &= \pi\left(p + h\right) \\ &= P + h_1\pi_x(p) + h_2\pi_y(p) + \frac{1}{2}h_1^2\pi_{xx}(p) \\ &\quad + h_1h_2\pi_{xy}(p) + \frac{1}{2}h_2^2\pi_{yy}(p) + O\left(\|h\|^3\right). \end{aligned}$$

By hypothesis (A3), the vectors $\pi_x(p)$, $\pi_y(p)$, $\pi_{xx}(p)$, $\pi_{xy}(p)$, and $\pi_{yy}(p)$ are linearly independent and form a basis for $\mathbb{R}^5$, which we call the *canonical embedding basis at* P. The little vector ΔP at P can be written conveniently in this basis:

$$\Delta P = \left(h_1, h_2, \frac{1}{2}h_1^2, h_1h_2, \frac{1}{2}h_2^2\right)_P + O\left(\|h\|^3\right). \tag{3}$$

Next, we look at the image $F(P) = \pi(f(p))$ in essentially the same way. The Taylor expansion of $F(P + \Delta P) = F(\pi(p + h)) = \pi(f(p + h))$ is given by

$$\begin{aligned} F(P + \Delta P) &= F(P) + h_1(\pi f)_x(p) + h_2(\pi f)_y(p) + \frac{1}{2}h_1^2(\pi f)_{xx}(p) \\ &\quad + h_1h_2(\pi f)_{xy}(p) + \frac{1}{2}h_2^2(\pi f)_{yy}(p) + O\left(\|h\|^3\right) \end{aligned} \tag{4}$$

for small h. At this point, we can define the Eckmann–Ruelle linearization. In their papers [1, 2], the Eckmann and Ruelle group discussed the best local linearization of the reconstructed dynamics. We formally define the "Eckmann–Ruelle linearization" to be the unique linear map with certain properties. In Theorem 2.2, we show that our definition provides the best local linearization in the case when two-dimensional dynamics is reconstructed in five-dimensional space. The work in [5] shows how to generate the Eckmann–Ruelle linearization in certain other cases.

Definition 2.1. *Assuming* (A1)–(A3), *we define the* Eckmann–Ruelle linearization at $P = \pi(p)$ *to be the unique linear map* $M = M(P)$ *on* $\mathbb{R}^5$ *satisfying*

$$\begin{aligned} M\pi_x(p) &= (\pi f)_x(p), & M\pi_{xx}(p) &= (\pi f)_{xx}(p), \\ M\pi_y(p) &= (\pi f)_y(p), & M\pi_{xy}(p) &= (\pi f)_{xy}(p), \\ & & M\pi_{yy}(p) &= (\pi f)_{yy}(p). \end{aligned} \tag{5}$$

M is well defined and unique because the vectors on the left-hand side of (5) form a basis for $\mathbb{R}^5$, by assumption (A3). While it may be convenient to think of M as a matrix, no coordinate system has yet been specified. The example below will represent M with respect to one particular coordinate system (the natural one for numerical computations), but in section 3 we will make use of a different matrix representation for M. Our proof of the local linearization theorem below does not require any matrix representation at all; it needs only the equations in (5).

For an example of an Eckmann–Ruelle linearization, we shall compute M explicitly with respect to standard coordinates in $\mathbb{R}^5$. Let the underlying dynamical system be the Hénon map $f(x, y) = (1.4 - x^2 + 0.3y, x)$. The simplest measurement function that satisfies hypotheses (A1)–(A3) is $\pi(x, y) = (x, y, x^2, y^2, xy)$. For this measurement function, we have

$$\begin{aligned} \pi_x(x, y) &= (1, 0, 2x, 0, y), \\ \pi_y(x, y) &= (0, 1, 0, 2y, x), \\ \pi_{xx}(x, y) &= (0, 0, 2, 0, 0), \\ \pi_{yy}(x, y) &= (0, 0, 0, 2, 0), \\ \pi_{xy}(x, y) &= (0, 0, 0, 0, 1), \end{aligned} \tag{6}$$

and with

$$(\pi f)(x, y) = (1.4 - x^2 + 0.3y, x, (1.4 - x^2 + 0.3y)^2, x^2, 1.4x - x^3 + 0.3xy)$$

we have

$$\begin{aligned} (\pi f)_x(x, y) &= (-2x, 1, 4x^3 - 5.6x - 1.2xy, 2x, 1.4 - 3x^2 + 0.3y), \\ (\pi f)_y(x, y) &= (0.3, 0, 0.84 - 0.6x^2 + 0.18y, 0, 0.3x), \\ (\pi f)_{xx}(x, y) &= (-2, 0, 12x^2 - 5.6 - 1.2y, 2, -6x), \\ (\pi f)_{yy}(x, y) &= (0, 0, 0.18, 0, 0), \\ (\pi f)_{xy}(x, y) &= (0, 0, -1.2x, 0, 0.3). \end{aligned} \tag{7}$$

M is the unique linear map that maps the vectors in (6) to corresponding vectors in (7). To find the matrix representation of M at $\pi(x, y)$ with respect to the standard basis in $\mathbb{R}^5$, we need to see where the standard unit vectors map. From the equations above, note that

$$(1, 0, 0, 0, 0) = \pi_x(x, y) - x\pi_{xx}(x, y) - y\pi_{xy}(x, y),$$

and thus

$$\begin{aligned} M(1, 0, 0, 0, 0) &= M\pi_x(x, y) - xM\pi_{xx}(x, y) - yM\pi_{xy}(x, y) \\ &= (\pi f)_x(x, y) - x(\pi f)_{xx}(x, y) - y(\pi f)_{xy}(x, y) \\ &= (0, 1, 1.2xy - 8x^3, 0, 1.4 + 3x^2). \end{aligned}$$

Similarly, we obtain mappings for the other standard basis vectors. Putting all of this together, we obtain the matrix representation for M at $P = \pi(x, y)$:

$$M(P) = \begin{pmatrix} 0 & 0.3 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1.2xy - 8x^3 & 0.84 + 0.6x^2 & -2.8 + 6x^2 - 0.6y & .09 & -1.2x \\ 0 & 0 & 1 & 0 & 0 \\ 1.4 + 3x^2 & 0 & -3x & 0 & 0.3 \end{pmatrix}.$$

Next, we prove that the Eckmann–Ruelle linearization does a very good job of linearizing the observed dynamics.

Theorem 2.2 (local linearizations for $\mathbb{R}^2 \to \mathbb{R}^5$). *Assume* (A1) *and* (A2). *Let* $P = \pi(p)$ *be a point of* $\pi(Sqr)$ *for which* (A3) *holds. The Eckmann–Ruelle linearization* $M = M(P)$ *defined by* (5) *is the unique linear map such that*

$$F(P + \Delta P) - F(P) - M\Delta P = O\left(\|\Delta P\|^3\right) \quad \text{as } \|\Delta P\| \to 0, \quad \text{where } P + \Delta P \in \pi(Sqr). \tag{8}$$

Proof. For small ΔP with $P + \Delta P \in \pi(Sqr)$, we can write $P + \Delta P = \pi(p + h)$, where $h = (h_1, h_2)$. Note that $F(P) = \pi(f(p))$ and $F(P + \Delta P) = \pi(f(p + h))$, and by (2), we have $O(\|h\|^3) = O(\|\Delta P\|^3)$.

Let M be any linear map. We expand $F(P + \Delta P) - F(P) - M\Delta P$ in terms of its Taylor series using (3) and (4):

$$\begin{aligned} F(P + \Delta P) - F(P) - M\Delta P = {} & h_1 \left((\pi f)_x(p) - M\pi_x(p)\right) \\ & + h_2 \left((\pi f)_y(p) - M\pi_y(p)\right) \\ & + \frac{1}{2} h_1^2 \left((\pi f)_{xx}(p) - M\pi_{xx}(p)\right) \\ & + h_1 h_2 \left((\pi f)_{xy}(p) - M\pi_{xy}(p)\right) \\ & + \frac{1}{2} h_2^2 \left((\pi f)_{yy}(p) - M\pi_{yy}(p)\right) \\ & + O\left(\|h\|^3\right). \end{aligned}$$

Since $\pi_x(p), \pi_y(p), \pi_{xx}(p), \pi_{xy}(p), \pi_{yy}(p)$ form a basis by (A3), there is only one linear map, namely, the Eckmann–Ruelle linearization, that can eliminate all five of the first- and second-order terms in this representation. All other linear maps must be either $O(\|h\|)$ or $O(\|h\|^2)$. Thus, it is clear that M satisfies (8) if and only if M is the Eckmann–Ruelle linearization defined by (5). ∎

Note that the error term in (8) of Theorem 2.2 is smaller than that for the usual Jacobian matrix, which would be $O(\|\Delta P\|^2)$ if it existed.

3. Lyapunov exponent formulas. Once the local linearization matrices have been computed at each point of the reconstructed trajectory, we must extract Lyapunov exponents from them. Mimicking the standard definition of Lyapunov exponents (see, for example,

[19, p. 31]), we define the *Eckmann–Ruelle–Lyapunov (ERL) exponents* of the reconstructed trajectory $P_0, P_1, \ldots$ in $\mathbb{R}^m$ to be the values obtained by the limit

$$h_{ER}(P_0, v) := \lim_{n\to\infty} \frac{1}{n} \ln \|M_{n-1} M_{n-2} \cdots M_1 M_0 v\| \tag{9}$$

for unit vectors $v \in \mathbb{R}^m$, where $M_i = M(P_i)$ is the best local linearization (i.e., Eckmann–Ruelle linearization) at the point $P_i = \pi(p_i)$ in the trajectory. This is exactly the same definition as for Lyapunov exponents, except that we use the M_i instead of Jacobian derivatives (which need not exist along the trajectory). In this section, we show that the matrix product in (9) can be written as an upper-triangular matrix. A straightforward calculation will then produce a formula for the limiting values of (9).

In practice, the limit in (9) can be computed using the treppen-iteration algorithm described in [1, 2] and elsewhere. This method uses QR matrix decomposition to convert the matrix product $M_{n-1} \cdots M_0$ into a product of upper-triangular matrices $R_{n-1} \cdots R_0$. The diagonal elements of the latter upper-triangular matrix are the products of the corresponding diagonal elements of the R_i. Then, we can read the ERL exponents right from the diagonal entries $(R_i)_{jj}$ of the intermediate matrices R_i:

$$\lambda_j = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| (R_i)_{jj} \right|.$$

Theorem 3.2 will justify the ability to read the exponents directly from the diagonal in this way. A proof of the theorem is given in [5].

Definition 3.1. *A sequence of real numbers* $\{r_n\}$ *has* (geometric) growth rate γ, *provided*

$$\lim_{n\to\infty} \frac{\ln |r_n|}{n} = \gamma.$$

Theorem 3.2. *For* $k = 1, 2, \ldots$, *let* A_k *be* $m \times m$ *upper-triangular matrices, and define* $S_n = A_n \cdots A_1$. *Assume that the magnitudes of the entries of* A_k *are bounded independent of* k, *and that the diagonal entries of* S_n *have growth rates* $\gamma_1, \ldots, \gamma_m$ *as* $n \to \infty$. *Then there exist vectors* $v_1, \ldots, v_m \in \mathbb{R}^m$ *such that for each* $i = 1, \ldots, m$, $\|S_n v_i\|$ *has growth rate* γ_i.

The next theorem is the main result of this paper, giving the formula for the exponents. Most of its hypotheses are used to guarantee that the Eckmann–Ruelle linearization exists at each point of the trajectory. Naturally, we must also assume the existence of Lyapunov exponents for the underlying trajectory. Theorems such as Oseledec's multiplicative ergodic theorem [1] guarantee the existence of Lyapunov exponents in appropriate circumstances. Since we will assume that Lyapunov exponents exist for the underlying trajectory, we shall not at this time add extra hypotheses to guarantee said existence. Specifically, given Lyapunov exponents $\lambda \geq \mu$, we assume the standard property that there is a unit tangent vector in $\mathbb{R}^2$ with growth rate μ and that almost every other tangent vector in $\mathbb{R}^2$ has growth rate λ.

Theorem 3.3 (Lyapunov exponent formula for $\mathbb{R}^2 \to \mathbb{R}^5$). *Assume* (A1), (A2), *and that the trajectory of* f *in Sqr,* $p_0, p_1, \ldots \in \mathbb{R}^2$, *has Lyapunov exponents* λ *and* μ. *Assume each point of the reconstructed trajectory,* $P_i = \pi(p_i)$ *in* $\mathbb{R}^5$, *satisfies* (A3). *Then, the reconstructed trajectory has ERL exponents* λ, μ, 2λ, $\lambda + \mu$, *and* 2μ.

Proof. Recall that the canonical embedding basis given in assumption (A3) from section 2 consists of the first- and second-order partial derivative vectors of π, namely, $\pi_x(p)$, $\pi_y(p)$, $\pi_{xx}(p)$, $\pi_{xy}(p)$, and $\pi_{yy}(p)$. In fact, the construction of the Eckmann–Ruelle linearization from the Taylor series of π and πf can be carried out with respect to any orthonormal set of coordinates. The uniqueness part of the local linearization Theorem 2.2 ensures that the resulting linear map will be the same. Thus, we may introduce convenient local coordinate systems at each point $p_i \in \mathbb{R}^2$ in the trajectory of p_0.

Without loss of generality, assume $\lambda \geq \mu$. At the initial point $p_0 \in \mathbb{R}^2$ of the underlying trajectory, choose a unit Lyapunov vector $v_0 \in \mathbb{R}^2$ corresponding to the exponent μ:

$$\lim_{n\to\infty} \frac{1}{n} \ln \left\| Df_{p_{n-1}} Df_{p_{n-2}} \cdots Df_{p_1} Df_{p_0} v_0 \right\| = \mu.$$

By hypothesis, almost every other vector in $\mathbb{R}^2$ has growth rate λ. Choose a unit vector w_0 perpendicular to v_0. This gives us an orthonormal basis $\{v_0, w_0\}$ for $\mathbb{R}^2$ based at p_0. Now, given a basis $\{v_n, w_n\}$ at $p_n \in \mathbb{R}^2$, we construct a basis at p_{n+1} by setting $v_{n+1} := Df_{p_n} v_n / \|Df_{p_n} v_n\|$ and choosing w_{n+1} to be the unit vector perpendicular to v_{n+1} that satisfies $\langle w_{n+1}, Df_{p_n} w_n \rangle > 0$. By induction, we have an orthonormal basis $\{v_n, w_n\}$ for $\mathbb{R}^2$ at each point p_n.

For each n, we write points $p' \in \mathbb{R}^2$ near p_n as $p' = (x, y)$, provided $p' = p_n + x v_n + y w_n$. We can describe the underlying dynamics f near p_n in terms of these bases at p_n and p_{n+1} by $f(x, y) = (g(x, y), h(x, y))$. Though we will not explicitly show it, this representation for f depends on the base point p_n, and the reader should keep in mind that the component functions g and h may look very different as we vary p_n. Note that for each n, $h_x(p_n) = 0$ because $Df_{p_n} v_n = \|Df_{p_n} v_n\| v_{n+1}$ has no y-component at p_{n+1}. Thus, we can write the Jacobian derivative of f at p_n as

$$Df_{p_n} = \begin{pmatrix} g_x(p_n) & g_y(p_n) \\ 0 & h_y(p_n) \end{pmatrix}.$$

A straightforward calculation using the chain rule produces the next set of equations, where the partial derivatives of g and h are evaluated at p_n:

$$\begin{aligned}
(\pi f)_x(p_n) &= g_x \pi_x(p_{n+1}), \\
(\pi f)_y(p_n) &= g_y \pi_x(p_{n+1}) + h_y \pi_y(p_{n+1}), \\
(\pi f)_{xx}(p_n) &= g_{xx} \pi_x(p_{n+1}) + h_{xx} \pi_y(p_{n+1}) + g_x^2 \pi_{xx}(p_{n+1}), \\
(\pi f)_{xy}(p_n) &= g_{xy} \pi_x(p_{n+1}) + h_{xy} \pi_y(p_{n+1}) + g_x g_y \pi_{xx}(p_{n+1}) \\
&\qquad + g_x h_y \pi_{xy}(p_{n+1}), \\
(\pi f)_{yy}(p_n) &= g_{yy} \pi_x(p_{n+1}) + h_{yy} \pi_y(p_{n+1}) + g_y^2 \pi_{xx}(p_{n+1}) \\
&\qquad + 2 g_y h_y \pi_{xy}(p_{n+1}) + h_y^2 \pi_{yy}(p_{n+1}).
\end{aligned}$$

For each n, let β_n denote the canonical embedding basis for $\mathbb{R}^5$ at $P_n = \pi(p_n)$. With respect to the canonical embedding basis at p_{n+1}, namely,

$$\beta_{n+1} = \{\pi_x(p_{n+1}), \pi_y(p_{n+1}), \pi_{xx}(p_{n+1}), \pi_{xy}(p_{n+1}), \pi_{yy}(p_{n+1})\},$$

the above equations become

$$\begin{aligned}
(\pi f)_x(p_n) &= (g_x, 0, 0, 0, 0)_{\beta_{n+1}},\\
(\pi f)_y(p_n) &= (g_y, h_y, 0, 0, 0)_{\beta_{n+1}},\\
(\pi f)_{xx}(p_n) &= (g_{xx}, h_{xx}, g_x^2, 0, 0)_{\beta_{n+1}},\\
(\pi f)_{xy}(p_n) &= (g_{xy}, h_{xy}, g_x g_y, g_x h_y, 0)_{\beta_{n+1}},\\
(\pi f)_{yy}(p_n) &= (g_{yy}, h_{yy}, g_y^2, 2g_y h_y, h_y^2)_{\beta_{n+1}}.
\end{aligned}$$

Moreover, the equations (5) that define the Eckmann–Ruelle linearization have a particularly nice representation using these bases. For example, the equation

$$M\pi_x(p_n) = (\pi f)_x(p_n) = g_x \pi_x(p_{n+1})$$

becomes

$$M(1, 0, 0, 0, 0)_{\beta_n} = (g_x, 0, 0, 0, 0)_{\beta_{n+1}}.$$

Continuing in this way, we obtain this matrix representation for the Eckmann–Ruelle linearization $M_n = M(P_n)$ with respect to β_n and β_{n+1}:

$$[M_n]_{\beta_n}^{\beta_{n+1}} = \begin{pmatrix} g_x & g_y & g_{xx} & g_{xy} & g_{yy} \\ 0 & h_y & h_{xx} & h_{xy} & h_{yy} \\ 0 & 0 & g_x^2 & g_x g_y & g_y^2 \\ 0 & 0 & 0 & g_x h_y & 2g_y h_y \\ 0 & 0 & 0 & 0 & h_y^2 \end{pmatrix}.$$

All of the terms above the diagonal come from combinations of first and second derivatives of f, and, because f is C^3 by hypothesis, they will be bounded independent of p_n. Note that the upper left 2×2 block is the Jacobian Df_{p_n} of the underlying dynamics, and the lower right 3×3 block contains only combinations of terms from Df_{p_n}.

It follows that the product $[M_{n-1} \cdots M_0]_{\beta_0}^{\beta_n}$ is upper-triangular when written with respect to the canonical embedding bases. Specifically, if we write

$$G_n = \prod_{i=0}^{n-1} g_x(p_i), \qquad H_n = \prod_{i=0}^{n-1} h_y(p_i),$$

then we have

$$[M_{n-1} \cdots M_0]_{\beta_0}^{\beta_n} = \begin{pmatrix} G_n & b_{12} & b_{13} & b_{14} & b_{15} \\ 0 & H_n & b_{23} & b_{24} & b_{25} \\ 0 & 0 & G_n^2 & b_{34} & b_{35} \\ 0 & 0 & 0 & G_n H_n & b_{45} \\ 0 & 0 & 0 & 0 & H_n^2 \end{pmatrix},$$

where the b_{ij} are numbers which depend solely on the underlying dynamical system f and the first n points of the trajectory of p_0 in *Sqr*.

Now we need to compute the growth rates of the diagonal terms. To do this, we compute the Lyapunov exponents λ and μ of the underlying dynamical system in terms of the components of the Jacobian matrices. Since $Df_{p_n}v_n = \|Df_{p_n}v_n\|v_{n+1}$, we have $g_x(p_n) = \|Df_{p_n}v_n\|$. It follows that

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\ln|G_n| &= \lim_{n\to\infty}\frac{1}{n}\ln\left(\prod_{i=0}^{n-1}\|Df_{p_i}v_i\|\right)\\ &= \lim_{n\to\infty}\frac{1}{n}\ln\|Df_{p_{n1}}\cdots Df_{p_0}v_0\|\\ &= \mu,\end{aligned} \tag{10}$$

because v_0 was chosen to be a Lyapunov vector for μ. Recall from [1, p. 632] that the growth rate of areas is given by the sum of the Lyapunov exponents. Thus,

$$\begin{aligned}\mu+\lambda &= \lim_{n\to\infty}\frac{1}{n}\ln|\det(Df^n{}_{p_0})|\\ &= \lim_{n\to\infty}\frac{1}{n}\ln\left(\prod_{i=0}^{n-1}|\det(Df_{p_i})|\right)\\ &= \lim_{n\to\infty}\frac{1}{n}\ln|G_nH_n|\\ &= \mu+\lim_{n\to\infty}\frac{1}{n}\ln|H_n|,\end{aligned}$$

and it follows that

$$\lambda = \lim_{n\to\infty}\frac{1}{n}\ln|H_n|. \tag{11}$$

Finally, we see that the diagonal terms of $[M_{n-1}\cdots M_0]_{\beta_0}^{\beta_n}$ have growth rates of λ, μ, 2λ, $\lambda+\mu$, and 2μ. It follows from Theorem 3.2 that the reconstructed trajectory $P_0, P_1, \ldots$ in $\mathbb{R}^5$ has these ERL exponents, completing the proof of Theorem 3.3. ■

4. Numerical computations. We have determined theoretically the computed Lyapunov exponents in our main case. However, we are aware that other numerical results have appeared in the literature that do not agree with our theorems, notably in [4]. Thus, we provide examples to demonstrate that numerical experiments can agree with our theorems. The reader may wish to contrast our results with those in [4].

Recall that the Eckmann–Ruelle procedure has three basic steps. In the first step, one reconstructs the attractor with a measurement function. In the second step, one determines a local linearization matrix at each point of the reconstructed trajectory. In the final step, one extracts the Lyapunov exponents from the linearization matrices obtained in the previous step.

For our numerical experiments, we assume that we are given the underlying dynamical system $f : \mathbb{R}^2 \to \mathbb{R}^2$ and a measurement function $\pi : \mathbb{R}^2 \to \mathbb{R}^5$ for reconstructing the attractor. We iterate the function f a number of times to remove transients. Then, continuing to

iterate f, we produce a trajectory $\{p_i\} \in \mathbb{R}^2$. Instead of storing the points p_i, we apply the measurement function to each p_i and store the vectors $P_i = \pi(p_i) \in \mathbb{R}^5$.

The procedure for computing the local linearization matrix at a base point P in $\mathbb{R}^5$ is straightforward and based on least-squares methods. The goal is to determine the 5×5 matrix M which best satisfies $M(P_i - P) \approx P_{i+1} - F(P)$ for all P_i close to P. Given a neighborhood radius ϵ, we search our data list for those P_i which are ϵ-close to P in $\mathbb{R}^5$. To avoid the false-near-neighbor problem, for each such P_i we also construct the vectors $P_{i+j} - F^j(P)$, $j = 1, \ldots, 4$, and check their lengths. If all of these vectors have length at most ϵ, then we shall accept P_i for the calculation. Since the linearization we seek satisfies $M(P_i - P) \approx P_{i+1} - F(P)$, we build matrices A and B for which the kth column of A is a vector $P_i - P$ and the kth column of B is the corresponding vector $P_{i+1} - F(P)$. Then, our local linearization matrix M will be the 5×5 solution of the matrix equation $MA = B$. We solve this equation using a singular value decomposition routine taken from [20]. We restrict the matrices A, B to have a minimum of 15 and maximum of 150 columns.

Once the local linearization matrices have been computed, we must compute the exponents. Recall from section 3 that the ERL exponents of the reconstructed trajectory $P_0, P_1, \ldots$ in $\mathbb{R}^5$ are the values obtained by the limit

$$h_{ER}(P_0, \nu) := \lim_{n\to\infty} \frac{1}{n} \ln \|M_{n-1} M_{n-2} \cdots M_1 M_0 \nu\|$$

for unit vectors $\nu \in \mathbb{R}^5$, where $M_i = M(P_i)$ is the 5×5 Eckmann–Ruelle linearization at the point P_i in the trajectory. To evaluate this limit, we employ the treppen-iteration algorithm suggested in [1, 2]. Given our sequence M_i of linearization matrices, we use the QR matrix decomposition to find orthogonal matrices Q_i and upper-triangular matrices R_i (with nonnegative diagonal elements) such that

$$M_i Q_{i-1} = Q_i R_i \qquad \text{for } i = 0, 1, 2, \ldots,$$

where we take Q_{-1} to be the $m \times m$ identity matrix. Then, we can write

$$M_{n-1} M_{n-2} \cdots M_1 M_0 = Q_{n-1} R_{n-1} R_{n-2} \cdots R_0.$$

The orthogonal matrix Q_{n-1} will not affect the matrix norm. Thus,

$$h_{ER}(P_0, \nu) = \lim_{n\to\infty} \frac{1}{n} \ln \|R_{n-1} R_{n-2} \cdots R_1 R_0 \nu\| .$$

The product $R_{n-1} \cdots R_0$ is upper-triangular, and its diagonal elements are the eigenvalues of the matrix, which we expect will grow like the Lyapunov exponents. As mentioned in section 3, Theorem 3.2 justifies reading the exponents directly from the diagonal:

$$\lambda_k = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln\left((R_i)_{kk}\right) \qquad \text{for } k = 1, \ldots, m.$$

Thus, we are able to compute the ERL exponents from our linearizations by using the diagonal elements from the QR decomposition.

We conducted numerical experiments on the example described in section 2.4, using the Hénon map $f(x, y) = (1.4 - x^2 + 0.3y, x)$ and the measurement function $\pi(x, y) = (x, y, x^2, y^2, xy)$. In that section, we computed the Eckmann–Ruelle linearization M explicitly with respect to standard coordinates in $\mathbb{R}^5$:

$$M(\pi(x,y)) = \begin{pmatrix} 0 & 0.3 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1.2xy - 8x^3 & 0.84 + 0.6x^2 & -2.8 + 6x^2 - 0.6y & .09 & -1.2x \\ 0 & 0 & 1 & 0 & 0 \\ 1.4 + 3x^2 & 0 & -3x & 0 & 0.3 \end{pmatrix}.$$

Experiments with trajectories of 300,000 data points in $\mathbb{R}^5$ indicate that the linearization matrices computed by our algorithm match this formula. For example, at the point $\pi(1.555478, 0.398567) \in \mathbb{R}^5$ on the reconstructed Hénon attractor, the linearization matrix for a neighborhood of radius $\epsilon = 0.01$ was computed to be

$$\begin{pmatrix} 0.000000 & 0.300000 & -1.000000 & 0.000000 & 0.000000 \\ 1.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ -29.365725 & 2.290703 & 11.478416 & 0.090305 & -1.866085 \\ 0.000000 & 0.000000 & 1.000000 & 0.000000 & 0.000000 \\ 8.657187 & -0.001385 & -4.666094 & 0.000326 & 0.300723 \end{pmatrix}.$$

This computed matrix differs from the matrix predicted by the previous formula by less than 0.0018 in each component, while the matrix norm of the difference is 0.0030. At another point, $\pi(-1.741541, 1.753985)$, on the reconstructed Hénon attractor in $\mathbb{R}^5$, the local linearization for $\epsilon = 0.01$ is computed to be

$$\begin{pmatrix} 0.000000 & 0.300000 & -1.000000 & 0.000000 & 0.000000 \\ 1.000000 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\ 38.590278 & 2.658478 & 14.345224 & 0.090317 & 2.089740 \\ 0.000000 & 0.000000 & 1.000000 & 0.000000 & 0.000000 \\ 10.498785 & -0.000234 & 5.224592 & 0.000067 & 0.300000 \end{pmatrix}.$$

This computed matrix differs from the predicted matrix by less than 0.0014 in each component, while the matrix norm of the difference is 0.0015. We can also graph the convergence (in matrix norm) of the computed local linearizations to the Eckmann–Ruelle linearization as the neighborhood radius shrinks to zero. See Figure 2.

Using 300,000 data points, we computed the Lyapunov exponents for this example. Recall that the true Lyapunov exponents of the Hénon map are approximately $\lambda = 0.42$ and $\mu = -1.62$. We computed ERL exponents of $0.886200 \approx 2\lambda$, $0.422459 \approx \lambda$, $-1.195450 \approx \lambda + \mu$, $-1.636780 \approx \mu$, and $-3.183227 \approx 2\mu$. These values are consistent with the Lyapunov exponent formula given in Theorem 3.3. Since the Lyapunov exponent formula holds for any generic map (specifically, for one-to-one maps with linearly independent derivative vectors) and is independent of the measurement function π used, we expect to compute these same exponents when different measurement functions are used. Changing the measurement function

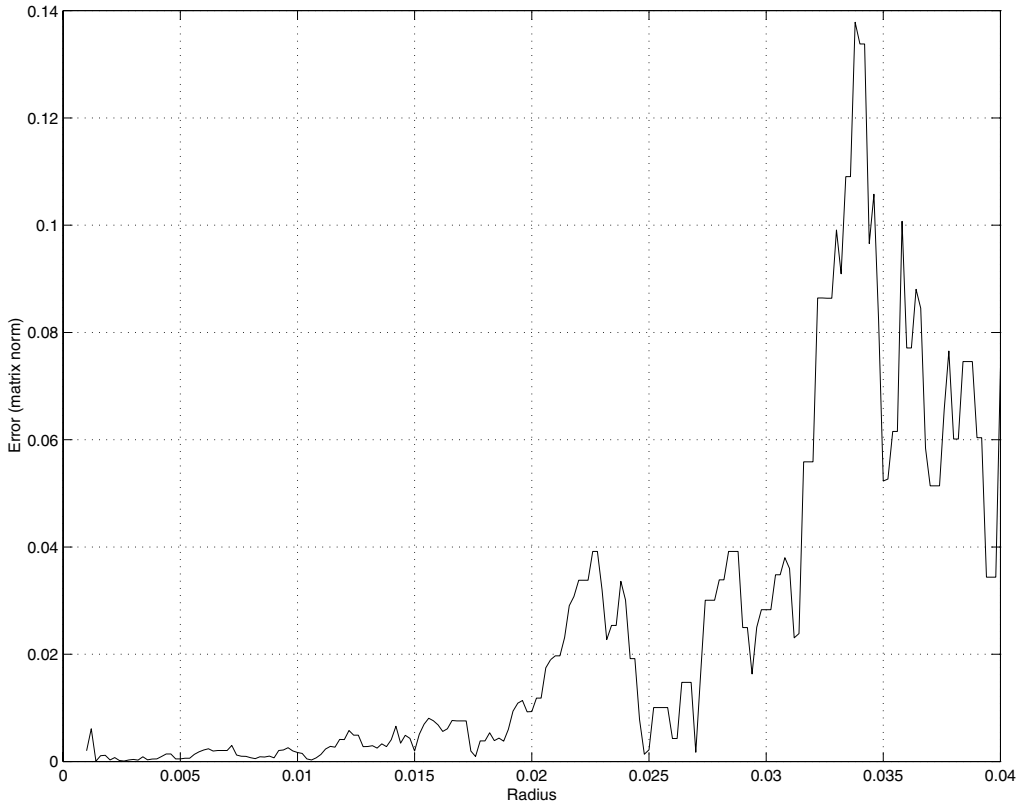


Figure 2. *Graph of the difference (in matrix norm) between the computed linearization and the Eckmann–Ruelle linearization at the point* $(-1.741541, 1.753985)$ *for the Hénon map reconstructed in* $\mathbb{R}^5$*. The calculation used one million data points.*

may alter the speed of convergence somewhat, but we still expect the results to be close.

We also computed the Lyapunov exponents from time-delay reconstructions of the Hénon system. First, we generated a time series by recording the x-coordinate of each point in a trajectory of the Hénon map. We collected 1.5 million data points in this way, and we then reconstructed the attractor in $\mathbb{R}^5$ using the standard time-delay method. Applying our computer program to this data set, we computed these exponents: $0.846902 \approx 2\lambda$, $0.417802 \approx \lambda$, $-1.173343 \approx \lambda + \mu$, $-1.623393 \approx \mu$, and $-2.848050 \approx 2\mu$. Notice that the correct exponents, λ and μ, are determined accurately, and there are small errors in the computed values for $\lambda + \mu$ and 2μ.

A similar time-delay reconstruction was performed using a different data set. We obtained 1.5 million data points by recording $y - xy$ for each point (x, y) in a trajectory of the Hénon map. (We used the same trajectory as in the previous time-delay reconstruction.) For this data set, we computed exponents $0.834372 \approx 2\lambda$, $0.418864 \approx \lambda$, $-1.202808 \approx \lambda+\mu$, $-1.623246 \approx \mu$, and $-2.886969 \approx 2\mu$. Again, the correct exponents appear. In addition, the $\lambda+\mu$ term appears to be exactly correct, while there has been a small improvement in the 2μ term.

The most likely explanation for the observed deviation of the 2μ term from its predicted value is that the neighborhood radii used in the computer program were too large to accurately

determine the 2μ term. (We allowed variable radii from $\epsilon = 0.0025$ to $\epsilon = 0.1$.) With more data and smaller neighborhood radii in the computations, that last exponent will converge to its proper value. Note that the true Lyapunov exponents appear in each of these examples, suggesting that they converge to their correct values rather quickly. This is not surprising if one looks back at the proofs of Theorem 2.2 in section 2 and Theorem 3.3 in section 3. Any matrix that does not map the first-order terms of the Taylor series correctly will have $O(\|\Delta P\|)$ error, instead of the $O(\|\Delta P\|^3)$ error of the Eckmann–Ruelle linearization. Fortunately, algorithms to compute local linearizations will instead find matrices with error $O(\|\Delta P\|^2)$ or better, and these matrices will map the first-order Taylor terms correctly. This ensures that the true Lyapunov exponents will be among the first of the computed exponents to converge.

5. Conclusions. In this paper, we investigated the expected output of the Eckmann–Ruelle procedure for computing Lyapunov exponents of reconstructed dynamics. We assumed ideal circumstances in order to determine what the algorithm should ideally produce. In particular, we assumed that we knew the dimension of the underlying dynamics, and we embedded in a carefully chosen reconstruction dimension. We assumed that the underlying Lyapunov exponents exist, that we had infinitely long data trajectories without noise, and that numerical methods converge to the proper objects. Therefore, since the Eckmann–Ruelle procedure attempts to find the best local linearization matrix at each point on the trajectory, we could use the unique Eckmann–Ruelle linearization for our Lyapunov exponent calculations. With these linearizations, we computed the Lyapunov exponents of the reconstructed dynamics. We showed that a two-dimensional system with Lyapunov exponents λ and μ, when reconstructed in a five-dimensional reconstruction space via a generic measurement function, will have exponents λ, μ, 2λ, $\lambda + \mu$, and 2μ, independent of the measurement function.

This result is important because it shows that, even under ideal conditions, if the underlying dynamics has a positive Lyapunov exponent, then the largest Lyapunov exponent for the reconstructed dynamics will always be spurious, independent of the measurement function. Thus, in practice, one should not use size criteria to distinguish which exponents are true and which are spurious. In the presence of noise and other practical and numerical considerations, the same conclusion should hold.

We recognize that there are still important questions to be answered, notably how to determine the dimension of the underlying dynamics. Our main point is that one should be careful when trying to tease out the true exponents from the spurious, even in situations where the underlying dimension is known.

Given the difficulties that many authors have found with separating the true and spurious exponents, and given observations such as those in the present paper or those in [4], which show that a spurious exponent can actually achieve arbitrary magnitude, perhaps the best method would be one that removes spurious exponents from the process altogether. As noted in the introduction, a number of authors have suggested projecting the reconstructed dynamics to the attractor's tangent space. However, there is no reason to expect the attractor to have a tangent space. Ott and Yorke [17] describe a generalized tangent space of an attractor such that the dimension of the tangent space can unavoidably change from point to point on the attractor. We intend to explore these directions in future work.

Acknowledgments. The authors wish to thank the referees for their insight and comments.

REFERENCES

[1] J.-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*, Rev. Modern Phys., 57 (1985), pp. 617–656.

[2] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, and S. Ciliberto, *Liapunov exponents from time series*, Phys. Rev. A (3), 34 (1986), pp. 4971–4979.

[3] M. Sano and Y. Sawada, *Measurement of the Lyapunov spectrum from a chaotic time series*, Phys. Rev. Lett., 55 (1985), pp. 1082–1085.

[4] W. D. Dechert and R. Gençay, *The topological invariance of Lyapunov exponents in embedded dynamics*, Phys. D, 90 (1996), pp. 40–55.

[5] J. A. Tempkin, *Spurious Lyapunov Exponents Computed Using the Eckmann-Ruelle Procedure*, Ph.D. dissertation, Department of Mathematics, University of Maryland, College Park, MD, 1999; available online at http://www.arXiv.org/nlin.CD/0003007.

[6] F. Takens, *Detecting strange attractors in turbulence*, in Lecture Notes in Math. 898, Springer-Verlag, New York, 1981, pp. 366–381.

[7] N. Packard, J. Crutchfield, J. D. Farmer, and R. Shaw, *Geometry from a time series*, Phys. Rev. Lett., 45 (1980), pp. 712–716.

[8] T. Sauer, J. A. Tempkin, and J. A. Yorke, *Spurious Lyapunov exponents in attractor reconstruction*, Phys. Rev. Lett., 80 (1998), pp. 4341–4344.

[9] T. Sauer and J. A. Yorke, *Reconstructing the Jacobian from data with observational noise*, Phys. Rev. Lett., 83 (1999), pp. 1331–1334.

[10] M. E. Mera and M. Morán, *Convergence of the Eckmann and Ruelle algorithm for the estimation of Liapunov exponents*, Ergodic Theory Dynam. Systems, 20 (2000), pp. 531–546.

[11] U. Parlitz, *Identification of true and spurious Lyapunov exponents from time series*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 2 (1992), pp. 155–165.

[12] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, Cambridge, UK, 1997.

[13] R. Brown, P. Bryant, and H. Abarbanel, *Computing the Lyapunov spectrum of a dynamical system from an observed time series*, Phys. Rev. A (3), 43 (1991), pp. 2787–2806.

[14] J. J. Healey, D. S. Broomhead, K. A. Cliffe, R. Jones, and T. Mullin, *The origins of chaos in a modified Van der Pol oscillator*, Phys. D, 48 (1991), pp. 322–339.

[15] A. Schenck zu Schweinsberg, T. Ritz, and U. Dressler, *Quasicontinuous control of a bronze ribbon experiment using time-delay coordinates*, Phys. Rev. E (3), 55 (1997), pp. 2145–2158.

[16] M. E. Mera and M. Morán, *Degrees of freedom of a time series*, J. Statist. Phys., 106 (2002), pp. 125–145.

[17] W. Ott and J. A. Yorke, *Learning about reality from observation*, SIAM J. Appl. Dyn. Syst., 2 (2003), pp. 297–322.

[18] T. Sauer, J. A. Yorke, and M. Casdagli, *Embedology*, J. Statist. Phys., 65 (1991), pp. 579–616.

[19] E. Ott, T. Sauer, and J. A. Yorke, *Coping with Chaos: Analysis of Chaotic Data and the Exploitation of Chaotic Systems*, Wiley Interscience, New York, 1994.

[20] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed., Cambridge University Press, Cambridge, UK, 1997.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 2, pp. 475–493

Stability Crossing Curves of Shifted Gamma-Distributed Delay Systems*

Constantin-Irinel Morărescu†, Silviu-Iulian Niculescu‡, and Keqin Gu§

Abstract. This paper characterizes the stability crossing curves of a class of linear systems with gamma-distributed delay with a gap. First, we describe the crossing set, i.e., the set of frequencies where the characteristic roots may cross the imaginary axis as the parameters change. Then, we describe the corresponding stability crossing curves, i.e., the set of parameters such that there is at least one pair of characteristic roots on the imaginary axis. Such stability crossing curves divide the parameter space $\mathbb{R}^2_+$ defined by the mean delay and the gap into different regions. Within each such region, the number of characteristic roots on the right half complex plane is fixed. This naturally describes the regions of parameters where the system is stable. The classification of the stability crossing curves is also discussed. Some illustrative examples (Cushing equation in biology, traffic flow models in transportation systems, and control over networks of a simplified helicopter model) are also presented.

Key words. stability, crossing curves, distributed delay, quasipolynomial, gamma-distribution

AMS subject classifications. 34K20, 34D99, 93D09, 93D99

DOI. 10.1137/060670766

1. Introduction.

The stability of dynamical systems in the presence of time delay is a problem of recurring interest (see, for instance, [11, 16, 8, 14] and the references therein). The presence of a time delay may induce instabilities and complex behaviors. Systems with distributed delays are present in many scientific disciplines such as physiology, population dynamics, and engineering.

One of the first studies devoted to population dynamics using a model with gamma-distributed delay is due to Cushing [5]. The linearization of this model is

$$\dot{x}(t) = -\alpha x(t) + \beta \int_{-\infty}^{t} g(t-\theta)x(\theta)d\theta, \tag{1.1}$$

where α is a constant defining the death rate per unit time, and β is a constant corresponding

*Received by the editors September 26, 2006; accepted for publication (in revised form) by B. Krauskopf March 20, 2007; published electronically June 29, 2007.
http://www.siam.org/journals/siads/6-2/67076.html

†Department of Mathematics, University "Politehnica" of Bucharest, 313, Splaiul Independenţei, RO-060032, Bucharest, Romania (constantin.morarescu@math.pub.ro). The work of this author was partially supported by a European Community Marie Curie Fellowship and in the framework of the CTS, contract HPMT-CT-2001-00278.

‡Corresponding author. Laboratoire de Signaux et Systèmes (L2S), Supélec, 3, rue Joliot Curie, 91190, Gif-sur-Yvette, France (Silviu.Niculescu@lss.supelec.fr). On leave from HeuDiaSyC (UMR CNRS 6599), Université de Technologie de Compiègne, Centre de Recherche de Royallieu, BP 20529, 60205, Compiègne, France. The work of this author was partially funded by the CNRS-US grant *Delays in interconnected dynamical systems: Analysis, and applications* (2005-2007).

§Department of Mechanical and Industrial Engineering, Southern Illinois University at Edwardsville, Edwardsville, IL 62026-1805 (kgu@siue.edu). The work of this author was partially funded by the CNRS-US grant *Delays in interconnected dynamical systems: Analysis, and applications* (2005-2007).

to the maternity function. The integration kernel of the distributed delay is the gamma-distribution [15, 4]

$$g(\xi) = \frac{a^{n+1}}{n!}\xi^n e^{-a\xi}. \tag{1.2}$$

Applying a Laplace transform of (1.1) with $g(\xi)$ as expressed in (1.2) yields a parameter-dependent polynomial characteristic equation of the form

$$D(s;\, \bar{\tau}, n) := (s+\alpha)\left(1 + s\frac{\bar{\tau}}{n+1}\right)^{n+1} - \beta = 0, \tag{1.3}$$

where s is the Laplace transform variable and $\bar{\tau} = (n+1)/a$ is the *mean delay*.

Cooke and Grossman [4] discussed the change of stability of (1.3) when one of the parameters, the mean delay value $\bar{\tau}$ or the exponent n, varies while the other is fixed.

Nisbet and Gurney [17] modified the gamma-distribution $g(\xi)$ expressed in (1.2) to the gamma-distribution with a gap

$$\hat{g}(\xi) = \begin{cases} 0, & \xi < \tau, \\ \frac{a^{n+1}}{n!}(\xi-\tau)^n e^{-a(\xi-\tau)}, & \xi \geq \tau, \end{cases} \tag{1.4}$$

to more accurately reflect the reality (see, for instance, [1, 15] for additional discussions). In this case, a simple computation shows that the mean delay is $\hat{\tau} = \tau + \frac{n+1}{a}$. The characteristic equation becomes a parameter-dependent quasipolynomial equation [1, 2] of the form

$$\hat{D}(s;\, \bar{\tau}, \tau, n) := (s+\alpha)\left(1 + s\frac{\bar{\tau}}{n+1}\right)^{n+1} - \beta e^{-s\tau} = 0. \tag{1.5}$$

We note that [2] pointed out some inaccuracies of the earlier results presented in [4] and [1]. It is also interesting to mention that Farkas, Farkas, and Szabó [7] studied the bifurcation problem of the predator-prey model also in the case of a gamma-distributed delay. More general bifurcation study of systems with distributed delay can be found in a book by Farkas [6].

More recently, and in a quite different field (engineering), it was pointed out that such gamma-distributed delays with a gap can also be encountered in the problem of controlling objects over communication networks [19]. More specifically, the overall communication delay in the network is modeled by a gamma-distributed delay with a gap, where the gap value corresponds to the minimal propagation delay in the network, which is always strictly positive. The stability problem of the closed-loop system in [19] reduces to a parameter-dependent characteristic quasipolynomial equation of the form

$$D(s;\, \bar{\tau}, \tau, n) := P(s)\left(1 + s\frac{\bar{\tau}}{n+1}\right)^{n+1} + Q(s)e^{-s\tau} = 0, \tag{1.6}$$

where $P(s)$ and $Q(s)$ are polynomials. Obviously, (1.5) is a special case of (1.6).

Another research area where a distributed delay appears naturally is the traffic flow dynamics. A simplified car-following model, where multiple vehicles in a ring have drivers with

identical behavior, and under the influence of a single constant time delay [3, 12], can be written as

$$\dot{x}_i(t) = \alpha_i(x_{i-1}(t-\tau) - x_i(t-\tau)), \quad i = 1, \ldots, p, \tag{1.7}$$

where p is the number of vehicles considered and $x_0 = x_p$. The left-hand side represents the acceleration of the ith vehicle, and the right-hand side expresses the velocity difference of consecutive vehicles.

When the delays are not assumed to be identical, several models in the literature are used to describe the dynamics of the model (see, for instance, [21] for some classifications and a large list of references). As suggested in [22], we can extend the previous models by incorporating a more general memory effect. Consider the system

$$\dot{x}_i(t) = \alpha_i \int_{-\infty}^{t} g(t-\theta)(x_{i-1}(\theta) - x_i(\theta))\mathrm{d}\theta, \tag{1.8}$$

where g is the delay distribution, which can represent both dead time and past memory. The corresponding characteristic equation of (1.8) is given by

$$\det[sI - (A_1 + A_2)G(s)] = 0, \tag{1.9}$$

where G denotes the Laplace transform of g. When g represents the gamma-distribution with a gap, we get

$$G(s) = \mathrm{e}^{-s\tau}\left(1 + s\frac{\bar{\tau}}{n+1}\right)^{-(n+1)}.$$

In the simplest case of two vehicles in a ring ($p = 2$, $i = 1, 2$, and $x_0 = x_2$), the matrices A_1 and A_2 are given by

$$A_1 = \begin{pmatrix} -\alpha_1 & 0 \\ 0 & -\alpha_2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & \alpha_1 \\ \alpha_2 & 0 \end{pmatrix}, \tag{1.10}$$

and the characteristic equation becomes

$$s\left(1 + s\frac{\bar{\tau}}{n+1}\right)^{n+1} + (\alpha_1 + \alpha_2)\mathrm{e}^{-s\tau} = 0, \tag{1.11}$$

which is again a special case of (1.6).

Finally, another interesting engineering example involving gamma-distributed delay is the machine tool vibration problem. The readers are referred to the nice paper by Stépán [24] for details. It is also interesting to mention that Insperger and Stépán [13] also used gamma-distributed delay in their numerical study of time delay systems.

In this paper, we will study the stability of systems with the characteristic equation (1.6) as the parameters $\bar{\tau}$ and τ vary. Specifically, we will describe the stability crossing curves, i.e., the set of parameters such that there exist at least one pair of characteristic roots on the imaginary axis. Such stability crossing curves divide the parameter space $\mathbb{R}^2_+$ into different regions. Within each such region, the number of characteristic roots on the right half complex plane is fixed. This naturally describes the regions of parameters where the system is stable.

It should be noted that there have been numerous works in the literature to describe the stability regions of parameter space, known as stability charts [23, 24]. These descriptions are typically valid for one specific system except that the parameters are allowed to vary. In a recent paper, Gu, Niculescu, and Chen [9] gave a characterization of the stability crossing curves for systems with two discrete delays as the parameters. One significant difference of [9] as compared to the stability charts is the fact that such characterization applies to any system within the class, i.e., any system with two delays. We note also the paper [20], where we can find some interesting characterizations that can be used for a large class of time delay systems (including distributed delay systems). However, the approach proposed in [20] requires rather heavy computation when dealing with quasipolynomials of a high degree.

The current paper follows the reasoning of [9], and our conclusion is valid for any system of the form (1.6).

The rest of this paper is organized as follows. Section 2 contains the problem statement and assumptions. Section 3 first discusses geometric characterization of the crossing curves. Next, the stability crossing curves together with their classifications are described. Several simple examples are described to illustrate the types of curves in our classification. Finally, tangent and smoothness properties and crossing direction are described. Section 4 includes four illustrative examples in some detail: a linearized first-order Cushing equation, a second-order system depicting some particular behavior, a simplified traffic flow model, and a control over networks of a simplified helicopter model. Some concluding remarks end the paper.

Throughout the paper, the following standard notation is used: $\mathbb{C}$ ($\mathbb{C}^+$, $\mathbb{C}^-$) is the set of complex numbers (with strictly positive and strictly negative real parts), and $j = \sqrt{-1}$. For $z \in \mathbb{C}$, $\angle(z)$, $\mathrm{Re}(z)$, and $\mathrm{Im}(z)$ define the argument, the real part, and the imaginary part of z. $\mathbb{R}$ ($\mathbb{R}^+$, $\mathbb{R}^-$) denotes the set of real numbers (greater than or equal to zero, less than or equal to zero). Next, $\mathbb{N}$ is the set of natural numbers, including zero, and $\mathbb{Z}$ is the set of integers. Finally, RHP denotes the right half plane of $\mathbb{C}$.

2. Problem formulation.

Consider a system with the characteristic equation

$$D(s;T,\tau) := P(s)(1+sT)^n + Q(s)\mathrm{e}^{-s\tau} = 0, \tag{2.1}$$

where the two parameters T and τ are nonnegative. We will describe the *stability crossing curves* $\mathcal{T}$, which is the set of (T,τ) such that (2.1) has imaginary solutions.

As the parameters (T,τ) cross the stability crossing curves, some characteristic roots cross the imaginary axis. Therefore, the number of roots on the right half complex plane are different on the two sides of the curves, from which we may describe the parameter regions of (T,τ) in $\mathbb{R}^2_+$ for the system to be stable.

Another related useful concept is the *crossing set* Ω, which is defined as the collection of all $\omega > 0$ such that there exists a parameter pair (T,τ) such that $D(j\omega;T,\tau) = 0$. In other words, as the parameters T and τ vary, the characteristic roots may cross the imaginary axis at $j\omega$ if and only if $\omega \in \Omega$.

We will restrict our discussions to the systems that satisfy the following assumptions:

Assumption I. $\deg(Q) < \deg(P)$.

Assumption II. $P(0) + Q(0) \neq 0$.

Assumption III. $P(s)$ and $Q(s)$ do not have common zeros.

Assumption IV. If $P(s) = p$, $Q(s) = q$, where p and q are constant real, then $|p| \neq |q|$.
Assumption V. $P(0) \neq 0$, $|P(0)| \neq |Q(0)|$.
Assumption VI. $P'(j\omega) \neq 0$ whenever $P(j\omega) = 0$.

Assumption I means that the time delay system represented by (2.1) is of retarded type. While not discussed here, it is possible to extend the analysis to systems of neutral type by relaxing this assumption to also allow $\deg(Q) = \deg(P)$, as long as $\lim_{s\to\infty} Q(s)/P(s) < 1$ is satisfied. Assumption II is made to exclude some trivial cases. If it is not satisfied, then $s = 0$ is a solution of (2.1) for arbitrary (T, τ), and therefore, the system can never be stable. Regarding Assumption III, if it is violated, we may find a common factor of the highest order $c(s) \neq$ constant of $P(s)$ and $Q(s)$. This would indicate that $D(s; T, \tau) = c(s)\hat{D}(s; T, \tau)$, where $\hat{D}(s; T, \tau)$ satisfies Assumption III, and our analysis can still proceed on $\hat{D}(s; T, \tau)$. Finally, Assumptions IV–VI are made to exclude some rare singular cases in order to simplify presentation.

Notice that we have restricted any element ω of the crossing set Ω to satisfy $\omega > 0$. Indeed, the discussion of $\omega < 0$ is redundant in view of the fact that $D(-j\omega; T, \tau)$ is the complex conjugate of $D(j\omega; T, \tau)$. Also, $\omega = 0$ is never an element of Ω in view of Assumption II.

3. Main results.

3.1. Crossing set and stability crossing curves. Consider a fixed $\omega > 0$. We first observe that as T and τ each vary within $[0, \infty)$, i.e., (T, τ) vary in $\mathbb{R}^2_+$, $|1 + j\omega T|^n \in [1, \infty)$, $|e^{j\omega\tau}| = 1$, and $\angle e^{j\omega\tau}$ may assume any nonnegative value by choosing appropriate τ. From this observation, it is not difficult to conclude the following proposition.

Proposition 3.1. *Given any $\omega > 0$, $\omega \in \Omega$ if and only if it satisfies*

$$0 < |P(j\omega)| \leq |Q(j\omega)|, \tag{3.1}$$

and all the corresponding T, τ can be calculated by

$$T = \frac{1}{\omega}\left(\left|\frac{Q(j\omega)}{P(j\omega)}\right|^{2/n} - 1\right)^{1/2}, \tag{3.2}$$

$$\tau = \tau_m = \frac{1}{\omega}(\angle Q(j\omega) - \angle P(j\omega) - n\arctan(\omega T) + \pi + m2\pi), \quad m = 0, \pm 1, \pm 2, \ldots. \tag{3.3}$$

Proof. For necessity of (3.1), let $\omega \in \Omega$, and apply modulus to (2.1). We obtain

$$|(1 + j\omega T)^n|\,|P(j\omega)| = |Q(j\omega)|\,. \tag{3.4}$$

This implies $|P(j\omega)| \leq |Q(j\omega)|$, because $|(1 + j\omega T)^n| \geq 1$. In addition, $|P(j\omega)| > 0$ is also necessary. Otherwise, $P(j\omega) = 0$, which implies $Q(j\omega) = 0$ in view of (3.4). But this violates Assumption III.

For sufficiency of (3.1), we need only to recognize that T and τ given by (3.2) and (3.3) make $s = j\omega$ a solution of (2.1). It is also easy to see by direct solution that T and τ given by (3.2) and (3.3) are all the solutions. ∎

Remark 1. If for all $\omega \in \mathbb{R}_+$, $|P(j\omega)| > |Q(j\omega)|$, then there does not exist any crossing root with respect to the imaginary axis, and the system (2.1) is *hyperbolic,* as suggested by Hale, Infante, and Tsen [10].

There are only a finite number of solutions to each of the two equations

$$P(j\omega) = 0 \tag{3.5}$$

and

$$|P(j\omega)| = |Q(j\omega)|, \tag{3.6}$$

because P and Q are both polynomials satisfying Assumptions I–IV. Therefore, Ω, which is the collection of ω satisfying (3.1), consists of a finite number of intervals. Denote these intervals as $\Omega_1, \Omega_2, \ldots, \Omega_N$. Then

$$\Omega = \bigcup_{k=1}^{N} \Omega_k.$$

Without loss of generality, we may order these intervals from left to right; i.e., for any $\omega_1 \in \Omega_{k_1}$, $\omega_2 \in \Omega_{k_2}$, $k_1 < k_2$, we have $\omega_1 < \omega_2$.

In order to give a geometric interpretation that allows for deriving the crossing set Ω, for $s = j\omega$, we rewrite (2.1) as

$$\left(-\frac{Q(j\omega)}{P(j\omega)}\right)^{1/n} \mathrm{e}^{-j\omega\tau/n} = 1 + j\omega T. \tag{3.7}$$

The equation (3.7) can be interpreted as the intersection between a circle with radius $|Q(j\omega)/P(j\omega)|^{1/n}$ and a vertical line passing through the point $(1, 0)$ in the complex plane. Therefore, the characterization of Ω can be easily derived from Figure 1. We will not restrict $\angle Q(j\omega)$ and $\angle P(j\omega)$ to a 2π range. Rather, we allow them to vary continuously within each interval Ω_k. Thus, for each fixed m, (3.2) and (3.3) represent a continuous curve. We denote such a curve as $\mathcal{T}_k^m$. Therefore, corresponding to a given interval Ω_k, we have an infinite number of continuous stability crossing curves $\mathcal{T}_k^m$, $m = 0, \pm 1, \pm 2, \ldots$. It should be noted that, for some m, part of the curve or the entire curve may be outside of the range $\mathbb{R}_+^2$ and therefore may not be physically meaningful.

The collection of all the points in $\mathcal{T}$ corresponding to Ω_k may be expressed as

$$\mathcal{T}_k = \bigcup_{m=-\infty}^{+\infty} \left(\mathcal{T}_k^m \bigcap \mathbb{R}_+^2\right).$$

Obviously, $\mathcal{T} = \bigcup_{k=1}^{N} \mathcal{T}_k$.

3.2. Classification of stability crossing curves. Let the left and right end points of interval Ω_k be denoted as ω_k^ℓ and ω_k^r, respectively. Due to Assumptions IV and V, it is not difficult to see that each end point ω_k^ℓ or ω_k^r must belong to one, and only one, of the following three types:

Type 1. It satisfies (3.6).

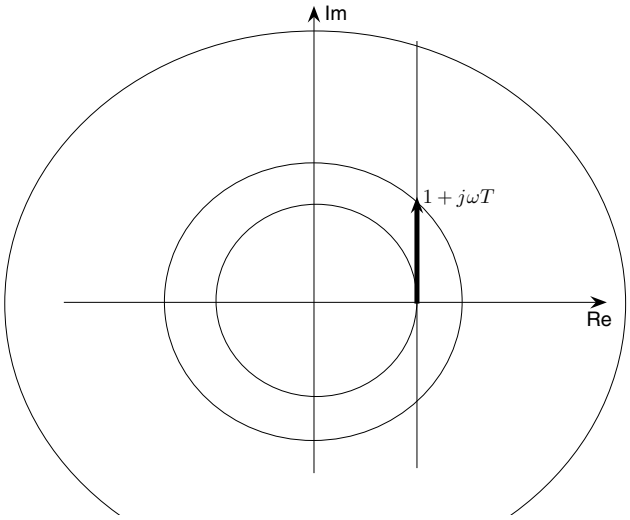


Figure 1. *The intersection is possible only if the radius $|Q(j\omega)/P(j\omega)|$ is larger than 1. The extreme cases for intersection are given by $|Q(j\omega)/P(j\omega)| = 1$ or $|Q(j\omega)/P(j\omega)| \to \infty \Leftrightarrow P(j\omega) \to 0$.*

Type 2. It satisfies (3.5).

Type 3. It equals 0.

Denote an end point as ω_0, which may be either a left end or a right end of an interval Ω_k. Then the corresponding points in $\mathcal{T}_k^m$ may be described as follows.

If ω_0 is of type 1, then $T = 0$. In other words, $\mathcal{T}_k^m$ intersects the τ-axis at $\omega = \omega_0$.

If ω_0 is of type 2, then as $\omega \to \omega_0$, $T \to \infty$ and

$$\tau \to \frac{1}{\omega_0}\left(\angle Q(j\omega_0) - \lim_{\omega\to\omega_0} \angle P(j\omega) - \frac{n\pi}{2} + \pi + m2\pi\right). \tag{3.8}$$

Obviously,

$$\lim_{\omega\to\omega_0} \angle P(j\omega) = \angle\left[\frac{d}{d\omega}P(j\omega)\right]_{\omega\to\omega_0} \tag{3.9}$$

if ω_0 is the left end point ω_k^ℓ of Ω_k, and

$$\lim_{\omega\to\omega_0} \angle P(j\omega) = \angle\left[\frac{d}{d\omega}P(j\omega)\right]_{\omega\to\omega_0} + \pi \tag{3.10}$$

if ω_0 is the right end point ω_k^r of Ω_k. In other words, $\mathcal{T}_k^m$ approaches a horizontal line.

Obviously, only ω_1^ℓ may be of type 3. Due to nonsingularity assumptions, if $\omega_1^\ell = 0$, we must have $0 < |P(0)| < |Q(0)|$. In this case, as $\omega \to 0$, both T and τ approach ∞. In fact, (T, τ) approaches a straight line with slope

$$\tau/T \to \frac{(\angle Q(0) - \angle P(0) - n\arctan\alpha + \pi + m2\pi)}{\alpha}, \tag{3.11}$$

where

$$\alpha = \left(\left| \frac{Q(0)}{P(0)} \right|^{2/n} - 1 \right)^{1/2}.$$

We say an interval Ω_k is of type ℓr if its left end is of type ℓ and its right end is of type r. We may accordingly divide these intervals into the following six types.

Type 11. In this case, $\mathcal{T}_k^m$ starts at a point on the τ-axis and ends at another point on the τ-axis.

Type 12. In this case, $\mathcal{T}_k^m$ starts at a point on the τ-axis, and the other end approaches ∞ along a horizontal line.

Type 21. This is the reverse of type 12. $\mathcal{T}_k^m$ starts at ∞ along a horizontal line and ends at the τ-axis.

Type 22. In this case, both ends of $\mathcal{T}_k^m$ approach horizontal lines.

Type 31. In this case, $\mathcal{T}_k^m$ begins at ∞ with an asymptote of slope expressed in (3.11). The other end is at the τ-axis.

Type 32. In this case, $\mathcal{T}_k^m$ again begins at ∞ with an asymptote of slope expressed in (3.11). The other end approaches ∞ along a horizontal line.

In what follows, we present two *academic* examples to illustrate some cases discussed above.

Example 3.1 (type 11). Let $n = 1$, $P(s) = s^2 + 3s + 2$, and $Q(s) = \sqrt{10}s$. Figure 2 (left) plots $|P(j\omega)|/|Q(j\omega)|$ against ω. From the plot, it can be seen that the crossing set Ω contains only one interval $\Omega = \Omega_1 = [1, 2]$ of type 11. Correspondingly, the stability crossing curves $\mathcal{T}$ are shown in Figure 2 (right), which consists of a series of curves with both ends on the τ-axis.

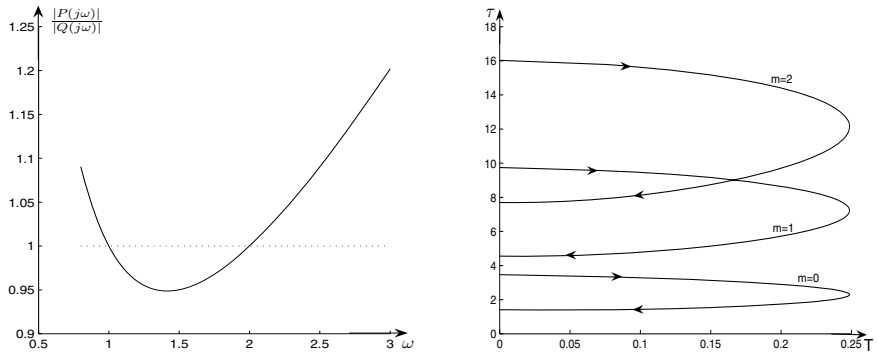


Figure 2. *Example* 3.1. *Left:* $|P(j\omega)|/|Q(j\omega)|$ *against* ω. *Right:* $\mathcal{T}_1^m$ *for* $m = 0, 1, 2$ *(type* 11*).*

Example 3.2 (types 22 and 32). Figure 3 (left) plots $|P(j\omega)|/|Q(j\omega)|$ against ω with $n = 1$:

$$P(s) = s^4 + 3s^2 + 2 \text{ and } Q(s) = s + 4. \tag{3.12}$$

In this case Ω contains three intervals: $\Omega_1 = (0, 1)$ (type 32), $\Omega_2 = (1, \sqrt{2})$ (type 22), and $\Omega_3 = (\sqrt{2}, 1.91]$ (type 21).

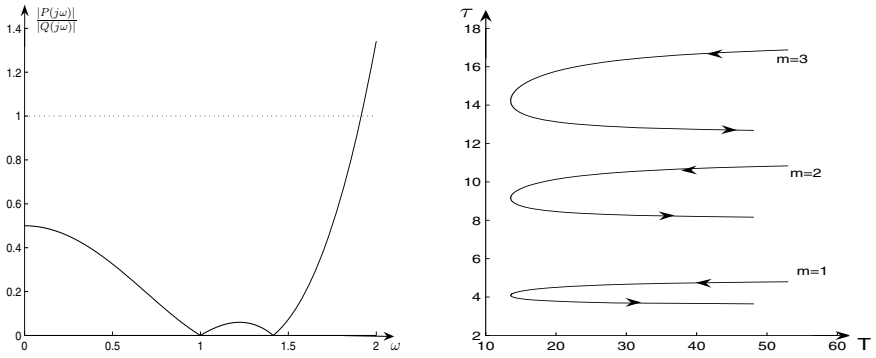


Figure 3. *Example* 3.2. *Left:* $|P(j\omega)|/|Q(j\omega)|$ *against* ω. *Right:* $\mathcal{T}_2^m$, $m = 1, 2, 3$ *(type* 22*).*

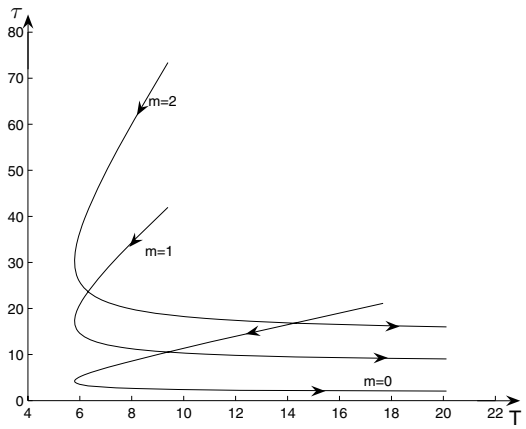


Figure 4. *Example* 3.2. $\mathcal{T}_2^m$, $m = 0, 1, 2$ *(type* 32*).*

The stability crossing curves consist of three series of curves. Since type 21 has already been shown in Example 3.1 above, here we will show only the two series corresponding to Ω_2 and Ω_1. The series corresponding to Ω_2 of type 22 is shown in Figure 3 (right). We can see that both ends approach infinity along the horizontal direction. The series corresponding to Ω_1 of type 32 is shown in Figure 4. The curves start from infinity in directions that can be calculated by (3.11) and end at infinity along the horizontal direction.

Remark 2. Starting from practical models encountered in the literature, we will illustrate other types in section 4. More precisely, the crossing sets of the examples we consider include intervals of the type 31 (linearized Cushing equation, simplified helicopter model), type 12, or type 21 (second-order example, simplified traffic flow model).

3.3. Tangents and smoothness. For a given k we will discuss the smoothness of the curves in $\mathcal{T}_k^m$ and thus of

$$\mathcal{T} = \bigcup_{k=1}^{N} \mathcal{T}_k = \bigcup_{k=1}^{N} \bigcup_{m=-\infty}^{+\infty} \left(\mathcal{T}_k^m \bigcap \mathbb{R}_+^2 \right).$$

In this part we use an approach based on the implicit function theorem.

For this purpose, we consider T and τ as implicit functions of $s = j\omega$ defined by (2.1). For given m and k, as $s = j\omega$ moves along the imaginary axis with $\omega \in \Omega_k$, $(T, \tau) = (T(\omega), \tau(\omega))$ moves along $\mathcal{T}_k^m$. For a given $\omega \in \Omega_k$, let

$$\begin{aligned}
R_0 &= \operatorname{Re}\left(\frac{j}{s}\frac{\partial D(s,T,\tau)}{\partial s}\right)_{s=j\omega} \\
&= \frac{1}{\omega}\operatorname{Re}\left\{\left[nTP(j\omega) + (1+j\omega T)P'(j\omega)\right]\right. \\
&\quad \left.\cdot (1+j\omega T)^{n-1} + (Q'(j\omega) - \tau Q(j\omega))\mathrm{e}^{-j\omega\tau}\right\}, \\
I_0 &= \operatorname{Im}\left(\frac{j}{s}\frac{\partial D(s,T,\tau)}{\partial s}\right)_{s=j\omega} \\
&= \frac{1}{\omega}\operatorname{Im}\left\{\left[nTP(j\omega) + (1+j\omega T)P'(j\omega)\right]\right. \\
&\quad \left.\cdot (1+j\omega T)^{n-1} + (Q'(j\omega) - \tau Q(j\omega))\mathrm{e}^{-j\omega\tau}\right\}, \\
R_1 &= \operatorname{Re}\left(\frac{1}{s}\frac{\partial D(s,T,\tau)}{\partial T}\right)_{s=j\omega} = \operatorname{Re}\left(n(1+j\omega T)^{n-1}P(j\omega)\right), \\
I_1 &= \operatorname{Im}\left(\frac{1}{s}\frac{\partial D(s,T,\tau)}{\partial T}\right)_{s=j\omega} = \operatorname{Im}\left(n(1+j\omega T)^{n-1}P(j\omega)\right), \\
R_2 &= \operatorname{Re}\left(\frac{1}{s}\frac{\partial D(s,T,\tau)}{\partial \tau}\right)_{s=j\omega} = -\operatorname{Re}\left(Q(j\omega)\mathrm{e}^{-j\omega\tau}\right), \\
I_2 &= \operatorname{Im}\left(\frac{1}{s}\frac{\partial D(s,T,\tau)}{\partial \tau}\right)_{s=j\omega} = -\operatorname{Im}\left(Q(j\omega)\mathrm{e}^{-j\omega\tau}\right).
\end{aligned}$$

Then, since $D(s; T, \tau)$ is an analytic function of s, T, and τ, the implicit function theorem indicates that the tangent of $\mathcal{T}_k^m$ can be expressed as

$$\begin{aligned}
\begin{pmatrix} \frac{\mathrm{d}T}{\mathrm{d}\omega} \\ \frac{\mathrm{d}\tau}{\mathrm{d}\omega} \end{pmatrix} &= \begin{pmatrix} R_1 & R_2 \\ I_1 & I_2 \end{pmatrix}^{-1} \begin{pmatrix} R_0 \\ I_0 \end{pmatrix} \\
&= \frac{1}{R_1 I_2 - R_2 I_1} \begin{pmatrix} R_0 I_2 - I_0 R_2 \\ I_0 R_1 - R_0 I_1 \end{pmatrix},
\end{aligned} \tag{3.13}$$

provided that

$$R_1 I_2 - R_2 I_1 \neq 0, \tag{3.14}$$

and $\mathrm{d}T/\mathrm{d}\omega$ and $\mathrm{d}\tau/\mathrm{d}\omega$ do not vanish simultaneously.

It follows that $\mathcal{T}_k$ is smooth everywhere except possibly at the points where either

$$R_1 I_2 - R_2 I_1 = 0 \tag{3.15}$$

or

$$\frac{\mathrm{d}T}{\mathrm{d}\omega} = \frac{\mathrm{d}\tau}{\mathrm{d}\omega} = 0. \tag{3.16}$$

From the above discussions, we can conclude the following proposition.

Proposition 3.2. *The curve $\mathcal{T}_k^m$ is smooth everywhere except possibly at the points corresponding to $s = j\omega$ in either of the following two cases:*

(1) *$s = j\omega$ is a multiple solution of (2.1);*

(2) *ω is a type 1 end point of Ω_k.*

Proof. From the above discussion, we need only to show that (3.15) or (3.16) can be satisfied only in the above two cases.

If (3.16) is satisfied, then, in view of (3.13), $R_0 = I_0 = 0$, which implies

$$\frac{\partial D}{\partial s} = 0.$$

This, together with $D = 0$, means that $s = j\omega$ is a multiple solution of (2.1) in case (1) above.

If condition (3.15) is satisfied, then

$$\frac{I_1}{R_1} = \frac{I_2}{R_2}$$

or

$$\angle\left(n(1+j\omega T)^{n-1}P(j\omega)\right) = \angle\left(-Q(j\omega)\mathrm{e}^{-j\omega\tau}\right).$$

But (2.1) implies

$$\angle\left((1+j\omega T)^{n}P(j\omega)\right) = \angle\left(-Q(j\omega)\mathrm{e}^{-j\omega\tau}\right).$$

Therefore, $\angle(1+j\omega T) = 0$, which in turn means $T = 0$. From this, we can conclude $|P(j\omega)| = |Q(j\omega)|$, and ω is a type 1 end point of Ω_k. ∎

3.4. Direction of crossing. Next we will discuss the direction in which the solutions of (2.1) cross the imaginary axis as (T, τ) deviates from the curve $\mathcal{T}_k^m$. We will call the direction of the curve that corresponds to increasing ω the *positive direction*. We will also call the region on the left-hand side as we head in the positive direction of the curve *the region on the left*.

To establish the direction of crossing we need to consider T and τ as functions of $s = \sigma + j\omega$, i.e., functions of two real variables σ and ω, and partial derivative notation needs to be adopted. Since the tangent of $\mathcal{T}_k^m$ along the positive direction is $\left(\frac{\partial T}{\partial \omega}, \frac{\partial \tau}{\partial \omega}\right)$, the normal to $\mathcal{T}_k^m$ pointing to the left-hand side of the positive direction is $\left(-\frac{\partial \tau}{\partial \omega}, \frac{\partial T}{\partial \omega}\right)$. Corresponding to a pair of complex conjugate solutions of (2.1) crossing the imaginary axis along the horizontal direction, (T, τ) moves along the direction $\left(\frac{\partial T}{\partial \sigma}, \frac{\partial \tau}{\partial \sigma}\right)$. So, as (T, τ) crosses the stability crossing curves from the right-hand side to the left-hand side, a pair of complex conjugate solutions of (2.1) cross the imaginary axis to the right half plane if

$$\left(\frac{\partial T}{\partial \omega}\frac{\partial \tau}{\partial \sigma} - \frac{\partial \tau}{\partial \omega}\frac{\partial T}{\partial \sigma}\right)_{s=j\omega} > 0, \tag{3.17}$$

i.e., the region on the left of $\mathcal{T}_k^m$ gains two solutions on the right half plane. If the inequality (3.17) is reversed then the region on the left of $\mathcal{T}_k^m$ loses two right half plane solutions. Similar to (3.13), we can express

$$\begin{pmatrix} \frac{\partial T}{\partial \sigma} \\ \frac{\partial \tau}{\partial \sigma} \end{pmatrix}_{s=j\omega} = \frac{1}{R_1 I_2 - R_2 I_1} \begin{pmatrix} R_0 R_2 + I_0 I_2 \\ -R_0 R_1 - I_0 I_1 \end{pmatrix}. \tag{3.18}$$

Using this we arrive at the following proposition.

Proposition 3.3. *Let $\omega \in (\omega_k^\ell, \omega_k^r)$ and $(T,\tau) \in \mathcal{T}_k$ such that $j\omega$ is a simple solution of the characteristic equation*

$$D(s;T,\tau) = 0,$$

given by (2.1) and

$$D(j\omega';T,\tau) \neq 0 \qquad \forall \omega' > 0, \quad \omega' \neq \omega$$

(i.e., (T,τ) is not an intersection point of two curves or different sections of a single curve of $\mathcal{T}$).

Then, as (T,τ) crosses the stability crossing curves from the right-hand side to the left-hand side at this point, a pair of solutions of (2.1) cross the imaginary axis to the right, through $s = \pm j\omega$ if $R_2I_1 - R_1I_2 > 0$. The crossing is to the left if the inequality is reversed.

Proof. Direct computation shows that

$$\left(\frac{\partial T}{\partial \omega}\frac{\partial \tau}{\partial \sigma} - \frac{\partial \tau}{\partial \omega}\frac{\partial T}{\partial \sigma}\right)_{s=j\omega} = \frac{(R_0^2 + I_0^2)(R_2I_1 - R_1I_2)}{(R_1I_2 - R_2I_1)^2}.$$

Therefore (3.17) can be written as $R_2I_1 - R_1I_2 > 0$. ∎

4. Illustrative examples. In order to illustrate the cases presented in the previous sections, we shall consider four examples: the linearized Cushing equation with a gap (first-order system), a simplified helicopter model (second-order system), a second-order system encountered in control engineering, and finally a simplified traffic flow model.

Example 4.1 (linearized Cushing equation with a gap). Cushing has formulated and analyzed some general population growth models [5], and some of them have been largely treated in the literature (see, for instance, [1, 2, 4, 15] and the references therein). One of these models leads to the characteristic equation $(s+\alpha)(1+sT)^n + \beta e^{-s\tau} = 0$, where α is the death rate per unit time, and β is a constant corresponding to the maternity function. Based on the particular form of the characteristic equation, it is easy to see that the only interesting case is $|\alpha| < |\beta|$. Otherwise, the crossing set Ω is *empty*.

If $|\alpha| < |\beta|$, then $\Omega = \Omega_1 = (0, \sqrt{\beta^2 - \alpha^2}]$, which is of type 31. The corresponding pairs (T,τ) are given by

$$T = \frac{1}{\omega}\left[\left(\frac{\beta^2}{\omega^2+\alpha^2}\right)^{1/n} - 1\right]^{1/2},$$

$$\tau_m = \frac{1}{\omega}\left[\angle\left(\frac{-\beta}{(\alpha+j\omega)(1+j\omega T)^n}\right) + 2m\pi\right].$$

According to Proposition 3.2, we get

$$\lim_{\omega \to \sqrt{\beta^2-\alpha^2}} T = 0, \quad \lim_{\omega \to 0} T = \infty, \quad \lim_{\omega \to 0} \tau_m = \infty,$$

and

$$\lim_{\omega \to \sqrt{\beta^2-\alpha^2}} \tau_m = \frac{1}{\sqrt{\beta^2-\alpha^2}}\left(2m\pi + \angle\left(\frac{-\beta}{\alpha}\right) - \arctan\frac{\sqrt{\beta^2-\alpha^2}}{\alpha}\right).$$

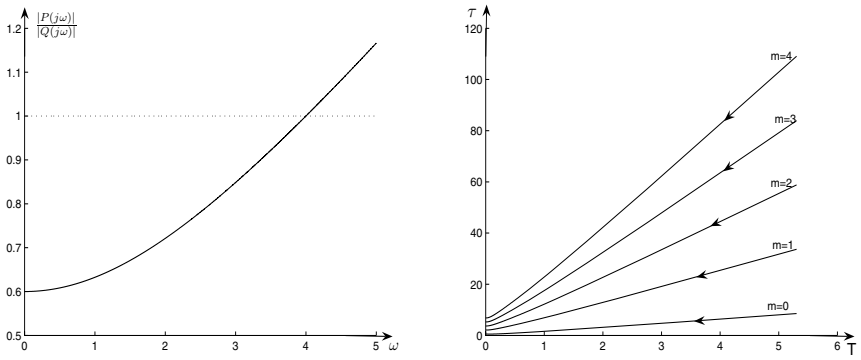


Figure 5. *Linearized Cushing equation example. Left:* $|P(j\omega)|/|Q(j\omega)|$ *against* ω. *Right:* τ_m, $m \in \{0, 1, 2, 3, 4\}$, *versus* T *when* $n = 1$.

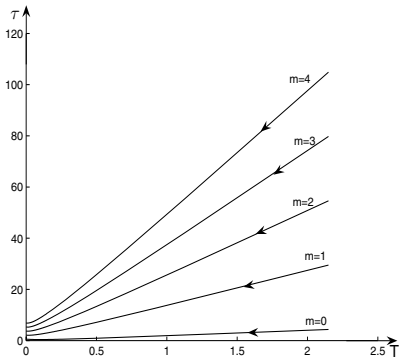


Figure 6. *Linearized Cushing equation example.* τ_m, $m \in \{0, 1, 2, 3, 4\}$, *versus* T *when* $n = 4$.

Also the slopes of the corresponding asymptotes are given by

$$\lim_{\omega \to 0} \frac{\tau}{T} = \frac{-n \arctan \left[\left(\frac{\beta^2}{\alpha^2} \right)^{1/n} - 1 \right]^{1/2} + \angle \left(\frac{-\beta}{\alpha} \right) + 2m\pi}{\left[\left(\frac{\beta^2}{\alpha^2} \right)^{1/n} - 1 \right]^{1/2}}.$$

Figures 5 (right) and 6 plot τ_m, $m \in \{0, 1, 2, 3, 4\}$, against T in the cases $n = 1$ and $n = 4$ for $\alpha = 3$ and $\beta = 5$, respectively. The crossing set $\Omega = (0, 4]$ is shown in Figure 5 (left). We observe that $\tau_{m+1}(\omega) > \tau_m(\omega)$ $\forall m \geq 0$ and $\omega \in \Omega$. Furthermore, it is easy to see that at $\omega = 2 \in \Omega$, for any m,

$$R_2 I_1 - R_1 I_2 = -13n \left(\frac{25}{13} \right)^{\frac{n-1}{n}} \left[\left(\frac{25}{13} \right)^{1/n} - 1 \right]^{1/2} < 0.$$

Therefore, we can conclude that as τ increases from zero, every time it crosses the stability crossing curve in Figure 5 (right) or Figure 6, (2.1) gains two additional right half plane

solutions. In addition, we can easily see that the system is stable when $T = 0$, $\tau = 0$. Therefore, the linearized Cushing equation is stable only in the region below the curve labeled "$m = 0$" and above the T-axis.

Example 4.2 (simplified helicopter model). Consider a helicopter model [18, 19] consisting of a fixed base and a rotary arm mounted on the base. The arm carries the helicopter body on one end and a counterweight on the other. The arm can make an elevation motion around an angle x. The corresponding nonlinear mathematical model is

$$J \cdot \ddot{x} = -g \cdot y \cdot (M + m) \cdot \sin x + 2 \cdot k_t \cdot r \cdot v(t), \tag{4.1}$$

where k_t and g represent the motor and the gravity constants, y is the distance between the rotation point and the rotary arm, r is the distance from the helicopter body to the fixed base, and m and M denote the mass of the helicopter blades (including the motors and the fixing devices) and the counterweight, respectively. J is the moment of inertia around the rotating point, and $v(t)$ is the corresponding voltage. We note that all of these values can be explicitly measured. Linearizing around the quiescent point, one gets

$$J \cdot \ddot{x} = -g \cdot y \cdot (M + m) \cdot x + 2 \cdot k_t \cdot r \cdot v, \tag{4.2}$$

and finally, after the damping factor identification, we obtain the helicopter transfer function

$$G(s) = \frac{0.2607}{s^2 + 0.07441s + 2.904}. \tag{4.3}$$

Considering a simple proportional-derivative (PD) controller (for improving the system response of the above helicopter laboratory experiment) $G_c(s) = (16.5s + 19.5)$ with a gamma-distribution with a gap $e^{-s\tau}/(1+sT)^n$ modeling the overall communication delay, one obtains the closed-loop characteristic equation given by

$$(s^2 + 0.07441s + 2.904)(1 + sT)^n + (4.3015s + 5.0836)\mathrm{e}^{-s\tau} = 0. \tag{4.4}$$

The crossing set Ω consists of one interval $(0, 5.0002]$ of type 31 (Figure 7 (left)).

Some stability crossing curves are plotted in Figure 7 (right).

Example 4.3 (controlling second-order systems). Consider (2.1) when $Q(s) = k_1 s + k_2$ and $P(s) = s^2 + 2$. It is easy to see that if $k_1 = 0$ and $T = 0$, the characteristic equation corresponds to the closed-loop system of a simple oscillator $1/(s^2 + 2)$ controlled by a delayed output feedback of the form $k_2 e^{-s\tau}$, that is,

$$\ddot{y}(t) + 2y(t) = u(t),$$

with

$$u(t) = -k_2 y(t - \tau).$$

It is important to point out that for very small delay values τ and very small gains k_2, the closed-loop system is asymptotically stable, but it is not asymptotically stable if the delay τ is equal to 0, that is, for the control law $u(t) = -k_2 y(t)$. We have the so-called *stabilizing effect* of the *delay* (see, for instance, [16] and the references therein on stabilizing oscillations by using delayed feedback laws).

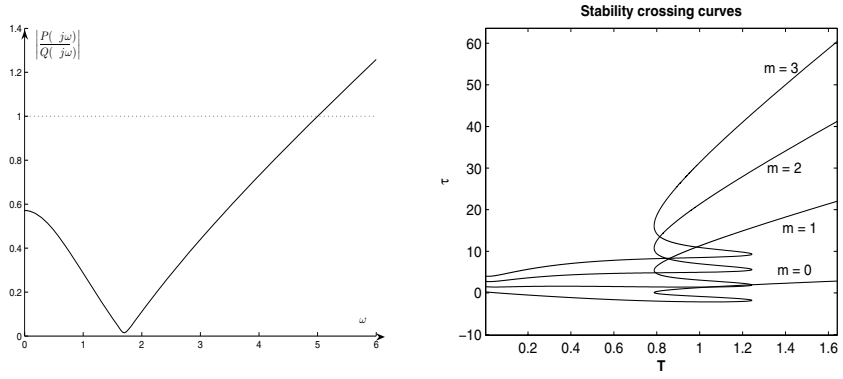


Figure 7. *Simplified helicopter model:* $\mathcal{T}_1^m$, $m = 0, 1, 2, 3$, *for* (4.4).

The more general system with the characteristic equation

$$(s^2 + 2)(1 + sT)^n + (k_1 s + k_2)\mathrm{e}^{-s\tau} = 0 \tag{4.5}$$

can be analyzed as follows.

Case 1: If $|k_2| < 2$ then the crossing set $\Omega = [\omega_+, \omega_-] \setminus \{\sqrt{2}\}$, where

$$\omega_\pm = \sqrt{\frac{k_1^2 + 4 \pm \sqrt{(k_1^2+4)^2 - 4(4 - k_2^2)}}{2}}.$$

We note that

$$\omega_- \le \sqrt{\frac{k_1^2 + 4 \pm \sqrt{(k_1^2+4)^2 - 16}}{2}} \le \sqrt{2} < \omega_+.$$

Therefore, Ω consists of two intervals of types 12 and 21, respectively. More details can be found below, where we will discuss the cases $k_2 = 0$ and $k_1 = 1$.

Case 2: If $|k_2| \ge 2$ then the crossing set $\Omega = (0, \omega_+] \setminus \{\sqrt{2}\}$, where $\omega_+ > \sqrt{2}$ is defined above. So, the crossing set Ω consists of two intervals of types 32 and 21, respectively.

Next we consider the following special case: $k_1 = 1$ and $k_2 = 0$. Using (3.5) and (3.6), we compute the crossing set $\Omega = \Omega_1 \cup \Omega_2$, where $\Omega_1 = [1, \sqrt{2})$ is of type 12 and $\Omega_2 = (\sqrt{2}, 2]$ is of type 21 (see also Figure 8 (left)). Simple computation shows that

$$T = \frac{1}{\omega}\sqrt{\left(\frac{\omega^2}{(2-\omega^2)^2}\right)^{1/n} - 1}$$

and

$$\tau_m = \frac{1}{\omega}\left(\angle\frac{j\omega}{(2-\omega^2)(1+j\omega T)^n} + 2m\pi\right).$$

According to the result of Proposition 3.2 we have $\lim_{\omega\to 1} T = 0$, $\lim_{\omega\to 2} T = 0$, $\lim_{\omega\to\sqrt{2}} T =$

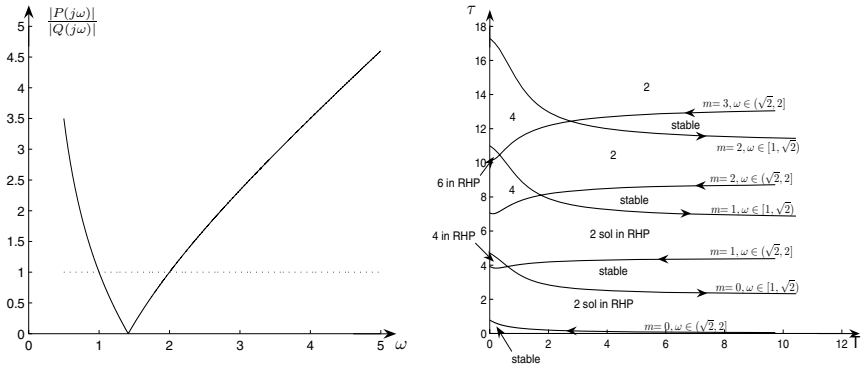


Figure 8. *Second-order system Example 4.3: Left:* $|P(j\omega)|/|Q(j\omega)|$ *versus* ω. *Right:* τ_m, $m = 0, 1, 2, 3$, *versus* T *when* $n = 1$.

∞, $\lim_{\omega\to 1} \tau_m = -\frac{\pi}{2} + 2m\pi$, $\lim_{\omega\to 2} \tau_m = \frac{\pi}{4} + m\pi$, and

$$\lim_{\omega\to\sqrt{2}-0} \tau_m = \frac{[2m - (n-1)/2]\pi}{\sqrt{2}},$$
$$\lim_{\omega\to\sqrt{2}+0} \tau_m = \frac{[2m - (n+1)/2]\pi}{\sqrt{2}}.$$

We will now calculate the direction of crossing. A direction calculation yields

$$\begin{aligned} R_2 I_1 - R_1 I_2 &= n(1+\omega^2 T^2)^{n-1}(2-\omega^2)^2 \operatorname{Im}(1 - j\omega T) \\ &= -n\omega T(1+\omega^2 T^2)^{n-1}(2-\omega^2)^2 < 0. \end{aligned}$$

Therefore, using Proposition 3.3, we can conclude that as we cross the stability crossing curve from its right-hand side to its left-hand side, a pair of complex conjugate solutions of $D = 0$ cross the imaginary axis from the right half complex plane to the left half plane.

The computations above show us that the following inequalities hold:

$$\tau_m(\sqrt{2}+0) < \tau_m(\sqrt{2}-0) < \tau_{m+1}(\sqrt{2}+0) \quad \forall m \in \mathbb{Z}.$$

This simply states that for large values of T the crossing toward stability and the crossing toward instability interlace. Considering the additional fact that the system is obviously stable for $\tau = 0$ and $T = 0$, and the fact that the stability crossing curve approaches horizontal, we can conclude that the system has an *infinite number* of stable regions. Figure 8 shows the case when $n = 1$.

Example 4.4 (traffic flow model). Finally we consider a *time delay microscopic system* including delayed reactions of the driver, and, as explained in the introduction, we will use a *distributed delay with a gap* for modeling human driver reactions with respect to the traffic behavior. Specifically, considering the traffic flow dynamics described by (1.8) with $\alpha_1 = \alpha_2 = 2$, the stability analysis leads to the conclusion that the system has only one stability region (see Figure 9).

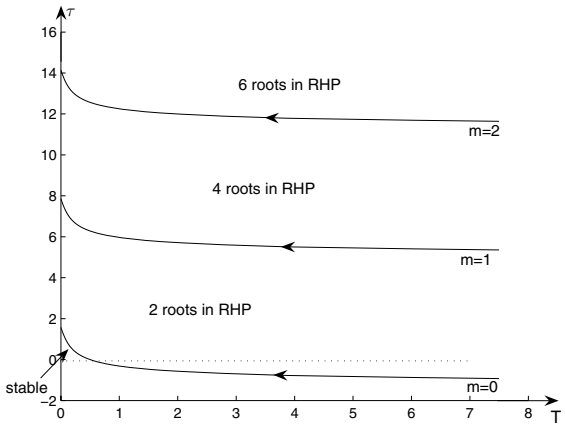


Figure 9. *Traffic flow model with two vehicles:* τ_m, $m \in \{0, 1, 2\}$, *versus* T *when* $n = 2$.

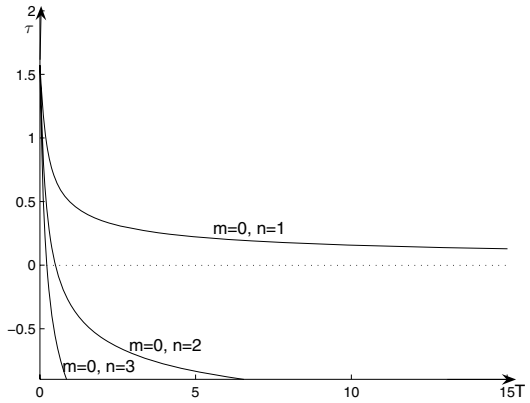


Figure 10. *Traffic flow model:* τ_1 *versus* T *when* $n \in \{1, 2, 3\}$.

More exactly, the crossing set Ω consists of one interval $(0, 4]$, and the crossing curves are described by the following equations:

$$T = \frac{1}{\omega}\sqrt{\left(\frac{16}{\omega^2}\right)^{1/n} - 1}, \tag{4.6}$$

$$\tau = \frac{1}{\omega}\left(\frac{\pi}{2} - n\arctan(\omega T) + 2m\pi\right), \qquad m = 0, 1, 2, \ldots. \tag{4.7}$$

We note that Assumption V is not satisfied in this case, and the shape of the crossing curves do not perfectly match the classification proposed in section 3. However, the ideas of our approach still apply.

It is also important to point out that the matrix $A_1 + A_2$ that defines the characteristic equation (1.9) always has an eigenvalue at the origin. This corresponds to the situation in

which the relative movement of one vehicle to the others is zero (the vehicles are either staying or moving with the same velocity).

On the other hand, varying n over positive integers, we can see that the stability region becomes *smaller* as the integer n increases (Figure 10).

5. Concluding remarks. This paper addressed the stability problem of shifted gamma-distributed delay systems. More specifically, we have characterized the geometry of the stability crossing curves in the parameter space defined by the gap and the corresponding mean delay. Several illustrative examples complete the presentation.

REFERENCES

[1] S. P. Blythe, R. M. Nisbet, and W. S. C. Gurney, *Stability switches in distributed delay models*, J. Math. Anal. Appl., 109 (1985), pp. 388–396.

[2] F. G. Boese, *The stability chart for the linearized Cushing equation with a discrete delay and with gamma-distributed delays*, J. Math. Anal. Appl., 140 (1989), pp. 510–536.

[3] R. E. Chandler, R. Herman, and E. W. Montroll, *Traffic dynamics: Analysis of stability in car following*, Oper. Res., 7 (1958), pp. 165–184.

[4] K. L. Cooke and Z. Grossman, *Discrete delay, distributed delay and stability switches*, J. Math. Anal. Appl., 86 (1982), pp. 592–627.

[5] J. M. Cushing, *Volterra integrodifferential equations in population dynamics*, in Mathematics of Biology, M. Iannalli, ed., Ligouri Editore, Naples, 1981, pp. 81–148.

[6] M. Farkas, *Periodic Motions*, Springer-Verlag, New York, 1994.

[7] A. Farkas, M. Farkas, and G. Szabó, *Multiparameter bifurcation diagrams in predator-prey models with time lag*, J. Math. Biol., 26 (1988), pp. 93–103.

[8] K. Gu, V. L. Kharitonov, and J. Chen, *Stability and Robust Stability of Time-Delay Systems*, Birkhäuser Boston, Boston, 2003.

[9] K. Gu, S.-I. Niculescu, and J. Chen, *On stability of crossing curves for general systems with two delays*, J. Math. Anal. Appl., 311 (2005), pp. 231–253.

[10] J. K. Hale, E. F. Infante, and F. S.-P. Tsen, *Stability in linear delay equations*, J. Math. Anal. Appl., 105 (1985), pp. 533–555.

[11] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Appl. Math. Sci. 99, Springer-Verlag, New York, 1993.

[12] D. Helbing, *Traffic and related self-driven many-particle systems*, Rev. Modern Phys., 73 (2001), pp. 1067–1141.

[13] T. Insperger and G. Stépán, *Comparison of the Fargue-type analytical approximation and the numerical semi-discretization method for delayed systems*, J. Comput. Appl. Mech., 4 (2003), pp. 203–215.

[14] Y. Kuang, *Delay Differential Equations with Applications in Population Dynamics*, Academic Press, Boston, 1993.

[15] N. MacDonald, *Biological Delay Systems: Linear Stability Theory*, Cambridge University Press, Cambridge, UK, 1989.

[16] S.-I. Niculescu, *Delay Effects on Stability. A Robust Control Approach*, Lecture Notes in Control and Inform. Sci. 269, Springer-Verlag, London, 2001.

[17] R. M. Nisbet and W. S. C. Gurney, *The formulation of age-structure models*, in Mathematical Ecology, T. G. Hallam and S. A. Levin, eds., Springer-Verlag, Berlin, 1986, pp. 95–115.

[18] Quanser, *3D Helicopter Experiment with DAQ Board and WinCon Software*, 2004, http://www.quanser.com.

[19] O. Roesch, H. Roth, and S.-I. Niculescu, *Remote control of mechatronic systems over communication networks*, in Proceedings of the IEEE International Conference on Mechatronics and Automation, Niagara Falls, Canada, 2005.

[20] R. Sipahi and N. Olgac, *Complete robustness of third-order LTI multiple time-delay systems*, Automatica J. IFAC, 41 (2005), pp. 1413–1422.

[21] R. Sipahi and S.-I. Niculescu, *A survey of deterministic time-delayed traffic flow models and arising problems due to delays*, in Proceedings of the 7th IFAC Workshop on Time-Delay Systems, Nantes, France, 2007 (to appear).
[22] R. Sipahi, S.-I. Niculescu, and F. M. Atay, *Effects of short-term memory of drivers on stability interpretations of traffic flow dynamics*, in Proceedings of the 2007 American Control Conference, New York, NY, 2007 (to appear).
[23] G. Stépán, *Retarded Dynamical Systems: Stability and Characteristic Function*, Wiley, New York, 1989.
[24] G. Stépán, *Delay-differential equation models for machine tool chatter*, in Dynamics and Chaos in Manufacturing Process, F. C. Moon, ed., Wiley, New York, 1998, pp. 165–192.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 2, pp. 494–547

Period-Doubling of Spiral Waves and Defects*

Björn Sandstede† and Arnd Scheel‡

Abstract. Motivated by experimental observations in the light-sensitive Belousov–Zhabotinsky reaction and subsequent numerical works, we discuss period-doubling bifurcations of spiral waves and other coherent structures. We report on explanations of the observed phenomena which involve a detailed analysis of spectra, and of the associated eigenfunctions, of defects on bounded and unbounded domains.

Key words. period-doubling, spiral waves, wave trains

AMS subject classifications. 37L10, 35K57, 34C37

DOI. 10.1137/060668158

1. Introduction.

Spiral waves arise in many biological, chemical, and physical systems. They rotate rigidly as functions of time, and a typical spatial profile of a planar spiral wave is shown in Figure 1(i). The importance of spiral waves is partly due to the fact that experimentally observed patterns are often organized by interacting spirals. Upon varying system parameters, spiral waves may destabilize, and the resulting instabilities lead often to more complex coherent patterns or to spatio-temporally disorganized dynamics. Examples of experimentally observed instabilities are meander instabilities [17, 20, 24], core [49] and far-field breakup [26], and period-doubling instabilities [27, 28, 47].

From a classical dynamical-systems viewpoint, we expect that the transition to complicated dynamics is initiated by a sequence of generic local or global bifurcations—saddle-node and Hopf bifurcations in the case of equilibria, and saddle-node, Hopf, and period-doubling bifurcations in the case of periodic orbits. Indeed, chemical reactions can be modeled by reaction-diffusion systems in bounded domains for which bifurcations can be reduced to finite-dimensional center manifolds, and where instabilities are therefore expected to be of the aforementioned type.

In a first attempt to understand spiral-wave instabilities, we can view spirals as time-periodic solutions while disregarding their spatial structure: note that spirals rotate rigidly as functions of time and that their wave pattern repeats itself after one period of rotation; see Figure 1(i). Thus, from this viewpoint, we expect to see Hopf and period-doubling bifurcations as typical precursors on the route to complicated spatio-temporal dynamics. Hopf

*Received by the editors August 24, 2006; accepted for publication (in revised form) by T. Kaper March 30, 2007; published electronically June 29, 2007.
http://www.siam.org/journals/siads/6-2/66815.html

†Department of Mathematics, University of Surrey, Guildford, GU2 7XH, UK (b.sandstede@surrey.ac.uk). This author was partially supported by a Royal Society–Wolfson Research Merit Award and by the NSF through grant DMS-0203854.

‡Department of Mathematics, University of Minnesota, Minneapolis, MN 55455 (scheel@ima.umn.edu). This author was partially supported by the NSF through grant DMS-0203301.

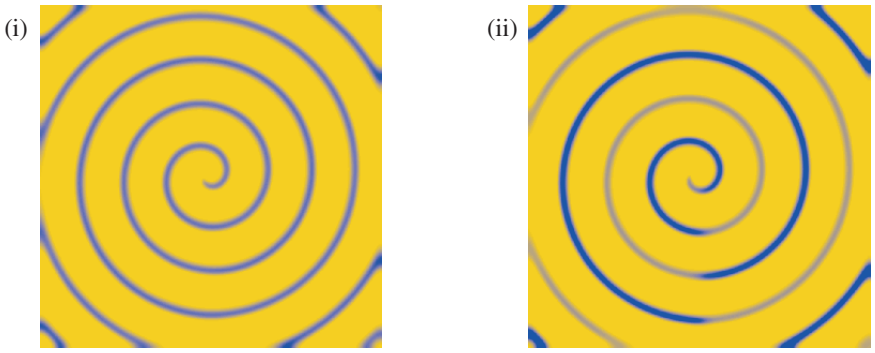


Figure 1. *Snapshots of the w-component of two-dimensional (2D) spiral waves in the Rössler equation* (1.3) *are shown for two different parameter values:* (i) *shows a rigidly-rotating spiral wave for* $C = 2.95$ *in the accompanying animation* (66815_01.mpg [542KB])*, while* (ii) *shows the spiral wave for* $C = 3.4$ *after a period-doubling bifurcation in the accompanying animation* (66815_02.mpg [653KB])*. The period-doubled spiral exhibits a line defect, which emanates from the core and ends at the bottom of the boundary, in order to accommodate the phase-shifted wave trains to either side.*

bifurcations have indeed been observed and give rise to meander instabilities [2]. What appears to be chaotic hypermeander of spiral tips has also been observed for parameter values far beyond the meandering transition, but, to our knowledge, the question of whether complicated hypermeander is actually caused by subsequent secondary bifurcations has not yet been settled.

More recently, a different type of instability has been observed both in experiments [27, 28, 47] and in numerical simulations [14]. The primary spiral destabilizes as illustrated in Figure 1(ii) and gives rise to a new spiral wave that emits wave trains with doubled wavelength and temporal period. An additional feature of this transition is the occurrence of a line defect that emerges from the spiral core and which accommodates the necessary mismatch of the phases of the oscillations that are emitted by the spiral core. The pattern still repeats itself but only after two rotations of the spiral core. Thus, the spiral wave, regarded as a time-periodic solution, has undergone a period-doubling bifurcation, and, in accordance with the literature, we will refer to this instability as *period-doubling* of spiral waves.

This apparently straightforward explanation is, however, too simple. Consider, for example, the spiral wave as a solution in a circular domain. We may then pass to a rigidly corotating coordinate frame in which the spiral wave becomes an equilibrium. In particular, we can obtain the Floquet multipliers of the spiral wave in the original laboratory frame simply by exponentiating the eigenvalues of the spiral wave in the corotating frame where the spiral is an equilibrium. An application of the spectral mapping theorem then shows that a simple eigenvalue $\rho = -1$ cannot occur for the exponential of the real linearization in the corotating frame. In other words, in rotationally symmetric domains, spiral waves are equilibria when considered in a corotating frame, which generically undergo only saddle-node or Hopf bifurcations. From this perspective, period-doubling ceases to be meaningful as an instability mechanism.

Thus, the only conceivable explanation left is that the instability is a Hopf bifurcation. Since the temporal period of the bifurcating patterns observed in experiments and numerical

simulations is approximately twice the period of the primary spiral, the Hopf frequency ω_{H} needs to be in a 2:1 resonance with the rotation frequency ω_* of the spiral wave so that

$$\omega_{\mathrm{H}} = \frac{\omega_*}{2}.$$

From a genericity viewpoint, it is therefore important to understand which mechanism enforces this 2:1 resonance of the Hopf frequency of eigenvalues and the rotation frequency of the spiral wave. This natural question is indeed the central point of this paper.

The seemingly artificial choice of the corotating frame can be put in a slightly more systematic context once the symmetries of the problem are taken into account. Posing the underlying reaction-diffusion system on a circular domain, rotations in SO(2) of the domain act on patterns and map solutions of the system to solutions. Spiral waves are relative equilibria with respect to this group action; that is, their time evolution is equivalent to the action of the group: spirals are rigidly rotating. The isotropy of the spiral waves that we consider is trivial: only a full rotation by 2π maps the spiral profile onto itself. Consequently, center manifolds near spiral waves are principal fiber bundles, given as a direct product of the underlying symmetry group SO(2) and a complement of the tangent space of the group orbit in the center eigenspace [43]. In particular, the center manifold is a globally trivial bundle, which provides yet another reason for why the case of a simple negative Floquet multiplier -1 is precluded for period-doubling bifurcations of spiral waves as this scenario requires the center manifold to be nonorientable. The structure of the principal fiber bundle can be understood by first constructing a center manifold in a Poincaré section, which is also a section to the group orbit, and then transporting the center manifold along the periodic orbit using the group action.

Symmetry is also the key to understanding the meandering patterns that arise at Hopf bifurcations. As first pointed out by Barkley [2], the meandering motion of spiral waves can be understood if we consider the spiral wave on the entire plane where, in addition to rotations, spatial translation of patterns maps solutions to solutions. The full symmetry group is therefore the special Euclidean group SE(2) of translations and rotations in the plane. Center manifolds near relative equilibria can still be described as principal fiber bundles $\mathrm{SE}(2)\times V$, where V corresponds to the Hopf eigenmodes. The dynamics on the center manifold is of skew-product form: After an appropriate reparametrization of time, the dynamics near onset is governed [7, 12, 43] by the ordinary differential equations (ODEs)

$$\begin{aligned}
\dot a &= \mathrm{e}^{\mathrm{i}\varphi}[v + \mathrm{O}(|v|^2)],\\
\dot\varphi &= \omega_*,\\
\dot v &= [\mu + \mathrm{i}\omega_{\mathrm{H}}]v - (1 - \mathrm{i}\alpha)|v|^2 v,
\end{aligned}$$

where φ denotes the phase of the spiral, that is, its angle of rotation relative to a fixed reference frame, $a = x + \mathrm{i}y$ is its position, and $v \in \mathbb{C}$ parametrizes a neighborhood of the origin in the Hopf eigenspace. Substituting the periodic orbit $v_*(t)$ with frequency $\omega_{\mathrm{H}} + \mu\alpha$ of the v-equation and the solution $\varphi_*(t) = \omega_* t$ of the φ-equation into the equation for a, and expanding the resulting equation in Fourier modes, we find that the solution $a_*(t)$ is given by

$$a_*(t) = \sum_{k=-\infty}^{\infty} a_k \frac{\mathrm{e}^{\mathrm{i}[\omega_* - k(\omega_{\mathrm{H}}+\mu\alpha)]t} - 1}{\omega_* - k(\omega_{\mathrm{H}} + \mu\alpha)} \tag{1.1}$$

so that the spiral tip stays bounded unless ω_* and ω_{H} are resonant. When

$$\omega_* = \ell\omega_{\mathrm{H}} \quad \text{for some} \quad \ell \in \mathbb{Z}, \tag{1.2}$$

then the tip position $a_*(t) = a_\ell t + \cdots$ is unbounded near $\mu = 0$, and the spiral wave drifts with velocity a_ℓ. The resonance $\ell = 1$ has been observed frequently in experiments [2, 46].

Returning to the period-doubling instability of spiral waves, we have already inferred that period-doubling ought to be a resonant Hopf bifurcation with $\ell = 2$ in (1.2). The drift predicted by (1.1) had not been observed originally in the experiments [47] or the numerical simulations [14]. Based on the theoretical predictions outlined above, we observed drift in the Rössler system

$$\begin{aligned} u_t &= 0.4\,\Delta u - v - w, \\ v_t &= 0.4\,\Delta v + u + 0.2\,v, \\ w_t &= 0.4\,\Delta w + uw - Cw + 0.2, \end{aligned} \tag{1.3}$$

upon varying C (and we report on these results in section 6 below). Independently, drift was also observed in [4] for the system (1.3).

In summary, the supposition of an exact 2:1 resonance of the Hopf frequency of eigenvalues and the rotation frequency of the spiral wave leads to the prediction that period-doubled spirals should drift, which was, in turn, verified in numerical simulations. Thus, the remaining key question is what enforces this resonance which seems to be nongeneric and should not occur in one-parameter systems such as (1.3).

At this point, it is time to emphasize that the reduction results for planar patterns in the presence of the noncompact Euclidean group hold only for localized rotating waves. For Archimedean spirals, the presence of essential spectrum on the imaginary axis prevents a reduction to a finite-dimensional system. While this issue may appear to be of a purely technical nature for meandering instabilities where theoretical predictions are in excellent agreement with experimental and numerical results, the situation is different for period-doubling instabilities. We shall argue that period-doubling of spirals is not caused by isolated point spectrum but instead by the essential spectrum of the asymptotic wave trains.

Specifically, we shall show that ordinary period-doubling bifurcations of wave trains in the traveling-wave equation create 2:1 resonances of the essential spectra of planar spiral waves that are spatially asymptotic to these wave trains in their far field. More precisely, period-doubling instabilities of wave trains manifest themselves for planar spiral waves in the form of curves of essential spectrum that cross the imaginary axis first at exactly $\Lambda = \pm \mathrm{i}\omega_*/2$. In particular, robust 2:1 resonances can occur in an open set of one-parameter families of reaction-diffusion systems. When posed on physically relevant bounded domains such as disks of radius R, we had shown previously in [35, 41, 42] that spiral spectra accumulate in the limit $R \to \infty$ onto the so-called absolute spectrum. We show here that absolute spectra of wave trains near period-doubling bifurcations are symmetric with respect to reflections across $\operatorname{Im}\Lambda = \mathrm{i}\omega_*/2$: One generic possibility is therefore that the absolute spectrum lies entirely on the line $\operatorname{Im}\Lambda = \mathrm{i}\omega_*/2$, leading again to a 2:1 resonance. The latter case occurs, in fact, for wave trains with small wave numbers near spatially homogeneous oscillations. Last, we shall also investigate the nature of the line defect apparent in Figure 1(ii).

Plan of the paper.
Section 2: Contains the analysis of spatio-temporal period-doubling of one-dimensional (1D) wave trains on which the remainder of this paper relies.
Section 3: Spatio-temporal period-doubling for 1D sources *(can be skipped at first reading)*.
Section 4: Extends the spectral results in section 2 from wave trains to planar Archimedean spiral waves and explains the 2:1 resonance of Hopf eigenvalues.
Section 5: Spatial dynamics is used to analyze line defects and boundary layers in period-doubling bifurcations near spatially homogeneous oscillations.
Section 6: Applies our results to the Rössler system (1.3) in which period-doubling had been observed previously.
Section 7: Discussion of results and their limitations.

2. Spatio-temporal period-doubling of wave trains. Our interest in this section is to study period-doubling bifurcations of wave trains and how these manifest themselves on the spectral level in different coordinate frames. The results obtained here are crucial for our analysis of period-doubling of spiral waves which we will undertake in section 4.

We consider reaction-diffusion systems

$$u_t = Du_{xx} + f(u;\mu), \qquad x \in \mathbb{R}, \tag{2.1}$$

for $u \in \mathbb{R}^n$ and $\mu \in \mathbb{R}$, where D is a positive diagonal matrix and the nonlinearity $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is smooth. We assume that (2.1) with $\mu = 0$ has a wave-train solution $u_{\rm wt}(kx - \omega t)$ for an appropriate wave number k and temporal frequency ω, where we assume that $u_{\rm wt}$ is 2π-periodic in its argument so that $u_{\rm wt}(\xi) = u_{\rm wt}(\xi + 2\pi)$ for all ξ.

2.1. Spatial and temporal period-doubling. If the wave number k vanishes, then $u(x,t) = u_{\rm wt}(-\omega t)$ is a spatially homogeneous oscillation which satisfies the ODE

$$u_t = f(u;\mu). \tag{2.2}$$

Period-doubling of $u_{\rm wt}(-\omega t)$ occurs when $\rho = -1$ is a temporal Floquet multiplier of the linearization of the period map associated with (2.2) about $u_{\rm wt}$. The multiplier $\rho = -1$ is generically simple, and the resulting purely temporal period-doubling leads to a spatially homogeneous oscillation with frequency close to $\omega/2$.

Next, assume that $k \neq 0$. In this case, we can pass from the laboratory frame x to the comoving frame $\xi = kx - \omega t$ in which (2.1) becomes

$$u_t = k^2 Du_{\xi\xi} + \omega u_\xi + f(u;\mu). \tag{2.3}$$

Note that $u_{\rm wt}(\xi)$ is an equilibrium solution of (2.3) with spatial period 2π, and we focus here on steady-state bifurcations of (2.3) which are captured by the traveling-wave ODE

$$k^2 Du_{\xi\xi} + \omega u_\xi + f(u;\mu) = 0. \tag{2.4}$$

Spatial period-doubling of the 2π-periodic orbit $u_{\rm wt}(\xi)$ of (2.4) occurs when

$$k^2 Dv_{\xi\xi} + \omega v_\xi + f_u(u_{\rm wt}(\xi);\mu)v = 0 \tag{2.5}$$

has a nonzero solution $v_{\mathrm{pd}}(\xi)$ with $v_{\mathrm{pd}}(\xi+2\pi) = -v_{\mathrm{pd}}(\xi)$ for all ξ, corresponding to a simple spatial Floquet multiplier at -1. This bifurcation corresponds to a generic pitchfork bifurcation of (2.3) when we pose it on the spatial interval $(0, 4\pi)$ with periodic boundary conditions. The $\mathbb{Z}_2$-symmetry that turns the steady-state bifurcation into a pitchfork is generated by the shift $\xi \mapsto \xi + 2\pi$ which also generates the isotropy group of the equilibrium u_{wt} of spatial period 2π when considered on the interval $(0, 4\pi)$. Lyapunov–Schmidt reduction for the nonlinear problem (2.4) on an appropriate function space of 4π-periodic function leads to a family of spatially period-doubled equilibria that bifurcate from u_{wt}. Center-manifold reduction, or Lyapunov–Schmidt reduction [19], shows that the principle of exchange of stability holds for the temporal dynamics of (2.3) on the space of 4π-periodic functions provided the cubic coefficient of the reduced equation is nonzero. In other words, the bifurcating pattern is stable as a solution to (2.3) if it exists for parameter values for which the primary pattern u_{wt} is unstable. We refer the reader to [8] for a discussion of the multiplicity of period-doubling eigenvalues using Evans functions.

Last, we interpret these results in the laboratory frame. Assuming that $k \neq 0$ and $\omega \neq 0$, we consider (2.1) on the interval $(0, 4\pi/k)$ with periodic boundary conditions. Equation (2.1) generates a compact semiflow Φ_t on $H^2_{\mathrm{per}}(0, \frac{4\pi}{k})$, and the wave train u_{wt} corresponds to a time-periodic solution with period $T = 2\pi/\omega$. We refer to eigenvalues ρ of the linearized period map $\Phi'_T(u_{\mathrm{wt}})$ as Floquet multipliers, which turn out to be conveniently related to the spectrum of the linearization

$$\lambda v = Dk^2 v_{\xi\xi} + \omega v_\xi + f_u(u_{\mathrm{wt}}(\xi); 0)v \tag{2.6}$$

of (2.3) with 4π-periodic boundary conditions about the equilibrium $u_{\mathrm{wt}}(\xi)$. Indeed, any nontrivial solution $v(\xi)$ to the eigenvalue problem (2.6) gives a solution $w(x, \cdot)$ of the eigenvalue problem for the period map of (2.1) in the laboratory frame via

$$w(x,t) = \mathrm{e}^{\lambda t} v(kx - \omega t), \qquad w(x,T) = \mathrm{e}^{\lambda T} v(kx - 2\pi)$$

and vice versa. Spatial period-doubling of (2.6) corresponds to $\lambda = 0$ and $v(\xi)$ with $v(\xi+2\pi) = -v(\xi)$ for all ξ. The resulting solution $w(x,t)$ satisfies $w(x,T) = -w(x,0)$ and therefore gives a simple Floquet multiplier $\rho = -1$. We refer to the occurrence of a simple Floquet multiplier $\rho = -1$ of $\Phi'_T(u_{\mathrm{wt}})$ as spatio-temporal period-doubling.

2.2. Essential spectra of wave trains. More generally, we can consider the linearization on the real line $x \in \mathbb{R}$. First, consider the linearization

$$v_t = Dk^2 v_{\xi\xi} + \omega v_\xi + f_u(u_{\mathrm{wt}}(\xi); 0)v, \qquad \xi \in \mathbb{R}, \tag{2.7}$$

in the comoving frame together with the associated eigenvalue problem

$$\lambda v = Dk^2 v_{\xi\xi} + \omega v_\xi + f_u(u_{\mathrm{wt}}(\xi); 0)v, \qquad \xi \in \mathbb{R}. \tag{2.8}$$

We write this equation as the first-order system

$$\mathbf{v}_x = \begin{pmatrix} 0 & 1 \\ k^{-2}D^{-1}[\lambda - f_u(u_{\mathrm{wt}}(\xi); 0)] & \omega k^{-2} D^{-1} \end{pmatrix} \mathbf{v} \tag{2.9}$$

and denote the associated 2π-period map by $\Psi_{2\pi}(\lambda)$. Spatial Floquet exponents ν/k of (2.8) or (2.9) are determined as roots of the Wronskian

$$d(\lambda,\nu) := \det\left[\mathrm{e}^{2\pi\nu/k} - \Psi_{2\pi}(\lambda)\right].$$

The Wronskian $d(\lambda,\nu)$ satisfies

$$\begin{array}{rcll} \overline{d(\lambda,\nu)} & = & d(\bar\lambda,\bar\nu) & \text{(complex conjugation)},\\ d(\lambda,\nu) & = & d(\lambda,\nu+\mathrm{i}k\ell) & \text{(artificial Floquet conjugation)} \end{array} \tag{2.10}$$

for all integers ℓ. Spatial Floquet exponents can also be found by seeking nontrivial solutions to (2.8) of the form

$$v(\xi) = \mathrm{e}^{\nu\xi/k}v_0(\xi), \qquad v_0(\xi+2\pi) = v_0(\xi) \quad \forall\xi, \tag{2.11}$$

where v_0 is a 2π-periodic solution of

$$\lambda v = D(k\partial_\xi+\nu)^2 v + \frac{\omega}{k}(k\partial_\xi+\nu)v + f_u(u_{\mathrm{wt}}(\xi);0)v. \tag{2.12}$$

Purely imaginary spatial Floquet exponents $\nu \in \mathrm{i}\mathbb{R}$ give eigenvalues λ of (2.8), and each eigenfunction (2.11) leads to a solution

$$v(\xi,t) = \mathrm{e}^{\lambda t}\mathrm{e}^{\nu\xi/k}v_0(\xi)$$

of (2.7). We record that spatial period-doubling as discussed in section 2.1 is equivalent to having a nontrivial solution v of (2.12) for $\lambda = 0$ and $\nu = \mathrm{i}k/2$.

In the laboratory frame, the relevant linearization is

$$u_t = Du_{xx} + f_u(u_{\mathrm{wt}}(kx-\omega t);0)u, \qquad x\in\mathbb{R}. \tag{2.13}$$

Temporal Floquet multipliers ρ and the associated Floquet exponents Λ in the laboratory frame are determined by bounded nontrivial solutions $u(x,t)$ of (2.13) with

$$u(x,T) = \rho u(x,0) = \mathrm{e}^{\Lambda T}u(x,0),$$

where $T = 2\pi/\omega$. It turns out that Λ is a temporal Floquet exponent if and only if there is a nontrivial solution of (2.13) of the form

$$u(x,t) = \mathrm{e}^{\Lambda t}\mathrm{e}^{\nu x}u_0(kx-\omega t)$$

with $\nu \in \mathrm{i}\mathbb{R}$, where u_0 is 2π-periodic in its argument. Solutions of this form for arbitrary $\nu\in\mathbb{C}$ are in one-to-one correspondence with the solutions (2.11) of (2.8) via

$$u(x,t) = \mathrm{e}^{\lambda t}\mathrm{e}^{\nu\xi/k}v_0(\xi) = \mathrm{e}^{[\lambda-\nu\omega/k]t}\mathrm{e}^{\nu x}v_0(kx-\omega t) = \mathrm{e}^{\Lambda t}\mathrm{e}^{\nu x}v_0(kx-\omega t)$$

with

$$\Lambda = \lambda - \frac{\omega\nu}{k} = \lambda - c_{\mathrm{p}}\nu, \tag{2.14}$$

where $c_\mathrm{p} = \omega/k$ is the phase speed of the wave train u_wt. Thus, the temporal Floquet exponents Λ in the laboratory frame are roots of

$$\mathcal{D}(\Lambda, \nu) := d(\Lambda + c_\mathrm{p}\nu, \nu).$$

Using (2.10), we see that $\mathcal{D}$ satisfies

$$\begin{array}{rcll} \overline{\mathcal{D}(\Lambda, \nu)} & = & \mathcal{D}(\bar{\Lambda}, \bar{\nu}) & \text{(complex conjugation)}, \\ \mathcal{D}(\Lambda, \nu) & = & \mathcal{D}(\Lambda - \mathrm{i}\omega\ell, \nu + \mathrm{i}k\ell) & \text{(Floquet conjugation)} \end{array} \tag{2.15}$$

for all integers ℓ. Typically, solutions of $d(\lambda, \nu) = 0$ come in curves $\lambda = \lambda_*(\nu)$, yielding also $\Lambda = \Lambda_*(\nu)$. For $\nu \in \mathrm{i}\mathbb{R}$, we refer to these curves as dispersion curves in the comoving and the laboratory frames, respectively. We say that a dispersion curve $\Lambda_*(\nu)$ is simple if

$$\partial_\Lambda \mathcal{D}(\Lambda, \nu) = \partial_\lambda d(\lambda, \nu) \neq 0$$

at $\Lambda = \Lambda_*(\nu)$ or $\lambda = \lambda_*(\nu)$. The derivative

$$c_\mathrm{g} := -\frac{\mathrm{d}\,\mathrm{Im}\,\Lambda}{\mathrm{d}\,\mathrm{Im}\,\nu}$$

is commonly referred to as the *group velocity* in the laboratory frame. The relation (2.14) can therefore be viewed as transforming the group velocity from the laboratory to the comoving frame by subtracting the speed of the frame.

Equation (2.14) implies that spatial period-doubling with $\lambda = 0$ and $\nu = \mathrm{i}k/2$ in the comoving frame becomes spatio-temporal period-doubling with $\Lambda = -\mathrm{i}\omega/2$ and $\nu = \mathrm{i}k/2$ in the laboratory frame. The observation that the composition of the two symmetries in (2.15) fixes $\mathrm{Im}\,\Lambda = -\mathrm{i}\omega/2$ leads us to the following lemma on robustness of period-doubling.

Lemma 2.1 (robustness of spatio-temporal period-doubling). *Assume that there is a simple dispersion curve $\Lambda(\nu)$ with*

$$\mathrm{Im}\,\Lambda(\mathrm{i}k/2) = -\frac{\mathrm{i}\omega}{2}; \tag{2.16}$$

then the dispersion curve is reflection symmetric about the line $\mathrm{Im}\,\Lambda = -\mathrm{i}\omega/2$ for ν close to $\mathrm{i}k/2$. Moreover, (2.16) is robust under sufficiently small perturbations of the parameter value μ and the coefficients $u_\mathrm{wt}(\xi)$ in (2.6).

Proof. From (2.15), we conclude that $\mathcal{D}(\Lambda, \nu) = 0$ if and only if $\mathcal{D}(\bar{\Lambda} - \mathrm{i}\omega, \bar{\nu} + \mathrm{i}k) = 0$. Upon substituting $\Lambda = -\mathrm{i}\omega/2 + l$ and $\nu = \mathrm{i}k/2 + \mathrm{i}\gamma$ with $\gamma \in \mathbb{R}$ into these identities, we see that $\mathcal{D}(-\mathrm{i}\omega/2 + l,\ \mathrm{i}k/2 + \mathrm{i}\gamma) = 0$ if and only if $\mathcal{D}(-\mathrm{i}\omega/2 + \bar{l},\ \mathrm{i}k/2 - \mathrm{i}\gamma) = 0$. Applying the implicit function theorem to both equations, and using uniqueness of solutions, we conclude that $l(-\gamma) = \bar{l}(\gamma)$ for all γ close to zero, which implies the asserted symmetry of the dispersion curve about the line $\mathrm{Im}\,\Lambda = -\mathrm{i}\omega/2$. Robustness with respect to parameter variations is again a consequence of the implicit function theorem. ∎

In preparation for the analysis in the following two sections, we examine the linearization in exponentially weighted spaces

$$L^2_\eta := \{u \in L^2_\mathrm{loc};\ |u|_{L^2_\eta} < \infty\}, \qquad |u|^2_{L^2_\eta} := \int_{\mathbb{R}} |u(x)|^2 \mathrm{e}^{-2\eta x}\,\mathrm{d}x. \tag{2.17}$$

The spectra in L^2_η can be computed in the same way as for $\eta = 0$ by solving (2.12) with $\nu \in \eta + \mathrm{i}\mathbb{R}$, which yields an η-dependent family of dispersion curves $\Lambda(\nu)$ with $\operatorname{Re}\nu = \eta$. The real part of these curves depends on η according to

$$\frac{\mathrm{d}\operatorname{Re}\Lambda}{\mathrm{d}\eta} = \frac{\mathrm{d}\operatorname{Re}\Lambda}{\mathrm{d}\operatorname{Re}\nu} = \frac{\mathrm{d}\operatorname{Im}\Lambda}{\mathrm{d}\operatorname{Im}\nu} = -c_{\mathrm{g}},$$

where we used the Cauchy–Riemann equations for the complex analytic function $\Lambda(\nu)$ in the second equality. In particular, if the group velocity c_{g} is positive, then positive weight rates $\eta > 0$, which predominantly measure mass accumulating at $x \to -\infty$, push dispersion curves $\Lambda(\nu)$ toward the stable direction since $\frac{\mathrm{d}\operatorname{Re}\Lambda}{\mathrm{d}\eta} < 0$. This can in fact be viewed as a justification of the terminology for c_{g} in the sense that the group velocity measures transport from negative to positive x.

From now on, we shall always denote the temporal Floquet exponents of wave trains in the comoving frame by λ and in the laboratory frame by Λ.

2.3. Absolute spectra of wave trains. When we pass to large bounded domains with separated boundary conditions, exponential weights generate equivalent topologies for each finite domain size L. In [34], we showed that the spectrum of the linearized period map, considered on large but finite domains with typical separated boundary conditions, converges in the limit of infinite domain size. We proved that this limit is given generically by the absolute spectrum which can be computed using only the Wronskian $\mathcal{D}(\Lambda, \nu)$ and which typically consists of a locally finite collection of semialgebraic curves.

Since the absolute spectrum is related to separated boundary conditions, it depends crucially on the frame in which the boundary conditions are imposed. As we are primarily interested in 1D sources and two-dimensional (2D) spiral waves for which only the laboratory frame is relevant, we shall compute the absolute spectrum of wave trains in this frame. To define absolute spectra in the laboratory frame, we fix a point $\Lambda \in \mathbb{C}$ and collect all roots ν of the Wronskian $\mathcal{D}(\Lambda, \nu)$ subject to $0 \le \operatorname{Im}\nu < k$. As shown in [38, section 3.4], these roots form a countable set $\{\nu_j\}_{j\in\mathbb{Z}}$ which depends on the choice of $\Lambda \in \mathbb{C}$. Taking the restriction on the imaginary part of the ν_j into account, we conclude from [23] that there are only finitely many roots ν_j, counted with multiplicity as solutions to an analytic equation, in any bounded region of the complex plane. Furthermore, [38, section 3.4] implies that there are infinitely many roots with negative real parts and infinitely many roots with positive real parts. We may therefore order the roots ν_j, repeated with multiplicity, according to their real part

(2.18)
$$\cdots \le \operatorname{Re}\nu_{-k} \le \operatorname{Re}\nu_{-k+1} \le \cdots \le \operatorname{Re}\nu_{-1} \le \operatorname{Re}\nu_0 \le \operatorname{Re}\nu_1 \le \cdots \le \operatorname{Re}\nu_k \le \operatorname{Re}\nu_{k+1} \le \cdots,$$

which gives a well-defined labeling up to shifts in the indices and up to the ambiguity of labeling roots with equal real parts. For $\operatorname{Re}\Lambda \gg 1$, each ν_j has nonzero real part since the essential spectrum would otherwise extend arbitrarily far to the right in the complex plane. We may therefore choose the labeling in (2.18) so that $\operatorname{Re}\nu_0 < 0 < \operatorname{Re}\nu_1$ for $\operatorname{Re}\Lambda \gg 1$. We then define the *absolute spectrum* in the laboratory frame as the set

$$\Sigma_{\mathrm{abs}} = \{\Lambda \in \mathbb{C};\ \operatorname{Re}\nu_0 = \operatorname{Re}\nu_1\}. \tag{2.19}$$

We say that the absolute spectrum is *simple* if $\operatorname{Re}\nu_{-1} < \operatorname{Re}\nu_{0,1} < \operatorname{Re}\nu_2$ and call points where $\nu_0 = \nu_1$ *edges* of the absolute spectrum. Edges in simple absolute spectrum are called simple edges, and it is straightforward to see that a unique curve of absolute spectrum emerges from each simple edge. More generally, the absolute spectrum comes in curves, being defined by a single real condition for the complex parameter Λ, and we may naturally parametrize these curves using the parameter

$$s = (\operatorname{Im}\nu_1 - \operatorname{Im}\nu_0)^2$$

so that edges correspond to $s = 0$.

Inspecting (2.19) shows that the absolute spectrum also respects the symmetries (2.15) of the essential spectrum, namely, complex conjugation and the artificial Floquet covering symmetry $\Lambda \mapsto \Lambda + \mathrm{i}\omega$. In particular, we have the following analogue of Lemma 2.1.

Lemma 2.2 (robustness of absolute spatio-temporal period-doubling). *Suppose that a simple edge of the absolute spectrum $\Lambda(0)$ is located at $\operatorname{Im}\Lambda(0) = -\omega/2$ for $\operatorname{Im}\nu_0 = k/2$; then the unique dispersion curve emanating from $\Lambda(0)$ is horizontal; that is, $\operatorname{Im}\Lambda(s) = -\omega/2$ for $s \approx 0$. Moreover, the same conclusion holds for sufficiently small perturbations of the parameter value μ and the coefficients $u_{\mathrm{wt}}(\xi)$ in (2.6). In particular, the absolute spectrum crosses at the sharp resonance $-\mathrm{i}\omega/2$ for an open subset of one-parameter families of reaction-diffusion systems.*

Proof. The proof is similar to the proof of Lemma 2.1 and will be omitted. ∎

We emphasize that the crossing of the essential spectrum at $\pm\mathrm{i}\omega/2$ does not necessarily enforce the absolute spectrum to cross at resonance. The other generic possibility is that the absolute spectrum consists locally of two curves which are symmetric about $\operatorname{Im}\Lambda = \pm\omega/2$ but do not contain any points with $\operatorname{Im}\Lambda = \pm\omega/2$.

2.4. Spatially homogeneous oscillations. We show here that the hypotheses stated in Lemmas 2.1 and 2.2 are met for wave trains with small wave numbers that accompany spatially homogeneous oscillations. Indeed, assume that (2.2) admits a solution $u_{\mathrm{wt}}(-\omega t)$ which undergoes a generic temporal period-doubling bifurcation at $\mu = 0$. Moreover, assume that the Floquet spectrum of the linearized period map $\Phi_T'(u_{\mathrm{wt}})$ of (2.1) is contained in the open left half-plane except for simple edges at $\Lambda = 0$, $\Lambda = \pm\mathrm{i}\omega/2$, and their Floquet conjugates (note that the absolute and essential spectra of homogeneous oscillations coincide since these waves are invariant under the spatial reflections $x \mapsto -x$).

Lemma 2.3. *Under the assumptions stated above, there exists a family of wave trains, parametrized by their wave number k with $k \approx 0$, each of which undergoes a spatio-temporal period-doubling which satisfies the hypotheses of Lemmas 2.1 and 2.2.*

We remark that the statement of the preceding lemma will be further extended in section 5.

Proof. The existence problem and the eigenvalue problem of wave trains with wave number $k = \varepsilon \approx 0$ yield the singularly perturbed boundary-value problems

$$\varepsilon^2 D\partial_\xi^2 u + \omega\partial_\xi u + f(u) = 0, \qquad D(\varepsilon\partial_\xi + \nu)^2 v + \omega(\partial_\xi + \nu/\varepsilon)v + f'(u)v = \lambda v,$$

respectively, with 2π-periodic boundary conditions in $\xi = \varepsilon x$. The eigenvalue problem can be rewritten in the form

$$D(\varepsilon\partial_\xi + \nu)^2 v + \omega\partial_\xi v + f'(u)v = \Lambda v,$$

using the definition (2.14) of Λ. Writing these second-order equations as first-order equations and reducing the dynamics to a slow manifold using geometric singular perturbation theory

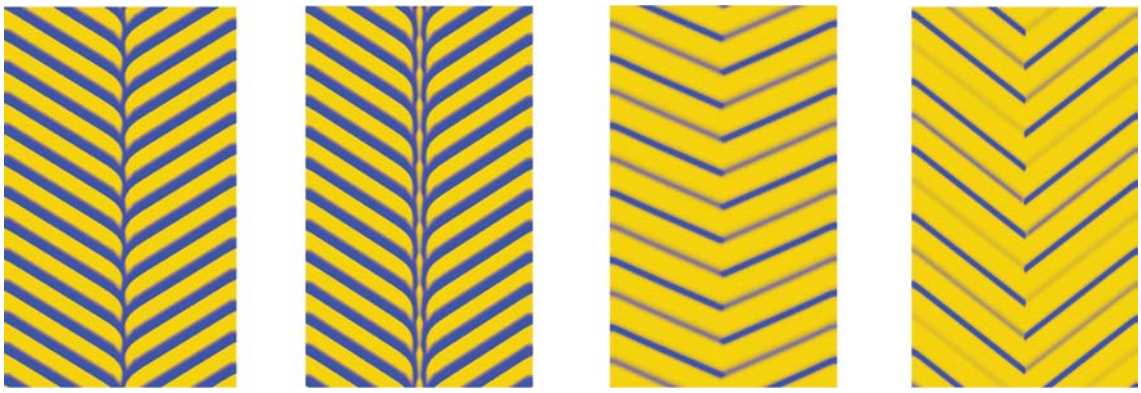


Figure 2. *From left to right: Plotted are space-time plots (time upward, space horizontal) of 1D spirals and 1D targets of the Brusselator with parameters as in* [38, *Appendix* B]*, and 1D spirals and moving period-doubled sources of the Rössler system* (6.1) *with* $C = 3.0$ *and* $C = 4.2$*, respectively.*

as in [38, section 3.3] shows that bounded solutions lie on the slow manifold and that the evolution on the slow manifold is obtained to leading order by formally setting $\varepsilon = 0$ in the system above. The reduced system therefore consists of a regular perturbation of a generic period-doubling bifurcation with a simple Floquet multiplier, which proves the claim about existence. The spectral problem with simple edges at $\Lambda = 0$ and $\Lambda = -\mathrm{i}\omega/2$ for $\nu = 0$ and $\nu = \mathrm{i}k/2$, respectively, is robust as well and yields the same spectral picture for $\varepsilon \approx 0$ with a possible offset in the real part of the period-doubling eigenvalue. ■

3. Period-doubling of sources in one space dimension. 2D spirals are defects in the sense that, far away from the location of the spiral tip, the medium resembles locally the essentially 1D planar wave trains that we encountered in the previous section. An additional property of 2D spirals is the active emission of wave trains in the sense that the group velocity of the planar wave trains that are observed in the far field points in the radial direction away from the center of the spiral.

Sources are 1D analogues of spiral waves, and we discuss in this section the 1D analogue (see Figure 2) of the period-doubling instability of 2D spiral waves. We are interested in seeing whether Floquet multipliers cross exactly at $\rho = -1$ or only nearby, and whether the bifurcating period-doubled sources will drift or not.

3.1. Sources. Recall the reaction-diffusion system

$$u_t = Du_{xx} + f(u;\mu), \qquad x \in \mathbb{R}. \tag{3.1}$$

Standing sources $u_*(x,t)$ are time-periodic solutions of (3.1) which converge to wave trains $u_{\mathrm{wt}}^{\pm}$ in the far field as $x \to \pm\infty$ whose group velocities $c_{\mathrm{g}}^{\pm}$, computed in the laboratory frame, point away from the interface so that $c_{\mathrm{g}}^{-} < 0$ and $c_{\mathrm{g}}^{+} > 0$. More precisely, we have

$$u_*(x,t) = u_*(x,t+T), \qquad |u_*(x,t) - u_{\mathrm{wt}}^{\pm}(k_\pm x - \omega_\pm t - \theta_\pm; k_\pm)| \to 0 \quad \text{as } x \to \pm\infty, \tag{3.2}$$

where $u_{\mathrm{wt}}(\xi;k)$ denotes a family of wave trains which are 2π-periodic in the argument ξ with temporal frequencies $\omega_\pm$, spatial wave numbers $k_\pm$, and phase corrections $\theta_\pm$ at $\pm\infty$,

respectively. Characteristic for sources is the requirement that the group velocities, computed in the laboratory frame, are directed away from the defect so that $c_g^- < 0$ and $c_g^+ > 0$. As before, we let $\omega = 2\pi/T$ denote the temporal frequency of the source. We showed in [38] that sources occur for open, nonempty classes of reaction-diffusion systems and generically for discrete sets of asymptotic wave numbers $k_\pm$.

We focus here exclusively on 1D targets and 1D spirals which are standing sources that are reflectionally symmetric so that

(3.3)
$$u_*(x,t) = u_*(-x,t) \qquad \text{(1D target)} \qquad \text{or} \qquad u_*(x,t) = u_*(-x,t+T/2) \qquad \text{(1D spiral)}$$

for all (x,t). Thus, 1D target patterns are even in x for each t, while 1D spirals are invariant when simultaneously reflecting the pattern and shifting in time by half the temporal period; see Figure 2. Reflectional symmetry implies in both cases that $k_+ = -k_- \neq 0$.

3.2. Spectra of sources on $\mathbb{R}$. Dynamic properties such as robustness, stability, and interaction with other defects are largely determined by spectral properties of the linearization Φ_T' of the time-T map of (3.1) about the defect. Throughout this section, we will switch back and forth between Floquet exponents Λ and Floquet multipliers $\rho = \mathrm{e}^{\Lambda T}$ in the spectrum of Φ_T'. We distinguish between the *point spectrum*, which consists of all $\rho \in \mathbb{C}$ for which $\Phi_T' - \rho$ is not invertible but still Fredholm of index zero, and the *essential spectrum*, which is the complement of the point spectrum in the spectrum. We showed in [38] that the Fredholm index of Φ_T' jumps precisely at the dispersion curves of the asymptotic 1D wave trains, computed in the frame of the defect. In particular, the essential spectrum of sources inherits the symmetry properties of the essential spectrum of the 1D wave trains.

Corollary 3.1 (robust period-doubling of sources). *There exists an open class of one-parameter families of reaction-diffusion systems where the spectrum of the linearization crosses the imaginary axis first at $\Lambda = \pm \mathrm{i}\omega/2$.*

The multiplicity of the essential spectrum depends on whether the underlying source is reversible in the sense of (3.3) or not: Since the asymptotic wave trains at $x = \pm\infty$ of 1D targets and 1D spirals are related by reflection symmetry, the essential spectra of both wave trains cross the imaginary axis simultaneously. Thus, the essential spectrum of reversible sources has geometric multiplicity two, which has implications for the actual bifurcation scenario which we will discuss in section 3.4.

We now discuss the point spectrum. There are no structural reasons that prevent 1D targets from having a simple point Floquet multiplier at $\rho = -1$ that crosses the imaginary axis, thus leading to a generic period-doubling bifurcation of time-periodic solutions of (3.1). Since 1D targets are symmetric under the spatial reflections defined by

$$(\mathcal{R}u)(x) := u(-x),$$

the linearized period map leaves the spaces $\mathrm{Fix}(\mathcal{R})$ and $\mathrm{Fix}(-\mathcal{R})$ of even and odd functions invariant [13]. The eigenfunction belonging to a simple multiplier $\rho = -1$ is therefore either even or odd: the bifurcating sources are 1D targets in the first case and 1D spirals in the second case, and they have approximately twice the temporal period in both cases.

Floquet multipliers of 1D spirals at $\rho = -1$ must, however, have geometric multiplicity two since the linearized period-T map $\Phi_T'(u_*)$ can be written as a square: Indeed, $\Phi_t(u)$ is

equivariant with respect to the spatial reflection $\mathcal{R}$ so that $\Phi_t(\mathcal{R}u) = \mathcal{R}\Phi_t(u)$ and therefore $\Phi'_{T/2}(\mathcal{R}u) = \mathcal{R}\Phi'_{T/2}(u)\mathcal{R}$. Using that $\Phi_{T/2}(u_*) = \mathcal{R}u_*$ for the 1D spiral u_*, we obtain

$$\Phi'_T(u_*) = \Phi'_{T/2}(\Phi_{T/2}(u_*))\Phi'_{T/2}(u_*) = \mathcal{R}\Phi'_{T/2}(u_*)\mathcal{R}\Phi'_{T/2}(u_*) = [\mathcal{R}\Phi'_{T/2}(u_*)]^2.$$

We first focus on the center subspace associated with critical point spectrum of $\mathcal{R}\Phi'_{T/2}(u_*)$. Clearly this subspace is invariant under $\Phi'_T(u_*)$. An eigenvalue -1 of $\Phi'_T(u_*)$ can be generated only by eigenvalues $\pm\mathrm{i}$ of $\mathcal{R}\Phi'_{T/2}(u_*)$, which come necessarily in complex conjugated pairs so that the eigenvalue -1 cannot be simple. If the essential spectrum of $\mathcal{R}\Phi'_{T/2}(u_*)$ is bounded away from $\pm\mathrm{i}$, then the essential spectrum of $\Phi'_T(u_*)$ is bounded away from -1 by Fredholm algebra properties, and the spectral projection P belonging to eigenvalues near -1 can be obtained by factoring

$$\Phi'_T(u_*) - \rho = [\mathcal{R}\Phi'_{T/2}(u_*) - \sqrt{\rho}\,]\,[\mathcal{R}\Phi'_{T/2}(u_*) + \sqrt{\rho}]$$

for ρ on a small circle Γ around -1, and computing

$$\begin{aligned} P &= \int_\Gamma [\rho - \Phi'_T(u_*)]^{-1}\,\mathrm{d}\rho \\ &= \int_\Gamma [\mathcal{R}\Phi'_{T/2}(u_*) - \sqrt{\rho}]^{-1}[\mathcal{R}\Phi'_{T/2}(u_*) + \sqrt{\rho}]^{-1}\,\mathrm{d}\rho \\ &= \int_{\sqrt{\Gamma}} \Big([\mathcal{R}\Phi'_{T/2}(u_*) - \sqrt{\rho}]^{-1} - [\mathcal{R}\Phi'_{T/2}(u_*) + \sqrt{\rho}]^{-1} \Big)\,\mathrm{d}\sqrt{\rho}. \end{aligned}$$

In particular, the spectral projection of $\Phi_T(u_*)$ associated with $\rho = -1$ is given by the sums of the spectral projections of $\mathcal{R}\Phi'_{T/2}$ associated with $\rho = \pm\mathrm{i}$. This shows that classical period-doubling in the form of a simple Floquet multiplier at -1 in the point spectrum cannot occur for 1D spirals.

Remark 3.2. *The preceding analysis also shows that the double Floquet multiplier at $\rho = -1$ for 1D spirals will generically split into two nonreal, complex conjugate multipliers since there is no structural reason which prevents the eigenvalues $\pm\mathrm{i}$ of $\mathcal{R}\Phi'_{T/2}(u_*)$ from moving off the imaginary axis, thus moving the multipliers of the square $\Phi'_T(u_*)$ off the negative real axis.*

In preparation for the discussion in section 3.3, we collect some properties of the spectra of symmetric sources in the exponentially weighted spaces

$$\hat{L}^2_\eta := \{u \in L^2_{\mathrm{loc}};\ |u|_{\hat{L}^2_\eta} < \infty\}, \qquad |u|^2_{\hat{L}^2_\eta} := \int_{\mathbb{R}} |u(x)|^2 \mathrm{e}^{-2\eta|x|}\,\mathrm{d}x.$$

The essential spectrum of the linearized period-T map of a symmetric source (3.3) on the space $\hat{L}^2_\eta$ is determined by the dispersion curves of the asymptotic wave trains in the spaces L^2_η from (2.17). Exploiting that the wave trains $u^\pm_{\mathrm{wt}}(x; k_\pm)$ are related by symmetry,

$$u^+_{\mathrm{wt}}(x; k_+) = u^-_{\mathrm{wt}}(-x; k_-), \qquad k_+ = -k_-,$$

it follows that the spatial Floquet exponents $\nu^\pm_j(\Lambda)$ that appear in the definition (2.19) of the absolute spectra of wave trains are related via

$$\nu^+_j(\Lambda) = -\nu^-_{1-j}(\Lambda) \qquad \forall j.$$

In particular, we have $\nu_0^+(\Lambda) = -\nu_1^-(\Lambda)$ and $\nu_1^+(\Lambda) = -\nu_0^-(\Lambda)$ so that the absolute spectra of the asymptotic wave trains coincide and so that, for each $\Lambda \notin \Sigma_{\mathrm{abs}}$, we can find a weight η with $\nu_0^+ < \eta < \nu_1^+$ such that the linearized period map $\Phi_T'(u_*) - \mathrm{e}^{\Lambda T}$ is Fredholm with index zero in $\hat{L}^2_\eta$. Note also that we can choose the weight η to be constant locally in Λ. We define the *extended point spectrum* to be the set of $\Lambda \notin \Sigma_{\mathrm{abs}}$ such that the linearized period map is not invertible in the space $\hat{L}^2_\eta$ with η chosen as described above. We can then also define geometric and algebraic multiplicities for elements of the extended point spectrum.

3.3. Spectra of sources on finite intervals. The resonant crossing of essential spectra of sources provides some evidence for why period-doubling can occur in a robust fashion. Experiments and numerical simulations are, however, posed on large but finite domains, typically with separated boundary conditions. The linearized period map on bounded domains is a compact operator, and the essential Floquet spectrum is therefore empty: Instead, the absolute spectra of the asymptotic wave trains become relevant. For simplicity, we restrict ourselves to Neumann boundary conditions, thus considering

$$\begin{aligned} u_t &= Du_{xx} + f(u;\mu), \qquad x \in (-L, L), \\ 0 &= u_x(\pm L, t), \end{aligned} \tag{3.4}$$

which are realistic for the experimental setup and which are also a standard choice for numerical simulations.

We shall focus exclusively on 1D targets and 1D spirals. We assume that there exists a symmetric source u_{so} such that $\Lambda = 0$ belongs to the extended point spectrum with algebraic and geometric multiplicity two. We showed in [38] that this assumption is satisfied for an open and nonempty set of reaction-diffusion systems and called this type of source elementary. To describe the influence of the boundary, we need an assumption on the boundary layer between wave trains and the boundary. We assume that (3.1) admits a *symmetric sink* u_{si}, that is, a solution of the form (3.2) which is even in x, with asymptotic wave numbers $k_{\mathrm{si}}^{\pm} := k_{\mathrm{so}}^{\mp}$. In particular, the group velocities of the sink point toward the center of the sink. We assume that $\Lambda = 0$ does not belong to either the absolute or the extended point spectrum of the linearization $\Phi_T'(u_{\mathrm{si}})$ of the period map. Again, this assumption is robust [38]. Since the sink is even, it gives solutions $u_{\mathrm{si}}^{\pm}$ of the system

$$\begin{aligned} u_t &= Du_{xx} + f(u;\mu), \qquad x \in \mathbb{R}^{\pm}, \\ 0 &= u_x(0, t) \end{aligned} \tag{3.5}$$

on the half-spaces $\mathbb{R}^+$ and $\mathbb{R}^-$ which satisfy Neumann boundary conditions at $x = 0$, and we refer to these two solutions on $\mathbb{R}^+$ and $\mathbb{R}^-$ as *boundary sinks* [38]. In this setup, we proved the following result on the existence and spectral properties of solutions on large bounded intervals.

Theorem 1 (see [38, section 6.8]). *Under the above assumptions, the reaction-diffusion system* (3.4) *has, for each* $L \gg 1$, *a unique time-periodic solution* $u_*(x, t; L)$ *which is close to the symmetric source* u_{so} *on* $(-L/2, L/2)$ *and to the appropriately translated boundary sinks* u_{si}^+ *and* u_{si}^- *on* $(-L, -L/2)$ *and* $(L/2, L)$, *respectively.*

The assumption on the existence of a symmetric sink can be verified in the special case of nearly homogeneous oscillations (see also section 5.4). Recall that a homogeneous oscillation with a simple Floquet exponent $\Lambda = 0$ is accompanied by a family of wave trains $u_{\mathrm{wt}}(kx - \omega(k)t)$ for small wave numbers $k \approx 0$.

Theorem 2 (see [6]). *Assume that there is a spatially homogeneous oscillation $u_{\mathrm{wt}}(-\omega(0)t)$ such that the Floquet multiplier $\Lambda = 0$ is a simple edge. For each $k \approx 0$, there exists a unique symmetric sink which is spatially asymptotic to the wave trains $u_{\mathrm{wt}}(\pm kx - \omega(\pm k)t)$ at $x = \pm\infty$.*

Next, we investigate the spectrum of the linearized period map near the truncated sources that we described in Theorem 1. An outline of the proof of the following theorem will be given in Appendix A.

Theorem 3. *Assume that the extended point spectrum of the sources u_{so} on $\mathbb{R}$ is discrete; then the spectrum of the period map of the truncated sources described in Theorem 1 converges locally uniformly in the symmetric Hausdorff distance to the disjoint union of the absolute spectrum Σ_{abs} of the wave trains u_{wt}, computed in the laboratory frame, and a discrete set of isolated points.*

The convergence toward the absolute spectrum is algebraic of order $\mathrm{O}(1/L)$, and the number of eigenvalues in any small neighborhood of any element of the absolute spectrum converges to infinity as $L \to \infty$. The discrete part of the limiting spectrum is the union of the extended point spectrum of the source u_{so} on $\mathbb{R}$ and the extended point spectra of the two symmetric boundary sinks $u_{\mathrm{si}}^{\pm}$ on $\mathbb{R}^{\pm}$ with Neumann boundary conditions. The convergence toward the discrete part is exponential in L, and the multiplicity of eigenvalues in any small neighborhood of the discrete part is finite and stabilizes as $L \to \infty$.

We remark that the absolute spectrum is close to the essential spectrum if the wave number of the asymptotic wave trains is sufficiently close to zero; see Lemma 2.3 and also section 5 below. We now discuss the implications of Theorem 3 for period-doubling of symmetric sources of (3.4).

We begin with 1D targets. Floquet exponents $\rho = -1$ in the Floquet point spectrum of a 1D target u_{so} on $\mathbb{R}$ will generically have multiplicity one and therefore persist as a simple multiplier $\rho \in \mathbb{R}^-$ near -1 for (3.4), with the eigenfunction lying again in the space of even or odd functions. Next, assume that the boundary sink u_{si}^+ on $\mathbb{R}^-$ with Neumann conditions has a simple Floquet multiplier $\rho = -1$ in its point spectrum. Since the sinks u_{si}^+ and u_{si}^- are related by reflection $x \mapsto -x$, the reflected sink u_{si}^- also has a simple Floquet multiplier $\rho = -1$, and Theorem 3 shows that the truncated source u_* has two Floquet multipliers near $\rho = -1$. Since 1D targets are symmetric under spatial reflections, the linearized period map leaves the spaces of even and odd functions invariant. The Floquet eigenfunctions of the boundary sinks on $\mathbb{R}^{\pm}$ yield one even and one odd eigenfunction of the truncated 1D target of (3.4), which can be seen via transversality arguments in a spatial-dynamics formulation of the eigenvalue problem. In particular, the two Floquet multipliers of the persisting source on the bounded interval are both real and close to -1 but may split on the negative real line. Thus, two period-doubling bifurcations will take place, both with multipliers at $\rho = -1$, one with an even and the other one with an odd eigenfunction. In summary, period-doubling of 1D targets for (3.1) via point eigenvalues persists with a sharp resonance at $\rho = -1$ for (3.4), and the resulting bifurcation leads to 1D targets and/or 1D spirals depending on the symmetries of the associated eigenfunctions.

Next, we consider 1D spirals. Symmetry enforces that Floquet multipliers $\rho \in \mathbb{R}^-$ of 1D spirals on $\mathbb{R}$ are double. This symmetry is also present for (3.4), and Remark 3.2 shows that a Floquet multiplier $\rho = -1$ therefore either persists as a double multiplier $\rho \in \mathbb{R}^-$ near -1 or will split into two complex conjugate multipliers. The same conclusion is true for the two multipliers near $\rho = -1$ that arise when the two boundary sinks undergo period-doubling with simple multipliers at $\rho = -1$. We expect that the two multipliers near -1 will generically split, so that there is no sharp resonance at $\rho = -1$ for the truncated 1D spiral of (3.4).

Last, we consider the absolute spectrum. The following corollary is a straightforward consequence of Theorem 3.

Corollary 3.3 (generic absolute period-doubling). *Resonant crossing of eigenvalues near the absolute spectrum at* $\Lambda = \pm i\omega/2 + O(1/L)$ *occurs in an open subset of one-parameter families of reaction-diffusion systems.*

For 1D spirals, we expect that, generically, the eigenvalues near the absolute spectrum will indeed move off the lines $\operatorname{Im}\Lambda = \pm\omega/2$. For 1D targets, we can, however, apply the same symmetry-based arguments as above which yield that the absolute eigenmodes decompose again into odd and even functions: This precludes movement of the associated Floquet multipliers off the lines $\operatorname{Im}\Lambda = \pm\omega/2$, and we therefore obtain a sharp resonance with multipliers on these lines.

3.4. Nonlinear bifurcations of 1D sources, and the role of group velocity. We now analyze the period-doubling instability of 1D targets and 1D spirals on the unbounded real line $x \in \mathbb{R}$ that arises when essential spectrum crosses the imaginary axis. We are interested in constructing coherent structures which are periodic in time and spatially asymptotic to period-doubled wave trains in the far field as shown in Figure 2. Our goal is to derive bifurcation and bifurcation failure results which are valid uniformly in the size of the domain. Our approach will also allow us to gain insight into the role of transport as represented by the group velocity of the linear period-doubling modes.

Throughout this section, we assume the existence of a family of wave trains with nonzero group velocity c_g which undergo a period-doubling instability with dispersion curve $\Lambda_{pd}(\nu)$ which satisfies

$$\Lambda_{pd}(ik_*/2) = -i\omega_*/2, \qquad c_g^{pd} = -\Lambda'_{pd}(ik_*/2) \neq 0, \qquad \operatorname{Re}\Lambda_{pd}(\nu) < 0 \quad \forall \nu \neq ik_*/2.$$

Furthermore, we assume that the period-doubling bifurcation is supercritical (more precisely, that the period-doubling bifurcation in the space of spatially periodic functions is a supercritical pitchfork bifurcation). We will now state three theorems on period-doubling bifurcations from defects on $\mathbb{R}$ which we shall prove later in this section.

Theorem 4 (bifurcation from 1D targets on $\mathbb{R}$). *Assume that there exists a* 1*D target with* $\Sigma_{ext} \cap i\mathbb{R} = \{0\}$, *where* $\Lambda = 0$ *has multiplicity two, whose asymptotic wave trains undergo period-doubling at* $\mu = 0$. *If* $c_g^{pd} < 0$, *then there exists a unique branch of bifurcating* 1*D target patterns and a unique branch of* 1*D spirals which are asymptotic to the period-doubled wave trains. If* $c_g^{pd} > 0$, *then* 1*D target patterns and* 1*D spirals that are asymptotic to the period-doubled wave trains do not exist near onset.*

Theorem 5 (bifurcation from 1D spirals on $\mathbb{R}$). *Assume that there exists a* 1*D spiral with temporal frequency* ω_* *and with* $\Sigma_{ext} \cap i\mathbb{R} = \{0\}$, *where* $\Lambda = 0$ *has multiplicity two, whose*

asymptotic wave trains undergo period-doubling at $\mu = 0$. *If* $c_g^{pd} < 0$, *then there exists a unique branch, up to spatial reflection, of bifurcating solutions which are asymptotic to the period-doubled wave trains. The wave speed* c_* *of the bifurcating solutions is close to zero with* $|c_*| \le K|\mu|$ *for some constant* K, *and their temporal frequency is close to* $\omega_*/2$ *in the comoving frame* $\xi = x - c_* t$. *If* $c_g^{pd} > 0$, *then standing or moving* $1D$ *target patterns or* $1D$ *spirals that are asymptotic to the period-doubled wave trains do not exist near onset.*

We have observed the bifurcation from 1D spirals to moving period-doubled sources described in Theorem 5 in numerical simulations of the Rössler system; see the two rightmost plots in Figure 2.

While bifurcations from sources to period-doubled sources occur on $\mathbb{R}$ if and only if the group velocity of the period doubling modes is directed toward the center of the defect, the following result shows that bifurcations from boundary sinks to period-doubled boundary sinks take place if and only if the group velocity is directed toward the boundary.

Theorem 6 (bifurcation of boundary layers on $\mathbb{R}^-$). *Assume that there exists a* $1D$ *boundary sink of* (3.5) *on* $\mathbb{R}^-$ *such that* $\Sigma_{ext} \cap i\mathbb{R} = \emptyset$. *If* $c_g^{pd} > 0$, *then there exists a unique branch of bifurcating boundary sinks which are asymptotic to the period-doubled wave trains. If* $c_g^{pd} < 0$, *then boundary sinks that are asymptotic to the bifurcating period-doubled wave trains do not exist near onset.*

Combining the statements on the bifurcation of coherent structures and boundary sinks, we see that we cannot expect the simultaneous bifurcation to both coherent structures on $\mathbb{R}$ and boundary layers on $\mathbb{R}^-$ near the onset of an essential instability. In particular, if period-doubled sources bifurcate on $\mathbb{R}$, then period-doubled boundary sinks will not be present, and consequently the period-doubled sources will not persist on large bounded domains with Neumann boundary conditions. We refer the reader to section 5.4 for the analysis of a scenario where sources on $\mathbb{R}$ persist on bounded domains due to instabilities of boundary sinks caused by the point spectrum.

The results stated above reflect an intuitive heuristic picture of transport. The linear group velocity of eigenmodes encodes the direction toward which a localized perturbation constructed from the eigenmode will propagate. Thus, if we begin with a source on $\mathbb{R}$, then the group velocities c_g^{pd} of the period-doubling modes at $x = \pm\infty$ determine whether period-doubling modes can propagate toward the core of the source or not: If the group velocity c_g^{pd} at $x = \infty$ is positive, then the period-doubling modes cannot propagate toward the core, and a nonlinear bifurcation to a period-doubled pattern is not possible. The same arguments apply to boundary sinks provided we interpret the boundary as the core of the coherent structure. Since transport occurs either away from the boundary toward the core, or else away from the core toward the boundary, we cannot expect the simultaneous bifurcation of sources and boundary sinks.

In the remainder of this section, we sketch the proofs of Theorems 4–6. The arguments are similar to those given in [36], and we will therefore refer the reader to [36] for the more technical aspects of the proofs. Since we are interested only in time-periodic solutions, we rewrite the

reaction-diffusion equation as a first-order evolution equation in the spatial variable x,

$$\begin{aligned} u_x &= v, \\ v_x &= D^{-1}[\omega u_\tau - cv - f(u;\mu)], \end{aligned} \tag{3.6}$$

where $(u,v)(\cdot) \in H^1(S^1,\mathbb{R}^n) \times H^{1/2}(S^1,\mathbb{R}^n)$ are 2π-periodic in τ for each fixed x. We showed in [30, 36, 37] that coherent structures can be found as intersections of stable and unstable manifolds of the periodic orbits (in the evolution variable x) that correspond to the asymptotic wave trains. Essential instabilities of the wave trains correspond to pitchfork bifurcations of the corresponding periodic orbits. Although (3.6) is ill-posed, the aforementioned stable and unstable manifolds exist, and the standard arguments for bifurcations in dynamical systems can be made rigorous [36, 37, 38, 39]. Equation (3.6) is invariant under the action of the symmetry group SO(2) via the time shift action $\mathcal{S}_\theta : (u,v)(\tau) \mapsto (u,v)(\tau+\theta)$ for each fixed $\theta \in \mathrm{SO}(2) \cong \mathbb{R}/2\pi\mathbb{Z}$. In particular, the subspace of solutions which are invariant under the shift $\mathcal{S}_\pi : \tau \mapsto \tau + \pi$ by half the period is invariant under the x-evolution. If we choose $\omega \approx \omega_*/2$, where ω_* is the temporal frequency of the primary source, then the primary wave trains and coherent structures lie in this fixed-point space, and we expect the period-doubled structures to bifurcate out of this subspace.

For the sake of clarity, we shall pretend in the rest of this section that the dynamics of (3.6) can be reduced to an appropriate six-dimensional invariant subspace, for instance, by using a suitable combination of Fourier modes in the periodic variable τ. We assume that this subspace intersects the fixed-point subspace of the shift by half a period in a four-dimensional subspace. Since all solutions truly depend on τ, we may furthermore factor out the free action of $\mathrm{SO}(2)/\mathbb{Z}_2$. We will now describe the dynamics of the reduced spatial dynamical system

$$U_x = F(U;\mu), \qquad U \in \mathbb{R}^3 \times \mathbb{R}^2, \tag{3.7}$$

in the reduced phase space which we assume is given by $\mathbb{R}^3 \times \mathbb{R}^2$. As mentioned above, the technical tools necessary to extend the analysis of this "toy problem" to the full system (3.6) have been described in [36, 37, 38, 39] to which we refer the reader for details.

The primary wave trains $u_{\mathrm{wt}}(kx-\tau)$ correspond to relative periodic orbits of (3.6) with respect to the shift symmetry $\mathcal{S}_\theta$ and therefore, upon factoring out the shift, to equilibria of the reduced spatial system (3.7) in $\mathbb{R}^3 \times \{0\}$. Any neutral Floquet exponent $\Lambda \in \mathrm{i}\omega\mathbb{Z}$ of the dispersion curve $\Lambda_*(\mathrm{i}\gamma)$ of the linearization of the reaction-diffusion system (2.1) about the asymptotic wave train u_{wt} gives a neutral Floquet exponent $\nu = \mathrm{i}\gamma \in \mathrm{i}\mathbb{R}$ of the corresponding relative periodic orbit of (3.6). By assumption, there are precisely two such neutral eigenvalues, namely, the phase eigenvalue $\Lambda = 0$ at $\nu = 0$ and the period-doubling mode $\Lambda = -\mathrm{i}\omega_*/2$ at $\nu = \mathrm{i}k_*/2$. Upon factoring out the shift symmetry $\mathrm{SO}(2)/\mathbb{Z}_2$, we see that the eigenvalue $\Lambda = 0$, which corresponds to shifts, is removed, while the period-doubling eigenvalue gives a Floquet exponent at $\nu = 0$ with eigenvector contained in $\{0\} \times \mathbb{R}^2$. Inside the invariant subspace $\mathbb{R}^3 \times \{0\}$, the equilibrium is hyperbolic with one unstable eigenvalue if $c_{\mathrm{g}} < 0$ and two unstable eigenvalues if $c_{\mathrm{g}} > 0$; see [38]. Similarly, besides the neutral eigenvalue $\nu = 0$, the equilibrium has one unstable eigenvalue on $\{0\} \times \mathbb{R}^2$ if $c_{\mathrm{g}}^{\mathrm{pd}} > 0$ and one stable eigenvalue if $c_{\mathrm{g}}^{\mathrm{pd}} < 0$.

The reversers $\mathcal{R}_0 : (u, v) \mapsto (u, -v)$ and $\mathcal{R}_\pi := \mathcal{R}_0\mathcal{S}_\pi$ each fix a three-dimensional subspace in $\mathbb{R}^6$ which is invariant under the action of SO(2), thus yielding a two-dimensional subspace in $\mathbb{R}^5$ whose intersection with the isotropy subspace $\operatorname{Fix}\mathcal{S}_\pi$ is one-dimensional. Similarly, the space $\{(u, v);\ v = 0\}$ of functions that satisfy Neumann boundary conditions corresponds to a two-dimensional subspace in $\mathbb{R}^5$ which intersects $\operatorname{Fix}\mathcal{S}_\pi$ in a line.

We now prove Theorem 4 for target patterns on $\mathbb{R}$. Before bifurcation for $\mu < 0$, 1D targets are found as intersections in $\mathbb{R}^5$ of the two-dimensional space $\operatorname{Fix}\mathcal{R}_0$ with the two-dimensional stable manifold of the equilibrium corresponding to the wave train with positive group velocity $c_{\mathrm{g}} > 0$. The assumption that the extended point spectrum in the origin has multiplicity two means that the intersection of the tangent spaces of the stable manifold and $\operatorname{Fix}\mathcal{R}_0$ is trivial and is broken with nonvanishing speed when we vary ω near $\omega_*/2$ [38].

First, assume $c_{\mathrm{g}}^{\mathrm{pd}} > 0$. From the preceding discussion of the dispersion relation, we see that the wave train is stable inside the one-dimensional center manifold for $\mu < 0$ before the onset of period-doubling. The stable manifold of the asymptotic wave train can therefore be continued smoothly through the bifurcation as a center-stable manifold. The assumption of minimal extended point spectrum implies that the intersection between $\operatorname{Fix}\mathcal{R}_0$ and the center-stable manifold of the equilibrium is transverse in the parameter ω at $\mu = 0$, and we conclude that the unique intersection persists through the bifurcation. Since this unique intersection is given by the primary 1D target pattern and therefore located inside $\operatorname{Fix}\mathcal{S}_\pi$, we conclude that period-doubled target patterns cannot bifurcate.

Next, assume that $c_{\mathrm{g}}^{\mathrm{pd}} < 0$. The wave train is then unstable inside the center manifold for $\mu < 0$ before bifurcation, which means that the stable manifold continues continuously through the bifurcation as the strong stable fiber. At $\mu = 0$, the strong stable fiber of the primary wave train crosses $\operatorname{Fix}\mathcal{R}_0$ transversely upon varying ω, and we conclude that the primary 1D target persists. On the other hand, the strong stable fiber of the bifurcating period-doubled wave trains is $\sqrt{\mu}$-close to the strong stable fiber of the primary wave train, and it therefore also crosses $\operatorname{Fix}\mathcal{R}_0$ transversely for $\tilde{\omega} = \omega + \mathrm{O}(\sqrt{\mu})$.

The same arguments apply when we replace $\operatorname{Fix}\mathcal{R}_0$ by $\operatorname{Fix}\mathcal{R}_\pi$, which completes the proof of Theorem 4.

The case of 1D spirals is similar. The primary 1D spirals are transverse intersection of the stable manifold of the wave train and the fixed-point space of the operator $\mathcal{R}_0\mathcal{S}_{\pi/2}$, which acts as a reverser in $\operatorname{Fix}\mathcal{S}_\pi$ but not in the entire phase space: Indeed, the flip symmetry of the shift $\mathcal{S}_{\pi/2}$ by half a period has order four after doubling the period and therefore cannot act as an involution when composed with the reverser $\mathcal{R}_0 : (u, v) \mapsto (u, -v)$. We therefore cannot expect to obtain period-doubled patterns as intersections with reversibility fixed-point spaces; instead the bifurcating patterns should drift. Thus, we transform into a comoving frame, include the wave speed $c \approx 0$ as an additional parameter, and seek intersections of the unstable manifold of the primary wave train with negative group velocity at $x = -\infty$ with the stable manifold of the primary wave train with positive group velocity at $x = \infty$. Since the intersection now occurs along flow lines of the differential equation, we lose one dimension for transversality, which is however compensated for by the additional parameter c. The existence and nonexistence proofs for period-doubled sources proceed now as before, and [33, Lemma 3.9] shows that the speed c of the bifurcating sources will be of the order $\mathrm{O}(\mu)$. We omit the straightforward adaptation of the arguments.

It remains to discuss boundary sinks on $\mathbb{R}_-$, which we seek as transverse intersections of the unstable manifold of the wave trains at $x = -\infty$ with positive group velocity and the boundary subspace. Since $c_g > 0$ at $x = -\infty$, we have transversality of the intersection for fixed ω, and we consequently find a family of boundary sinks, parametrized by their temporal frequency ω. For $c_g^{pd} > 0$, we find a family of period-doubled boundary sinks by continuing the strong unstable manifold of the wave trains continuously through the bifurcation as the strong unstable manifold of the period-doubled wave trains after bifurcation. For $c_g^{pd} < 0$, period-doubled boundary sinks can bifurcate only near certain discrete values of ω where the transversality conditions are violated, and an additional extended point spectrum occurs in the origin: An example where this can occur is near $k = 0$, and we refer the reader to section 5.4 for an analysis of the resulting scenario.

3.5. Nonlinear bifurcations of 1D sources on finite intervals. We now describe bifurcations on large bounded domains induced by the crossing of the absolute spectrum. Theorem 3 shows that there will be a large number of eigenvalues near each point of the absolute spectrum. Thus, for large domain diameters $L \gg 1$, we expect a sequence of bifurcations with a delayed onset $\mu_*(L) = \mu_*^\infty + O(1/L^2)$ of the instability compared with the crossing of the absolute spectrum at $\mu = \mu_*^\infty$. The small-amplitude regime of this bifurcation sequence can be analyzed using the methods described in [40] for the analogous case of a pitchfork bifurcation (matching with the reversibility lines here is equivalent to matching with the boundary conditions described there). As in [40], we expect that the amplitude of the bifurcating pattern in the far field scales with $\sqrt{\mu - \mu_*(L)}L^{3/2}$. Instead of carrying out the analysis of the entire bifurcation sequence, we focus here on the first bifurcation.

We start with the case of a 1D target pattern. If the absolute spectrum crosses the axis at $\rho = -1$, then section 3.3 shows that clusters of Floquet multipliers pass through $\rho = -1$ on the real axis. From the expansion at the edge of the absolute spectrum [34, section 5.4], we conclude that the first instability induced by the absolute spectrum occurs on a two-dimensional center-eigenspace, with multipliers passing through -1 within $O(1/L)$ of each other as functions of the bifurcation parameter μ, where the reflection symmetry acts trivially in one direction and nontrivially in the other direction. In physical space, this can be interpreted as synchronizing the instability in the far field without a phase shift or with a phase shift of π. The resulting bifurcation can be analyzed using Lyapunov–Schmidt reduction (for finite, large L) and exploiting the symmetry. We find again that 1D target patterns bifurcate in the space of symmetric functions, while 1D spirals, which are invariant under the flip symmetry $(x, t) \mapsto (-x, t+T)$, bifurcate in the space on which the symmetry acts nontrivially. We remark that the actual bifurcation is guaranteed on both spaces by degree arguments and the fact that the leading multipliers actually cross the imaginary axis. The absolute period-doubling of a 1D target pattern already shows "nongenericity" in the sense that we would typically expect the leading multipliers to be simple at resonance $\rho = -1$. The bifurcation analysis as described above remains valid only in a very small range of parameter values since the spectral gap to the next multiplier is only of order $1/L^2$.

Next, we consider 1D spirals. For 1D spirals, we cannot eliminate the translation symmetry by restricting to an appropriate fixed-point space since the isotropy of a 1D spiral is trivial for each fixed time t, and instead we need to consider the entire center manifold at once. Near a period-doubling bifurcation induced by the absolute spectrum, the linearization $\Phi_T'(u_*)$ has

four Floquet multipliers in an $O(1/L^2)$-neighborhood of the unit circle given by $\rho = 1$ from temporal translation, $\rho = O(e^{-\delta L})$ for the translation eigenvalue, and $\rho = -1-\mu+O(1/L)$ for the period-doubling multiplier, which has geometric multiplicity two as shown in section 3.3. The associated generalized eigenspace is therefore four-dimensional and can be parametrized by $\partial_t u_*$, $\partial_x u_*$, and the two period-doubling eigenfunctions $v_{\rm pd}$ and $\bar{v}_{\rm pd}$. If we denote the associated coordinates by $(\tau, \xi, v) \in S^1 \times \mathbb{R} \times \mathbb{C}$ for the temporal phase τ, the spatial translation ξ, and the complex Hopf amplitude v, then we see upon using [44, Theorem 2.9] that the vector field on the center manifold for $L \gg 1$ is given by

$$\begin{aligned} \dot\tau &= \omega_L(\tau, \xi, v, \mu), \\ \dot\xi &= g_L(\tau, \xi, v, \mu), \\ \dot v &= h_L(\tau, \xi, v, \mu), \end{aligned} \tag{3.8}$$

where

$$\omega_L(\tau, \xi, 0, 0) = \omega_* + O(e^{-\delta L}), \qquad g_L(\tau, \xi, 0, 0) = O(e^{-\delta L}), \qquad g(\tau + 2\pi, \xi, v, \mu) = g_L(\tau, \xi, v, \mu),$$
$$h_L(\tau, \xi, v, \mu) = [i\omega_*/2 + O(\mu + e^{-\delta L})]v + \gamma_L(\tau, \mu)|v|^2 v + O(|v|^5).$$

We do not know whether the center manifold exists in a uniform neighborhood of the source or whether the Taylor expansion on the center manifold converges as $L \to \infty$. If we assume that the temporal average of $\gamma_L(\tau, 0)$ is strictly negative uniformly in $L \gg 1$, then we obtain $\dot\xi \approx \alpha_L \mu$ for a constant α_L due to resonant terms of the form $g_1 e^{i\tau}\bar{v}^2$ in $g_L(\tau, \xi, v, \mu)$ as in the analysis in the introduction or in [44, section 7]. If α_L is not zero, the bifurcating 1D spirals should therefore drift on $(-L, L)$ and eventually leave the local center manifold. We believe that this drift can be followed on a global group-invariant center manifold until the effects of the boundary become of the order of the drift speed μ.

4. Period-doubling of spiral waves. The spectral analysis of planar spiral waves is in many respects analogous to that of 1D spirals. We consider the reaction-diffusion equation (2.1)

$$u_t = D\Delta u + f(u; \mu), \qquad (x, y) \in \mathbb{R}^2, \tag{4.1}$$

first on the plane, and subsequently on large disks $B_R(0)$ of radius $R \gg 1$ together with appropriate boundary conditions. An Archimedean spiral wave is a rigidly rotating solution of the form

$$u(x, y, t) = u_*(r, \varphi - \omega t), \qquad (x, y) = (r\cos\varphi, r\sin\varphi),$$

which converges to 1D wave trains $u_{\rm wt}$,

$$|u_*(r, \cdot - \omega_* t) - u_{\rm wt}(k_* r + \theta(r) + \cdot - \omega_* t)|_{C^1(S^1)} \to 0 \text{ as } r \to \infty,$$

where $k_* \neq 0$ denotes the asymptotic wave number of the wave trains, and $\theta(r)$ is a smooth phase correction with $\theta'(r) \to 0$ as $r \to \infty$. We shall assume that the essential spectrum of the asymptotic 1D wave train is simple at $\Lambda = 0$ and that the group velocity c_g of the wave train $u_{\rm wt}$, computed in the laboratory frame, is positive.

Spiral waves are equilibria in the corotating frame $\psi = \varphi - \omega_* t$, where they satisfy the elliptic system

$$D\left[u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\psi\psi}\right] + \omega_* u_\psi + f(u(r,\psi);\mu) = 0 \tag{4.2}$$

with

$$|u_*(r,\cdot) - u_{\mathrm{wt}}(k_* r + \theta(r) + \cdot)|_{C^1(S^1)} \to 0 \text{ as } r \to \infty. \tag{4.3}$$

The convergence assumed in (4.3) implies that the asymptotic shape of the spiral u_* is indeed given by the 1D wave-train solution u_{wt}, while the asymptotic wave number k_* and the temporal frequency ω_* are related via $\omega_* = \omega_{\mathrm{nl}}(k_*)$.

Next, we linearize (4.1) about the spiral wave in the corotating frame, which is equivalent to linearizing (4.2) about $u_*(r,\psi)$. The resulting operator $\mathcal{L}_*$ is given by

$$\mathcal{L}_* u = D\left[u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\psi\psi}\right] + \omega_* u_\psi + f_u(u_*(r,\psi);\mu)u,$$

which is a closed operator on $L^2(\mathbb{R}^2,\mathbb{R}^n)$. If we take the formal limit $r \to \infty$ in the eigenvalue equation

$$D\left[u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\psi\psi}\right] + \omega_* u_\psi + f_u(u_*(r,\psi);\mu)u = \Lambda u, \tag{4.4}$$

we obtain the limiting equation

$$Du_{rr} + \omega_* u_\psi + f_u(u_{\mathrm{wt}}(k_* r + \psi);\mu)u = \Lambda u \tag{4.5}$$

with 2π-periodic boundary conditions in ψ. If we set $\psi \mapsto -\omega_* t$, we recover the Floquet eigenvalue problem (2.13) of the 1D wave trains which we discussed in section 2.2. The Floquet symmetry of the Floquet eigenvalue problem (2.13) is reflected in the invariance of the asymptotic spiral eigenvalue problem (4.5) under the substitution

$$u(r,\psi) \longmapsto u(r,\psi)\mathrm{e}^{\mathrm{i}\ell\psi}, \qquad \Lambda \longmapsto \Lambda + \mathrm{i}\omega_*\ell \tag{4.6}$$

for each $\ell \in \mathbb{Z}$. We emphasize that the transformation (4.6) for the essential spectrum will generate new curves of spectrum for (4.4). The new eigenvalues are generated by the asymptotic SO(2)-symmetry of (4.4) and *not* by an artificial Floquet symmetry as for the wave trains. Indeed, the spiral wave is an equilibrium, and each Λ obtained from (4.6) belongs to a different eigenfunction.

We proved in [42] that these formal considerations can be made precise in the following sense. The operator $\mathcal{L}_* - \Lambda$ is Fredholm if and only if Λ does not belong to the Floquet spectrum of the linearized period map of the asymptotic wave trains in the laboratory frame, that is, if $\mathrm{e}^{2\pi\Lambda/\omega_*}$ is not in the spectrum of the period map of

$$u_t = Du_{xx} + f_u(u_{\mathrm{wt}}(k_* x - \omega_* t);\mu)u.$$

In particular, the essential spectrum is vertically periodic in the complex plane with period $\mathrm{i}\omega_*$. Moreover, for spirals emitting wave trains, which by definition have $c_g > 0$ at $\nu = 0$, the essential spectrum of the spiral wave in a neighborhood of the origin, and consequently all its vertical translates, moves into the left half-plane when $\mathcal{L}_*$ is posed on the spaces with exponentially weighted norms

$$|u|^2_{L^2_\eta} = \int_{\mathbb{R}^2} \left| u(x,y) \mathrm{e}^{-\eta\sqrt{x^2+y^2}} \right|^2 \mathrm{d}x\,\mathrm{d}y$$

for sufficiently small positive rates $\eta > 0$. We may therefore define the geometric and algebraic multiplicities of $\Lambda = 0$ and $\Lambda = \pm\mathrm{i}\omega_*$ as eigenvalues of $\mathcal{L}_*$ posed on L^2_η for small $\eta > 0$. On this space, $\partial_\psi u_*$ provides an eigenfunction of $\mathcal{L}_*$ with $\Lambda = 0$, while $\partial_x u_*$ and $\partial_y u_*$ generate eigenfunctions belonging to $\Lambda = \pm\mathrm{i}\omega_*$. We proved in [42] that spiral waves are robust provided these eigenvalues are algebraically simple as eigenvalues in L^2_η.

The preceding characterization of the essential spectrum of spiral waves by the spectrum of the asymptotic wave trains shows that essential spatio-temporal period-doubling of spiral waves is a robust phenomenon.

Corollary 4.1 (robust period-doubling of planar spirals). *There exists an open class of one-parameter families of reaction-diffusion systems such that the essential spectrum of $\mathcal{L}_*$ crosses the imaginary axis first at $\Lambda = \pm\mathrm{i}\omega_*/2 + \mathrm{i}\omega_*\ell$ for each $\ell \in \mathbb{Z}$.*

In preparation for a discussion of the spectra of spirals under truncation to large bounded disks, we record that the absolute spectrum of (4.4) coincides with the absolute spectrum of the asymptotic wave trains, computed in the laboratory frame, which is again vertically periodic in the complex plane with period $\mathrm{i}\omega_*$. In [42], we showed that for each $\Lambda \notin \Sigma_{\mathrm{abs}}$ there exists an exponential weight η such that $\mathcal{L}_* - \Lambda$ is Fredholm with index zero on the space L^2_η. We define the extended point spectrum as the set of $\Lambda \notin \Sigma_{\mathrm{abs}}$ for which the kernel of $\mathcal{L}_* - \Lambda$ is nontrivial on L^2_η, with η chosen as above. It is not difficult to see that the kernel does not depend on the choice of the weight [42].

The persistence of period-doubling on large bounded disks of radius $R \gg 1$ with Neumann boundary conditions is now very similar in spirit to the situation in one space dimension. First, we address the persistence of the spiral wave on disks $B_R(0)$ for $R \gg 1$. Similarly to Theorem 1, we assume the existence of a 1D boundary sink that connects the asymptotic wave trains at $x = -\infty$ with Neumann conditions at $x = 0$ such that $\Lambda = 0$ does not belong to its extended point spectrum. Moreover, we assume robustness of the spiral on the plane; that is, we require that $\Lambda = 0$ is algebraically simple in L^2_η for $\eta > 0$ small. Under these conditions, the spiral wave persists as a rigidly rotating solution of the reaction-diffusion system for all sufficiently large R [42].

Next, we consider the spectrum of the linearization (4.4) about the truncated spiral wave on $B_R(0)$ with Neumann conditions at $r = R$, for which a result completely analogous to Theorem 3 holds.

Theorem 7 (see [42]). *Assume that the extended point spectrum of the spiral wave is discrete; then the spectrum of the truncated 2D spiral converges locally uniformly in the symmetric Hausdorff distance to the disjoint union of the absolute spectrum Σ_{abs} of the asymptotic wave trains, computed in the laboratory frame, and a discrete set of isolated eigenvalues with finite multiplicity. Convergence to the absolute spectrum is algebraic of order $\mathrm{O}(1/R)$, and*

the number of eigenvalues inside any small disk that contains a point in the absolute spectrum converges to infinity as $R \to \infty$. The discrete part of the limiting spectrum is the union of the extended point spectrum of the spiral and the extended point spectrum of the boundary sink on $\mathbb{R}^-$. Convergence toward elements of the discrete part of the limiting spectrum is exponential in R, and the multiplicity of eigenvalues in a sufficiently small disk about an element of the extended point spectra converges to the sum of the multiplicities in the extended point spectra of planar spiral and boundary sink.

Remark 4.2. *We remark that it has recently been shown numerically* [45] *and analytically* [41] *that infinitely many discrete eigenvalues in the extended point spectrum of spiral waves can accumulate at edges of the absolute spectrum.*

Corollary 4.3 (absolute period-doubling on bounded domains). *Resonant crossing of eigenvalues at $\Lambda = \pm i\omega_*/2 + i\omega_*\ell + O(1/R)$ with $\ell \in \mathbb{Z}$ of spiral waves on disks of radius $R \gg 1$ occurs in an open subset of one-parameter families of reaction-diffusion systems.*

An interesting feature of period-doubling of spiral waves is the shape of its period-doubling eigenfunctions. We first describe the shape of eigenfunctions for the essential spectrum (although this is somewhat irrelevant to bifurcations taking place on large disks). Consider the eigenvalue problem (2.8) of the wave train $u_{\rm wt}$ in the 1D comoving frame and assume that it has period-doubling eigenvalues given by

$$\lambda(\nu) = -(c_{\rm g}^{\rm pd} - c_{\rm p})(\nu - ik_*/2) + d(\nu - ik_*/2)^2 + O(|\nu - ik_*/2|^3), \qquad c_{\rm p} = \frac{\omega_*}{k_*},$$

for $\nu \approx ik_*/2$ with associated eigenfunctions given by

$$v(y) = e^{\nu y} u_{\rm pd}(k_* y; \nu), \qquad u_{\rm pd}(k_* y; \nu) = u_{\rm pd}(k_* y + 2\pi; \nu) \quad \forall y$$

when written in the variable $y = \xi/k_*$. In the laboratory frame $x = y + c_{\rm p} t$, we obtain the critical dispersion curve

$$\Lambda(\nu) = -\frac{i\omega_*}{2} - c_{\rm g}^{\rm pd}(\nu - ik_*/2) + d(\nu - ik_*/2)^2 + O(|\nu - ik_*/2|^3) \tag{4.7}$$

with eigenfunctions

$$v(x,t) = e^{\Lambda(\nu)t} e^{\nu x} u_{\rm pd}(k_*(x - c_{\rm p} t); \nu).$$

The eigenfunction for the spiral is now obtained as in [35] by substituting $t = -\psi/\omega_*$ and $x = r$, which gives

$$u(r,\psi) = e^{-\Lambda(\nu)\psi/\omega_*} e^{\nu r} u_{\rm pd}(k_* r + \psi; \nu)$$

for the solution of (4.4). We evaluate this expression at the critical wave number $\nu = ik_*/2$ to get

$$u_0(r,\psi) = e^{i(k_* r - \psi)/2} u_{\rm pd}(k_* r + \psi) = e^{-i\psi/2} v_{\rm pd}(k_* r + \psi),$$

where we substituted the real-valued function

$$v_{\rm pd}(\xi) := e^{i\xi/2} u_{\rm pd}(\xi)$$

with $v_{\rm pd}(\xi + 2\pi) = -v_{\rm pd}(\xi)$ for all ξ, which corresponds to the period-doubling solution of (2.5). Exploiting the Floquet symmetry (4.6), we find the additional eigenfunctions

$$u_\ell(r,\psi) = e^{i\psi(\ell - 1/2)} v_{\rm pd}(k_* r + \psi) \tag{4.8}$$

belonging to $\Lambda = -\mathrm{i}\omega_*/2 + \mathrm{i}\omega_*\ell$ for $\ell \in \mathbb{Z}$, and in particular the complex conjugate

$$u_1(r,\psi) = \mathrm{e}^{\mathrm{i}\psi/2} v_{\mathrm{pd}}(k_* r + \psi)$$

of $u_0(r,\psi)$. To get real-valued solutions, we add up u_0 and u_1 and solve the time-dependent linearized problem with initial data $u_0 + u_1$ to get

$$u(r,\psi,t) = \mathrm{e}^{-\mathrm{i}\omega_* t/2}\mathrm{e}^{-\mathrm{i}\psi/2} v_{\mathrm{pd}}(k_* r+\psi) + \mathrm{e}^{\mathrm{i}\omega_* t/2}\mathrm{e}^{\mathrm{i}\psi/2} v_{\mathrm{pd}}(k_* r+\psi) = \cos\left(\frac{\psi+\omega_* t}{2}\right) v_{\mathrm{pd}}(k_* r+\psi).$$

In the laboratory frame $\varphi = \psi + \omega_* t$, we finally obtain the real perturbation

$$u(r,\varphi,t) = \cos\left(\frac{\varphi}{2}\right) v_{\mathrm{pd}}(k_* r + \varphi - \omega_* t). \tag{4.9}$$

If we formally add the solution (4.9) multiplied by a small amplitude $\sqrt{\epsilon}$ to the original spiral wave, we obtain

$$u_*(r,\varphi-\omega_* t) + \sqrt{\epsilon}\cos\left(\frac{\varphi}{2}\right) v_{\mathrm{pd}}(k_* r + \varphi - \omega_* t) \tag{4.10}$$

in the spiral far field. In particular, the amplitude of the period-doubling mode vanishes along the stationary line $\varphi = \pi$, while the spiral is rotating. The temporal frequency of the perturbation is $\omega_*/2$ since v_{pd} has period 4π. The pattern described by (4.10) looks exactly like those observed experimentally in [28, 29, 47], numerically in [14], and here in Figure 1(ii).

Although this computation is formal, the shape of eigenfunctions resulting from the absolute spectrum on large bounded domains can be computed similarly. Assume therefore that $|c_{\mathrm{g}}^{\mathrm{pd}}| \ll 1$, as is the case, for instance, near spatially homogeneous oscillations. In this case, the absolute spectrum has a branch point Λ_{bp} close to the tip of the period-doubling instability at $\Lambda = \mathrm{i}\omega_*/2 + \mu$. This branch point corresponds to a root of the equation $\mathrm{d}\Lambda/\mathrm{d}\nu = 0$, with $\Lambda(\nu)$ as in (4.7), and is therefore given by

$$\Lambda_{\mathrm{bp}} = \frac{\mathrm{i}\omega_*}{2} - \frac{[c_{\mathrm{g}}^{\mathrm{pd}}]^2}{4} + \mu \qquad \text{with} \qquad \nu_{\mathrm{bp}} = \frac{c_{\mathrm{g}}^{\mathrm{pd}}}{2d}.$$

Following the above computation gives

$$\cos(\varphi/2)\mathrm{e}^{\nu_{\mathrm{bp}} r} v_{\mathrm{pd}}(k_* r + \varphi - \omega_* t)$$

for the perturbation of the primary spiral-wave profile. In particular, we observe the stationary line of vanishing amplitude for the period-doubling mode, and in addition an exponential decay or growth of the eigenfunction depending on whether the group velocity of the period-doubling mode is negative or positive, respectively.

At the onset of the absolute instability on large disks, there are five eigenvalues in the vicinity of the imaginary axis, namely, $\Lambda = 0$ induced by rotation, $\Lambda = \pm\mathrm{i}\omega_* + \mathrm{O}(\mathrm{e}^{-\delta R})$ induced by translation, and $\Lambda_{\mathrm{pd}} = \pm\mathrm{i}\omega_*/2 + \mathrm{O}(1/R^2)$ near the branch point of the absolute spectrum that induces the period-doubling of the wave trains. In [43], we showed that resonant Hopf bifurcations of this type will typically lead to a slow drift of the spiral wave with drift

speed $O(\mu)$ (see also the discussion in section 1). Based on this prediction, we verified that drift indeed occurs in the Rössler system and report on these computations in section 6. Independently, drift was also observed numerically in [4].

We remark that the region of validity of our drift analysis is very small in parameter space since the eigenvalue at the edge of the absolute spectrum is $O(1/R^2)$-close to other eigenvalues that subsequently cross the imaginary axis.

Last, we comment on the role played by the other eigenfunctions $u_\ell(r,\psi)$ given in (4.8). Proceeding as above, we see that the sum of the eigenfunctions $u_{\ell+1}$ and $u_{-\ell}$ for positive integers ℓ generates patterns of the form

$$u_*(r,\varphi-\omega_* t)+\sqrt{\epsilon}\cos\left(\frac{(2\ell+1)\varphi}{2}\right)v_{\mathrm{pd}}(k_* r+\varphi-\omega_* t), \tag{4.11}$$

which exhibit $2\ell+1$ stationary line defects at $\varphi=\frac{2n+1}{2\ell+1}\pi$ for $n=0,\dots,2\ell$. Interestingly, none of the associated eigenvalues at $\Lambda=\mathrm{i}\omega_*(\ell+1/2)$ affects the expected drift in any way as the resulting Hopf frequencies $\omega_{\mathrm{H}}:=(\ell+1/2)\omega_*$ cannot satisfy the required resonance condition (1.2), except when $\ell=0$, which is therefore solely responsible for the occurrence of drift.

5. Defects near period-doubling of homogeneous oscillations.

In this section, we study patterns that are created near the onset of period-doubling of a family of wave trains. We restrict ourselves to the onset of period-doubling at homogeneous oscillations where $k\approx 0$ and therefore $c_{\mathrm{g}}=c_{\mathrm{g}}^{\mathrm{pd}}=0$.

5.1. Derivation of amplitude equations.

We shall assume that $u_0(\tau)$ has minimal period 2π and satisfies

$$\omega_0 u_\tau = f(u;\mu) \tag{5.1}$$

for $\mu=0$ and some $\omega_0\neq 0$. Furthermore, we assume that $\rho=1$ and $\rho=-1$ are geometrically and algebraically simple Floquet multipliers of the linearization

$$\omega_0 u_\tau = f_u(u_0(\tau);0)u$$

of (5.1) about $u_0(\tau)$; the associated nontrivial solutions of the linearization are given by $u_0'(\tau)$ and $u_{\mathrm{pd}}(\tau)$, respectively. The associated solutions to the adjoint equation

$$\omega_0 w_\tau = -f_u(u_0(\tau);0)^* w$$

will be denoted by $\psi_0(\tau)$ and $\psi_{\mathrm{pd}}(\tau)$, respectively.

Simplicity of $\rho=1$ implies that the periodic orbit $u_0(\tau)$ persists for all μ close to zero with temporal frequency $\omega=\omega_0(\mu)$, and we assume that the unique Floquet multiplier $\rho_{\mathrm{pd}}(\mu)$ near $\rho=-1$ of the persisting wave train satisfies $\rho_{\mathrm{pd}}'(0)<0$. The simplicity of $\rho=1$ also implies that the partial differential equation (PDE)

$$u_t = Du_{xx}+f(u;\mu) \tag{5.2}$$

with $\mu=0$ has a one-parameter family of traveling waves $u(x,t)=u_0(\omega t-kx;k)$, defined for $|k|\ll 1$, near u_0, where $\omega=\omega_{\mathrm{nl}}(k)$ with $\omega_{\mathrm{nl}}(0)=\omega_0$ is a smooth, even function of k [38,

section 3.3]. We assume that the nonlinear dispersion relation $\omega_{\rm nl}(k)$ is nondegenerate so that $\omega_{\rm nl}''(0) \neq 0$. Last, the linearization

$$u_t = Du_{xx} + f_u(u_0(\omega_0 t); 0)u$$

of (5.2) about $u_0(\omega_0 t)$ can be reduced, via spatial Fourier transform, to the ODE

$$u_t = [D\nu^2 + f_u(u_0(\omega_0 t); 0)]u. \tag{5.3}$$

The simplicity of the multipliers $\rho = \pm 1$ implies that (5.3) has unique Floquet exponents, given by $\lambda_0 = d_0\nu^2 + \mathrm{O}(\nu^4)$ and $\lambda_{\rm pd} = \pi\mathrm{i} + d_1\nu^2 + \mathrm{O}(\nu^4)$ for appropriate constants $d_0, d_1 \in \mathbb{R}$, for $|\nu| \ll 1$, which correspond to $\rho = \pm 1$. We assume that $d_0, d_1 > 0$.

We are interested in coherent structures near the homogeneous oscillations. Thus, for ω close to ω_0, we introduce the new time variable $\tau = \omega t$ and seek solutions $u(x, \tau)$ of the PDE

$$\omega u_\tau = Du_{xx} + f(u; \mu) \tag{5.4}$$

that are 4π-periodic in τ.

Theorem 8. *Under the above hypotheses, the following is true for all μ sufficiently close to zero and ω close to ω_0: Solutions $u(x, \tau)$ of (5.4) with period 4π in τ whose time slices $u(x, \cdot)$ are, for each $x \in \mathbb{R}$, close to an appropriate τ-translate of $u_0(\cdot)$ are in one-to-one correspondence with small bounded solutions of the ODE*

$$\begin{aligned}
\phi_x &= \kappa,\\
\kappa_x &= \frac{1}{d_0}\left[-\bar\omega + \frac{1}{2}\omega_{\rm nl}''(0)\kappa^2 + b_0A^2\right] + \mathrm{O}(|A|^3 + |\kappa|^3 + B^2 + \bar\omega^2),\\
A_x &= B,\\
B_x &= \frac{1}{d_1}\left[(-\rho_{\rm pd}'(0)\mu + b_1\bar\omega + b_2\kappa^2 + b_3A^2)A + b_4\kappa B\right]\\
&\quad + \mathrm{O}(|A|(A^4 + A^2|\kappa| + \mu^2 + \bar\omega^2) + |B|\kappa^2 + B^2(|\kappa| + |A|)),
\end{aligned} \tag{5.5}$$

where $\omega = \omega_0(\mu) + \bar\omega$. The right-hand side of (5.5) does not depend on ϕ and is equivariant under the reflection $(\phi, \kappa, A, B) \mapsto (\phi, \kappa, -A, -B)$, which corresponds to the time shift by 2π, and reversible under $x \mapsto -x$ with reverser $(\phi, \kappa, A, B) \mapsto (\phi, -\kappa, A, -B)$. The solution of (5.4) associated with a solution (ϕ, κ, A, B) of (5.5) has temporal period 2π if and only if $(A, B) = 0$.

Equation (5.5) is the steady-state equation associated with the formal amplitude equation[1]

$$\begin{aligned}
\phi_t &= d_0\phi_{xx} - \frac{1}{2}\omega_{\rm nl}''(0)\phi_x^2 - b_0A^2,\\
A_t &= d_1A_{xx} + \left[\rho_{\rm pd}'(0)\mu - \hat b_1\phi_{xx} - \hat b_2\phi_x^2 - \hat b_3A^2)\right]A - b_4\phi_xA_x
\end{aligned} \tag{5.6}$$

[1]The coefficients $\hat b_j$ can be obtained from the b_j's upon solving the equation for κ in (5.5) for $\bar\omega$ and substituting into the equation for A.

for the phase ϕ and the period-doubling mode A. A similar complex version of (5.6) has been analyzed in [9, 10, 11], where it was derived from a combustion model using formal multiscale expansions to describe the interaction of Burgers and Hopf modes. We also refer the reader to [22] for the derivation of other amplitude equations for systems with conservation laws.

Before embarking on the proof of the preceding theorem, we consider spectral PDE stability of the bounded solutions $u_*(x,\tau)$ of (5.4) described by Theorem 8. A complex number λ is a Floquet exponent of $u_*(x,\tau)$ if and only if there exists a nontrivial 4π-periodic solution $u(x,\tau)$ of

$$\lambda u + \omega u_\tau = D u_{xx} + f_u(u_*(x,\tau);\mu)u. \tag{5.7}$$

Floquet exponents of $u_*(x,\tau)$ near the origin are captured by the following result.

Theorem 9. *Under the hypotheses of Theorem 8, assume that $U_*(x) = (\kappa_*, A_*, B_*)(x)$ is a small bounded solution of (5.5) corresponding to a 4π-periodic solution $u_*(x,\tau)$ of (5.4). If we write (5.5) as*

$$\begin{pmatrix} d_0\phi_{xx} \\ d_1 A_{xx} \end{pmatrix} = G(\phi_x, A, A_x, \mu, \bar\omega),$$

then Floquet exponents λ of (5.7) near the origin are in one-to-one correspondence, counting multiplicity, with solutions λ near the origin of the reduced PDE eigenvalue problem

(5.8)

$$\begin{pmatrix} d_0\phi_{xx} \\ d_1 A_{xx} \end{pmatrix} = \mathrm{D}_{(\kappa,A,B)} G(\kappa_*(x), A_*(x), B_*(x), \mu, \bar\omega) \begin{pmatrix} \phi_x \\ A \\ A_x \end{pmatrix} + \lambda \begin{pmatrix} 1 & 0 \\ b_1 A_*(x) & 1 \end{pmatrix} \begin{pmatrix} \phi \\ A \end{pmatrix}$$
$$+ \lambda \left[\mathrm{O}(|\lambda| + |\mu| + |\bar\omega|) \begin{pmatrix} \phi \\ \phi_x \\ A \\ A_x \end{pmatrix} + \mathrm{O}(\|U_*\|) \begin{pmatrix} \phi_x \\ A \\ A_x \end{pmatrix} + \begin{pmatrix} \mathrm{O}(\|U_*\|) \\ \mathrm{O}(\|\kappa_*\|^2 + \|A_*\|^2 + \|B_*\|) \end{pmatrix} \Phi \right].$$

Furthermore, $u_*(x,\tau)$ does not have any Floquet exponents in the right half-plane other than those captured by (5.8) (or those obtained from the trivial Floquet symmetry).

Proof of Theorems 8 and 9. We proceed using spatial dynamics as in [6, section 8.1] and therefore write (5.4) as

$$\begin{aligned} u_x &= v, \\ v_x &= D^{-1}[\omega u_\tau - f(u;\mu)] \end{aligned} \tag{5.9}$$

on the space $\mathcal{X} := H^1_{\mathrm{per}}(0,4\pi) \times H^{1/2}_{\mathrm{per}}(0,4\pi)$. Thus, we regard (5.9) as a dynamical system in the spatial evolution variable x, acting on 4π-periodic functions $\mathbf{u} = (u,v) \in \mathcal{X}$ of the rescaled temporal variable τ. Important features of (5.9) are its equivariance under the shifts

$$\mathcal{S}_\phi : \quad \mathcal{X} \longrightarrow \mathcal{X}, \quad \mathbf{u}(\cdot) \longmapsto \mathbf{u}(\cdot - \phi)$$

for each fixed $\phi \in [0,4\pi]/\!\sim$ and reversibility in x with reverser $\mathcal{R} : (u,v) \mapsto (u,-v)$.

Equation (5.9) has, for $\mu = 0$ and $\omega = \omega_0$, an S^1-group orbit of stationary solutions given by $\mathcal{S}_\phi \mathbf{u}_0$, where

$$\mathbf{u}_0 := \begin{pmatrix} u_0 \\ 0 \end{pmatrix}.$$

Each of these solutions has isotropy $\mathbb{Z}_2$ generated by $\mathcal{S}_{2\pi}$. We first concentrate on a neighborhood of $\mathbf{u}_0$ and write $\mathbf{u} = \mathbf{u}_0 + \mathbf{v}$ so that $\mathbf{v} \in \mathcal{X}$ satisfies

$$\mathbf{v}_x = \mathcal{B}_0 \mathbf{v} + (\omega - \omega_0)\mathcal{N}(\mathbf{u}_0 + \mathbf{v}) + \mathcal{G}(\mathbf{v}; \mu) \tag{5.10}$$

with

$$\mathcal{B}_0 = \begin{pmatrix} 0 & 1 \\ D^{-1}[\omega_0 \partial_\tau - f_u(u_0(\cdot); 0)] & 0 \end{pmatrix}, \qquad \mathcal{N} = \begin{pmatrix} 0 & 0 \\ D^{-1}\partial_\tau & 0 \end{pmatrix},$$
$$\mathcal{G}(\mathbf{v}; \mu) = \begin{pmatrix} 0 \\ -D^{-1}[f(u + u_0(\cdot); \mu) - f(u_0(\cdot); 0) - f_u(u_0(\cdot); 0)u] \end{pmatrix}$$

for $\mathbf{v} = (u, v)$. This is the system considered in [6, section 8.1]. Here, we have the additional simplification that both the wave number k_0 and the group velocity c_g vanish. As in [6, section 8.1], the operator $\mathcal{B}_0$ is closed and densely defined on $\mathcal{X}$ and has only discrete spectrum. Exploiting our hypotheses, we see that $\mathcal{B}_0$ has, in contrast to [6, section 8.1], two geometrically simple eigenvalues at $\nu = 0$ with eigenfunctions $(u_0', 0)$ and $(u_{\mathrm{pd}}, 0)$ (compared with a unique geometrically simple eigenvalue in [6, section 8.1]). Each of these eigenvalues has algebraic multiplicity two with generalized eigenfunctions given by $(0, u_0')$ and $(0, u_{\mathrm{pd}})$, respectively. The associated eigenfunctions of the adjoint operator $\mathcal{B}_0^*$ are given by

$$\psi_0 = \begin{pmatrix} 0 \\ -D\psi_0 \end{pmatrix}, \qquad \psi_1 = \begin{pmatrix} -D\psi_0 \\ 0 \end{pmatrix}, \qquad \psi_0^{\mathrm{pd}} = \begin{pmatrix} 0 \\ D\psi_{\mathrm{pd}} \end{pmatrix}, \qquad \psi_1^{\mathrm{pd}} = \begin{pmatrix} D\psi_{\mathrm{pd}} \\ 0 \end{pmatrix}, \tag{5.11}$$

where ψ_0 and ψ_{pd} have been defined at the beginning of section 5.1. The remaining spectrum of $\mathcal{B}_0$ on $\mathcal{X}$ is bounded away from the imaginary axis.

Using spatial center-manifold theory as in [6, section 8.1], we conclude that there exists a four-dimensional center manifold associated with (5.9) which contains all solutions of (5.9) that stay near the S^1-orbit $\{\mathcal{S}_\phi \mathbf{u}_0;\ \phi \in [0, 4\pi]/\!\sim\}$ of equilibria for all x. The center manifold can be constructed so that it is invariant under the shifts $\mathcal{S}_\theta$ and the reverser $\mathcal{R}$. In particular, the vector field on the center manifold is reversible and equivariant under shifts. Upon inspecting the operator $\mathcal{B}_0$ and exploiting the invariance under shifts in τ, we find that the center manifold can be parametrized by the coordinates (ϕ, κ, A, B) via

$$\mathbf{u} = \mathcal{S}_\phi \left[\begin{pmatrix} u_0(\mu) \\ 0 \end{pmatrix} - \kappa \begin{pmatrix} 0 \\ u_0'(\mu) \end{pmatrix} + A \begin{pmatrix} u_{\mathrm{pd}}(\mu) \\ 0 \end{pmatrix} + B \begin{pmatrix} 0 \\ u_{\mathrm{pd}}(\mu) \end{pmatrix} + \bar{\omega} \begin{pmatrix} u_\omega(\mu) \\ 0 \end{pmatrix} + \mathcal{H}_0(\kappa, A, B, \mu, \bar{\omega}) \right], \tag{5.12}$$

where $u_0(\mu)$ denotes the μ-dependent spatially homogeneous oscillation, where $u_{\mathrm{pd}}(\mu)$ is the μ-dependent eigenmode associated with $u_0(\mu)$ which causes period doubling at $\mu = 0$, and where we use the parameter $\bar{\omega} := \omega - \omega_0(\mu)$. The function u_ω is the unique 2π-periodic solution of the system

$$[\omega_0 \partial_\tau - f_u(u_0(\tau; \mu); \mu)]u_\omega = -u_0' + \frac{\langle \psi_0, u_0' \rangle_{L^2(0,2\pi)}}{\langle \psi_0, Du_0' \rangle_{L^2(0,2\pi)}} Du_0' \tag{5.13}$$

with $\langle u_\omega, u_0' \rangle_{L^2(0,2\pi)} = 0$. The function $\mathcal{H}_0$ is smooth and takes values in the generalized hyperbolic eigenspace E^{h} of $\mathcal{B}_0$, and its derivative with respect to each of its arguments vanishes at the origin $(\kappa, A, B, \mu, \bar{\omega}) = 0$. Indeed, in these coordinates, the shifts $\mathcal{S}_\theta$ are represented by

$$\mathcal{S}_\theta : (\phi, \kappa, A, B) \longmapsto (\phi + \theta, \kappa, A, B),$$

and equivariance implies that the reduced vector field, and the center-manifold parametrization $\mathcal{H}_0$, can therefore not depend on ϕ, as claimed. The isotropy group generated by $\mathcal{S}_{2\pi}$ and the reverser $\mathcal{R}$ are represented by

$$\mathcal{S}_{2\pi} : (\phi, \kappa, A, B) \longmapsto (\phi, \kappa, -A, -B), \qquad \mathcal{R} : (\phi, \kappa, A, B) \longmapsto (\phi, -\kappa, A, -B).$$

In particular, the reduced vector field will be equivariant under $\mathcal{S}_{2\pi}$ and reversible under $\mathcal{R}$. We shall now argue that the vector field for (ϕ, κ, A, B) is necessarily of the form

$$\begin{aligned}
\phi_x &= \kappa + g_1(\kappa, A, B, \mu, \bar{\omega}),\\
\kappa_x &= \frac{1}{d_0}\left[-\bar{\omega} + \frac{1}{2}\omega_{\mathrm{nl}}''(0)\kappa^2 + b_0 A^2\right] + \mathrm{O}(|A|^3 + |\kappa|^3 + B^2 + \bar{\omega}^2),\\
A_x &= B + g_2(\kappa, A, B, \mu, \bar{\omega}),\\
B_x &= \frac{1}{d_1}\left[(-\rho_{\mathrm{pd}}'(0)\mu + b_1\bar{\omega} + b_2\kappa^2 + b_3 A^2)A + b_4\kappa B\right]\\
&\quad + \mathrm{O}(|A|(A^4 + A^2|\kappa| + \mu^2 + \bar{\omega}^2) + |B|\kappa^2 + B^2(|\kappa| + |A|))
\end{aligned} \tag{5.14}$$

for appropriate constants $b_j \in \mathbb{R}$, where the functions g_1 and g_2 are smooth and respect the symmetries and the reverser, and their first two derivatives vanish at $(\kappa, A, B, \mu, \bar{\omega}) = 0$. Indeed, the linear terms in (5.14) can be computed as in [6, section 8.1] by substituting (5.12) into (5.10) and projecting using the adjoint eigenfunctions (5.11). In particular, the coefficient b_1 is given by

$$b_1 = \frac{\langle \psi_{\mathrm{pd}}, \partial_\tau u_{\mathrm{pd}} - f_{uu}(u_0; 0)[u_\omega, u_{\mathrm{pd}}] \rangle_{L^2(0,4\pi)}}{\langle \psi_{\mathrm{pd}}, u_{\mathrm{pd}} \rangle_{L^2(0,4\pi)}}. \tag{5.15}$$

The functions g_1 and g_2 must vanish to second order due to the facts that the diagonal of $\mathcal{B}_0$ vanishes and the nonlinearity appears only in the v-component. Checking compatibility of the monomial terms with the involution $\mathcal{S}_{2\pi}$ and the reverser $\mathcal{R}$, we find that the equations for κ and B must be of the specified form. The special form of the equation for κ when $A = B = 0$ is a consequence of [6, section 8.1] and our choice of $\bar{\omega}$ as the offset from the μ-dependent temporal frequency of spatially homogeneous oscillations. Last, to bring (5.14) into the form (5.5), we introduce the new coordinates

$$\tilde{\kappa} = \kappa + g_1(\kappa, A, B, \mu, \bar{\omega}), \qquad \tilde{B} = B + g_2(\kappa, A, B, \mu, \bar{\omega}).$$

Upon dropping the tildes, we arrive at (5.5) as claimed, which completes the proof of Theorem 8.

To prove Theorem 9, we record that the solutions described by Theorem 8 are uniformly close to the homogeneous oscillations $u_0(\tau)$ whose Floquet multipliers are contained strictly

inside the unit disk with the exception of multipliers close to $\rho = \pm 1$. Floquet multipliers near $\rho = \pm 1$ can be captured by a spatial center-manifold reduction for the eigenvalue problem

$$\mathbf{v}_x = [\mathcal{B}_0 + \bar{\omega}\mathcal{N} + \mathrm{D}\mathcal{G}(\mathbf{u};\mu)]\,\mathbf{v} + \lambda \begin{pmatrix} 0 & 0 \\ D^{-1}\partial_\tau & 0 \end{pmatrix} \mathbf{v}, \tag{5.16}$$

which is carried out simultaneously with the reduction for the existence problem (5.10). Following the same strategy as above, we find that the center manifold for the eigenvalue problem (5.16) about a solution $\mathbf{u}$ from (5.12) corresponding to a solution (κ_*, A_*, B_*) of (5.5) is parametrized by

$$\begin{aligned} \mathbf{v} = &\left[-\begin{pmatrix} u_0' \\ 0 \end{pmatrix} - \kappa_* \begin{pmatrix} 0 \\ u_0'' \end{pmatrix} + A_* \begin{pmatrix} u_{\mathrm{pd}}' \\ 0 \end{pmatrix} + B_* \begin{pmatrix} 0 \\ u_{\mathrm{pd}}' \end{pmatrix} + \partial_\tau \mathcal{H}_0(\kappa_*, A_*, B_*, \mu, \bar{\omega}) \right] \Phi \\ &- \kappa \begin{pmatrix} 0 \\ u_0' \end{pmatrix} + A \begin{pmatrix} u_{\mathrm{pd}} \\ 0 \end{pmatrix} + B \begin{pmatrix} 0 \\ u_{\mathrm{pd}} \end{pmatrix} + \mathrm{D}_{(\kappa,A,B)} \mathcal{H}_0(\kappa_*, A_*, B_*, \mu, \bar{\omega}) \begin{pmatrix} \kappa \\ A \\ B \end{pmatrix} \\ &+ \lambda \left[\begin{pmatrix} u_\omega(\mu) \\ 0 \end{pmatrix} \Phi + \mathcal{H}_{10} \begin{pmatrix} \kappa \\ A \\ B \end{pmatrix} + \mathcal{H}_{11}(\kappa_*, A_*, B_*, \mu, \bar{\omega}, \lambda) \begin{pmatrix} \Phi \\ \kappa \\ A \\ B \end{pmatrix} \right], \end{aligned} \tag{5.17}$$

where $\mathcal{H}_{10}$ and $\mathcal{H}_{11}$ map into the hyperbolic eigenspace E^{h} of $\mathcal{B}_0$ and where $H_{11}(0) = 0$. In particular, for $\lambda = 0$, we obtain precisely the linearization of the reduced vector field about (κ_*, A_*, B_*), and it remains to calculate the λ-dependent terms. Using that $\mathcal{H}_{10}$ maps into E^{h} and that $(u_\omega, 0) \in E^{\mathrm{h}}$ by construction (5.13), we obtain the desired expression (5.8) upon substituting (5.17) into (5.7) and projecting using the adjoint eigenfunctions (5.11). We emphasize that the coefficient in front of the $\lambda A_* \Phi$ term is equal to b_1 as computed in (5.15). ■

The coefficients appearing in (5.5), and, in particular, the coefficient b_1, are in general nonzero. It will often be more convenient to express the term $\bar{\omega}A$ in terms of $\kappa_x A$. Thus, we write (5.5) as

$$\begin{aligned} \phi_x &= \kappa, \\ \kappa_x &= \frac{1}{d_0}\left[-\bar{\omega} + \frac{1}{2}\omega_{\mathrm{nl}}''(0)\kappa^2 + b_0 A^2 \right] + \mathrm{O}(|A|^3 + |\kappa|^3 + B^2 + \bar{\omega}^2), \\ A_x &= B, \\ B_x &= \left(-\frac{\rho_{\mathrm{pd}}'(0)}{d_1}\mu + \tilde{b}_1\kappa_x + \tilde{b}_2\kappa^2 + \tilde{b}_3 A^2 \right) A + \tilde{b}_4 \kappa B \\ &\quad + \mathrm{O}(|A|(A^4 + A^2|\kappa| + \mu^2 + \bar{\omega}^2) + |B|\kappa^2 + B^2(|\kappa| + |A|)), \end{aligned} \tag{5.18}$$

where

$$\tilde{b}_1 := -\frac{d_0 b_1}{d_1}, \qquad \tilde{b}_2 := \frac{b_2 + \frac{1}{2}\omega_{\mathrm{nl}}''(0) b_1}{d_1}, \qquad \tilde{b}_3 := \frac{b_3 + b_0 b_1}{d_1}, \qquad \tilde{b}_4 := \frac{b_4}{d_1}.$$

The coefficients appearing in (5.18) have the following interpretation: b_0 describes how the temporal frequency of period-doubled spatially homogeneous oscillations changes with the amplitude of the period-doubling mode. The coefficient $\tilde{b}_2$ encodes the wave number dependence of the onset of period-doubling, and $\tilde{b}_3$ reflects whether the period-doubling bifurcation is subcritical or supercritical. Last, $\tilde{b}_4$ gives the dependence of the linear group velocity of the period-doubling mode on the wave number of the underlying wave train. We shall assume that the period-doubling bifurcation is supercritical and that the homogeneous oscillations destabilize before the wave trains with nonzero wave number.

Hypothesis 1. *We assume that* $\tilde{b}_2 > 0$ *and* $\tilde{b}_3 > 0$.

Since we already assumed that $\omega''_{\mathrm{nl}}(0) \neq 0$, we can arrange to have $\omega''_{\mathrm{nl}}(0) > 0$, possibly after replacing κ by $-\kappa$. Using this normalization together with $\tilde{b}_3 > 0$, an appropriate change of the parameters and the dependent and independent variables transforms (5.18) into

$$\begin{aligned} \kappa_x &= -\bar{\omega} + \kappa^2 + bA^2 + \mathrm{O}(|A|^3 + |\kappa|^3 + B^2 + \bar{\omega}^2), \\ A_x &= B, \\ B_x &= [-\mu + a\kappa_x + d\kappa^2 + A^2]A + c\kappa B \\ &\quad + \mathrm{O}(|A|(A^4 + A^2|\kappa| + \mu^2 + \bar{\omega}^2) + |B|\kappa^2 + B^2(|\kappa| + |A|)), \end{aligned} \tag{5.19}$$

where we use the same letters for the new transformed variables and omit the equation for ϕ as it decouples. Hypothesis 1 translates into $d > 0$.

The long wavelength scaling

$$(\kappa, A, B, \mu, \bar{\omega}, x) \longrightarrow \left(\epsilon\kappa, \epsilon A, \epsilon^2 B, \epsilon^2 \mu, \epsilon^2 \Omega, \frac{x}{\epsilon}\right) \tag{5.20}$$

transforms (5.19) into

$$\begin{aligned} \kappa_x &= -\Omega + \kappa^2 + bA^2 + \mathrm{O}(\epsilon), \\ A_x &= B, \\ B_x &= [-\mu + a\kappa_x + d\kappa^2 + A^2]A + c\kappa B + \mathrm{O}(\epsilon). \end{aligned} \tag{5.21}$$

This system is equivariant under the reflection $\mathcal{S}_{2\pi} : (\kappa, A, B) \mapsto (\kappa, -A, -B)$ and reversible with reverser

$$\mathcal{R} : (\kappa, A, B) \longmapsto (-\kappa, A, -B).$$

We set $\epsilon = 0$ in the following and focus on the resulting system

$$\begin{aligned} \kappa_x &= -\Omega + \kappa^2 + bA^2, \\ A_x &= B, \\ B_x &= [-\mu + a\kappa_x + d\kappa^2 + A^2]A + c\kappa B \end{aligned} \tag{5.22}$$

or, alternatively, upon substituting the equation for κ_x into the last equation, on

$$\begin{aligned} \kappa_x &= -\Omega + \kappa^2 + bA^2, \\ A_x &= B, \\ B_x &= [-(\mu + a\Omega) + (a + d)\kappa^2 + (1 + ab)A^2]A + c\kappa B. \end{aligned} \tag{5.23}$$

The reversibility of the full problem (5.21) will allow us to show persistence of the solutions of (5.22) that we shall construct below for $\epsilon > 0$.

Last, we discuss the PDE stability of bounded solutions $U_* = (\kappa_*, A_*, B_*)$ to (5.21) as given by Theorem 9. Using that any such solution U_* is of order ϵ as a solution to (5.19) due to the rescaling (5.20), it is not difficult to see that any eigenvalue λ of the reduced eigenvalue problem (5.8) which lies near the origin and has $\operatorname{Re}\lambda \geq 0$ is necessarily of order $\mathrm{O}(\epsilon^2)$; see [6, proof of Lemma 8.2] for a similar argument. Thus, the rescaling (5.20) for U_* together with the rescaling

$$(\Phi, \kappa, A, B, \lambda, x) \longrightarrow \left(\Phi, \epsilon\kappa, \epsilon A, \epsilon^2 B, \epsilon^2 \Lambda, \frac{x}{\epsilon}\right)$$

for the linearization captures all unstable Floquet exponents near the origin, while transforming (5.8) into

(5.24)
$$\Lambda \begin{pmatrix} \frac{1}{d_0} & 0 \\ -aK_0 A_* & \frac{1}{d_1} \end{pmatrix} \begin{pmatrix} \Phi \\ A \end{pmatrix} = \left[\begin{pmatrix} \partial_{xx} - 2\kappa_* \partial_x & -2bA_* \\ (-2(a+d)\kappa_* A_* - cB_*)\partial_x & \partial_{xx} - c\kappa_* \partial_x + (\mu + a\Omega) - (a+d)\kappa_*^2 - (1+ab)A_*^2 \end{pmatrix} + \mathrm{O}(\epsilon) \right] \begin{pmatrix} \phi \\ A \end{pmatrix}$$

for a certain constant $K_0 > 0$ that arises due to the coordinate transformations leading from (5.18) to (5.19). Since we will not need the precise value of K_0, we will not compute it.

5.2. Wave trains. We first investigate equilibria of (5.22), which correspond to wave trains of the original reaction-diffusion system (5.4). Equilibria (κ, A, B) have $B = 0$ and satisfy

$$\Omega = \kappa^2 + bA^2, \qquad [-\mu + d\kappa^2 + A^2]A = 0. \tag{5.25}$$

Equilibria $U_0 = (\kappa, 0, 0)$ with $A = 0$ exist for all wave numbers κ with frequency offset given by $\Omega = \kappa^2$. Thus, their group velocity is given by

$$c_{\mathrm{g}}^0 = \frac{\mathrm{d}\Omega}{\mathrm{d}\kappa} = 2\kappa.$$

The linearization of (5.22) about these solutions is given by

$$L_0 = \begin{pmatrix} 2\kappa & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\mu + d\kappa^2 & c\kappa \end{pmatrix},$$

from which we see that they are hyperbolic except when $\kappa = 0$ or $\kappa = \sqrt{\mu/d}$. The bifurcation at $\kappa = \sqrt{\mu/d}$ is a pitchfork which corresponds to the period-doubling bifurcation which we analyze next.

The equilibria bifurcating at $\kappa = \sqrt{\mu/d}$ can be found by solving (5.25) with $A \neq 0$. We find equilibria

$$U_{\mathrm{pd}} = \left(\kappa, \pm\sqrt{\mu - d\kappa^2}, 0\right)$$

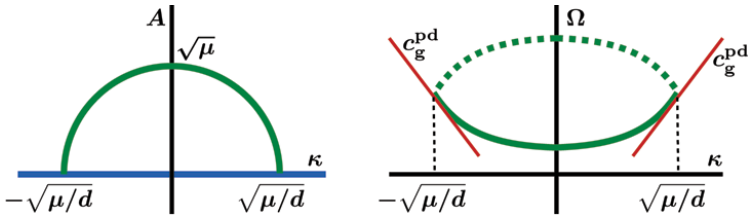


Figure 3. *The bifurcation diagram of the wave trains (left) and their nonlinear dispersion relation (right) are shown: The solid dispersion curve is for* $1 - bd > 0$*, while the dashed curve is for* $1 - bd < 0$*.*

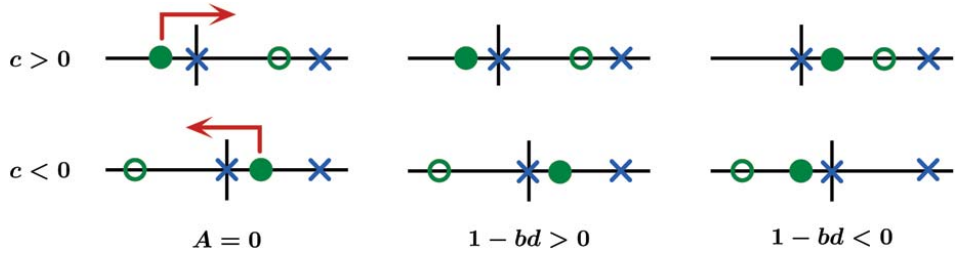


Figure 4. *The spectra of the linearizations* L_0 *and* L_{pd} *of* (5.22) *about the equilibria* U_0 *(left) and* U_{pd} *(center and right), respectively, together with the phase eigenvalue at the origin from the trivial equation* $\phi_x = \kappa$ *are plotted for different signs of the parameters. The open and closed circles denote eigenvalues coming from the period-doubling amplitude* A*, while crosses denote eigenvalues coming from the phase* ϕ*. The closed circles correspond to the eigenvalue* ν_{pd} *that triggers the period-doubling bifurcation; the arrows denote its movement upon decreasing* κ *through* $\sqrt{\mu/d}$*.*

defined for $\kappa^2 < \mu/d$, where

$$\Omega = b\mu + (1 - bd)\kappa^2;$$

see Figure 3. The group velocity of the period-doubled wave trains is therefore given by

$$c_{\mathrm{g}}^{\mathrm{pd}} = \frac{\mathrm{d}\Omega}{\mathrm{d}\kappa} = 2(1 - bd)\kappa.$$

Near $d\kappa^2 = \mu$, the linearization L_{pd} of (5.22) about U_{pd} has eigenvalues near 2κ and $c\kappa$ in addition to the pitchfork eigenvalue given by

$$\nu_{\mathrm{pd}} = -\frac{c_{\mathrm{g}}^{\mathrm{pd}} A^2}{2c\kappa^2} + \mathrm{O}(A^3) = -\frac{(1 - bd)A^2}{c\kappa} + \mathrm{O}(A^3).$$

The spectrum of L_{pd} is illustrated in Figure 4. We remark that the spatial eigenvalue structure reveals in particular that the period-doubled wave trains must be PDE unstable near onset for $c > 0$ and $1 - bd < 0$.

5.3. Coherent structures. Our goal in this section is to shed some light on the nature of the line defect that appears in Figure 1(ii). The line defect mediates between a period-doubled wave train and its 2π-translate. Thus, we shall discuss coherent structures that are spatially asymptotic as $x \to \pm\infty$ to the period-two wave train U_{pd} and its 2π-translate $-U_{\mathrm{pd}}$: These

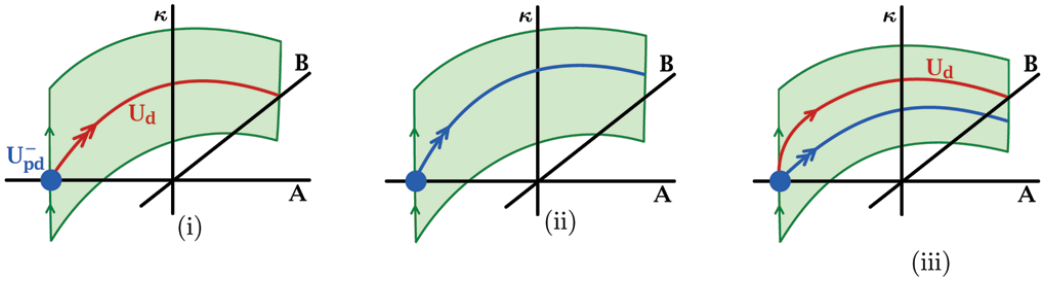


Figure 5. *Figure* (i) *shows part of the reversible heteroclinic orbit* $U_{\rm d}$ *of* (5.26) *for* $b = 0$ *that connects the equilibria* $U_{\rm pd}^-$ *and* $U_{\rm pd}^+$. *Figures* (ii) *and* (iii) *contain the unfolding for* $b < 0$ *and* $b > 0$, *respectively, upon setting* $\Omega := b\mu$: *The reversible heteroclinic orbit persists only for* $b > 0$.

structures correspond to heteroclinic orbits between $U_{\rm pd}$ and $-U_{\rm pd}$ of the spatial differential equation (5.23)

$$\begin{aligned} \kappa_x &= -\Omega + \kappa^2 + bA^2, \\ A_x &= B, \\ B_x &= [-(\mu + a\Omega) + (a+d)\kappa^2 + (1+ab)A^2]A + c\kappa B. \end{aligned} \tag{5.26}$$

There are various limiting cases in which a perturbation analysis is possible. We focus on the perturbation from $b = 0$ as it is the most illuminating case.

When $b = 0$, (5.26) admits the semihyperbolic equilibria $U_{\rm pd}^{\pm} = (0, \pm\sqrt{\mu}, 0)$ for $\Omega = 0$ and $\mu > 0$ which correspond to spatially homogeneous period-doubled wave trains of the reaction-diffusion system (5.2). These equilibria are connected by the heteroclinic orbit

$$U_{\rm d}(x) = \left(0, \sqrt{\mu}\tanh\sqrt{\frac{\mu}{2}}x, \frac{\mu}{\sqrt{2}}\operatorname{sech}^2\sqrt{\frac{\mu}{2}}x\right);$$

see Figure 5(i). This orbit is reversible under the reverser $\mathcal{RS}_{2\pi} : (\kappa, A, B) \mapsto (-\kappa, -A, B)$. We discuss now in what sense the reversible connection $U_{\rm d}$ persists upon varying b near zero, while fixing all other parameters including μ. We focus on the persistence of reversible orbits which are obtained as intersections of unstable manifolds with the B-axis. The following analysis is similar to the one given in [38, section 7].

The first case is $0 < b \ll 1$: We pick $\Omega = b\mu$ so that (5.26) becomes

$$\begin{aligned} \kappa_x &= \kappa^2 + b(A^2 - \mu), \\ A_x &= B, \\ B_x &= [-\mu + (a+d)\kappa^2 + A^2 + ab(A^2 - \mu)]A + c\kappa B. \end{aligned} \tag{5.27}$$

In particular, the equilibrium $U_{\rm pd}^- = (0, -\sqrt{\mu}, 0)$ persists as a semihyperbolic equilibrium for all $b > 0$. We wish to determine how the κ-component of the strong unstable manifold at $x = 0$ depends on b upon varying b near zero. To this end, we record that the adjoint variational equation

$$W_x = -\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ cB_{\rm d}(x) & [-\mu + 3A_{\rm d}^2(x)]A & 0 \end{pmatrix}^* W$$

associated with the linearization of (5.27) about U_{d} at $b = 0$ has the solution $W(x) = (1, 0, 0)$. The Melnikov integral associated with the derivative of the right-hand side of (5.27) with respect to b is therefore given by

$$M := \int_{-\infty}^{0} \left\langle W(x), (A_{\mathrm{d}}^2(x) - \mu) \begin{pmatrix} 1 \\ 0 \\ a \end{pmatrix} \right\rangle \mathrm{d}x = \int_{-\infty}^{0} [A_{\mathrm{d}}^2(x) - \mu]\, \mathrm{d}x = -\sqrt{2\mu} < 0. \tag{5.28}$$

Thus, the unfolding of the heteroclinic orbit near $b = 0$ is as shown in Figure 5, and we conclude that the reversible heteroclinic orbit between the semihyperbolic equilibria $U_{\mathrm{pd}}^{\pm}$ persists only for $b > 0$, but not for $b < 0$. The resulting coherent structure of the reaction-diffusion system is a contact defect in the classification of [38] as it mediates between two wave trains with zero group velocity.

The remaining case is $-1 \ll b < 0$: We set $\Omega = b\mu + (1 - bd)\eta^2$ and treat η as an independent parameter with $\eta \approx 0$ so that (5.26) becomes

$$\begin{aligned} \kappa_x &= -(1 - bd)\eta^2 + \kappa^2 + b(A^2 - \mu), \\ A_x &= B, \\ B_x &= [-\mu - a(1 - bd)\eta^2 + (a + d)\kappa^2 + A^2 + ab(A^2 - \mu)]A + c\kappa B. \end{aligned} \tag{5.29}$$

The parameter η unfolds the saddle-node bifurcation occurring at $b = 0$, leading therefore to the equilibria

$$U_{\mathrm{pd}}^{\pm}(\eta) = (\eta, \pm\sqrt{\mu - d\eta^2}, 0)$$

near $U_{\mathrm{pd}}^{\pm}$. On account of the results in section 5.2, we know that the equilibria $U_{\mathrm{pd}}^{\pm}(\eta)$ have positive group velocity for $\eta > 0$ and negative group velocity for $\eta < 0$. We focus on finding reversible heteroclinic orbits that connect the hyperbolic equilibrium $U_{\mathrm{pd}}^{-}(-\eta)$ at $x = -\infty$ to the hyperbolic equilibrium $U_{\mathrm{pd}}^{+}(\eta)$ at $x = \infty$ for $\eta > 0$. The resulting coherent structure of the reaction-diffusion system is a source in the classification of [38] as it connects a wave train with negative group velocity at $x = -\infty$ to a wave train with positive group velocity at $x = \infty$. To find sources, we note that the behavior of the κ-component of the strong unstable manifold under changes of b at $x = 0$ is, for $\eta = 0$, again determined by the Melnikov integral $M < 0$ in (5.28). Thus, as far as sources are concerned, the unfolding of the heteroclinic orbit near $b = 0$ is as shown in Figure 6. We conclude that, for each $b < 0$ close to zero, there is a unique $\eta > 0$ with a reversible heteroclinic orbit connecting $U_{\mathrm{pd}}^{-}(-\eta)$ at $x = -\infty$ to $U_{\mathrm{pd}}^{+}(\eta)$, while no such connection exists for $b > 0$.

Theorem 10. *For $0 < b \ll 1$, the amplitude equation (5.26) has contact defects that connect the period-doubled spatially homogeneous wave train U_{pd} at $x = -\infty$ and its 2π-time translate at $x = \infty$. For $-1 \ll b < 0$, (5.26) admits sources that connect period-doubled wave trains $U_{\mathrm{pd}}(x)$ with negative group velocity at $x = -\infty$ to the reflected wave trains $U_{\mathrm{pd}}(-x)$ with positive group velocity at $x = \infty$. For a, b, c, and d sufficiently close to zero, both defects are spectrally stable.*

Proof. The existence part has already been proved, and we therefore focus on spectral stability.

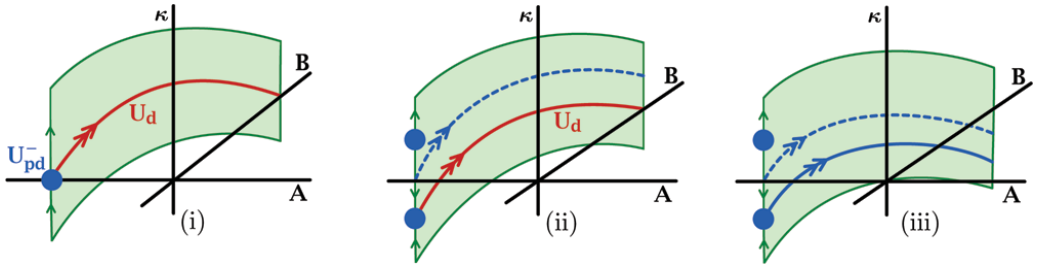


Figure 6. *Figure* (i) *shows part of the reversible heteroclinic orbit* U_{d} *of* (5.26) *for* $b = 0$ *that connects the equilibria* U_{pd}^- *and* U_{pd}^+. *Figures* (ii) *and* (iii) *contain the unfolding for* $b < 0$ *and* $b > 0$, *respectively, upon setting* $\Omega = b\mu + (1 - bd)\eta^2$ *with* $\eta \neq 0$: *The reversible source persists for* $b < 0$.

We consider sources first and set $a = c = d = 0$, $b = -\delta$, and $\Omega = b + \eta^2 = -\delta + \eta^2$ for $\delta > 0$ small. We also rescale the A-equation so that $\mu = 1$. Thus, (5.29) becomes

$$\begin{aligned} \kappa_x &= -\eta^2 + \kappa^2 + \delta(1 - A^2), \\ A_x &= B, \\ B_x &= [A^2 - 1]A \end{aligned}$$

so that $A_*(x) = \tanh(x/\sqrt{2})$ independently of η and δ. The source $U_{\mathrm{d}} = (\kappa_*, A_*, \partial_x A_*)$ decays exponentially to zero as $x \to \pm\infty$ with a rate independent of $\delta \geq 0$ since it lies by construction in the strong unstable and stable manifolds of the asymptotic semihyperbolic equilibria. The reduced PDE eigenvalue problem (5.24) about U_{d} is given by

$$\Lambda \begin{pmatrix} \phi \\ A \end{pmatrix} = \begin{pmatrix} d_0[\partial_{xx} - 2\kappa_*\partial_x] & 2\delta A_* \\ 0 & d_1[\partial_{xx} + 1 - A_*^2] \end{pmatrix} \begin{pmatrix} \phi \\ A \end{pmatrix}. \tag{5.30}$$

Since the constant functions are admissible eigenfunctions for sources according to the counting arguments presented in [38], we see that $\Lambda = 0$ is an eigenvalue with geometric multiplicity two. This is in line with [38, Lemma 4.4], which asserts that sources must have two eigenvalues at the origin. It remains to show that the algebraic multiplicity of $\Lambda = 0$ is two and that there are no other eigenvalues in the closed right half-plane. To prove this claim, we set $\delta = 0$ to get

$$\Lambda \begin{pmatrix} \phi \\ A \end{pmatrix} = \begin{pmatrix} d_0\partial_{xx} & 0 \\ 0 & d_1[\partial_{xx} + 1 - A_*^2] \end{pmatrix} \begin{pmatrix} \phi \\ A \end{pmatrix} =: \begin{pmatrix} \mathcal{L}_0 & 0 \\ 0 & \mathcal{L}_1 \end{pmatrix} \begin{pmatrix} \phi \\ A \end{pmatrix}. \tag{5.31}$$

Sturm–Liouville theory implies that $\mathcal{L}_1$ has a simple eigenvalue $\Lambda = 0$ and no other spectrum in the closed right half-plane. Similarly, $\mathcal{L}_0$ has the eigenvalue $\Lambda = 0$ with eigenfunction $\phi(x) = 1$ and no other spectrum in the closed right half-plane. Since the perturbation leading from (5.31) to (5.30) is small and decays with uniform exponential rate in x, we can apply standard Evans-function theory [18] to conclude that (5.30) with $0 < \delta \ll 1$ has precisely two eigenvalues near the origin, counting multiplicity, which are therefore given by the eigenvalues at $\Lambda = 0$ mentioned above. The same argument applies when perturbing from $(a, c, d) = 0$, which completes the proof for sources.

It remains to consider the contact defects. We set $a = c = d = 0$, $b = \delta$, and $\Omega = b = \delta$ for $\delta > 0$ small and again rescale the A-equation so that $\mu = 1$. The existence problem (5.29) becomes

$$\begin{aligned} \kappa_x &= \kappa^2 - \delta(1 - A^2), \\ A_x &= B, \\ B_x &= [A^2 - 1]A \end{aligned}$$

so that $A_*(x) = \tanh(x/\sqrt{2})$ independently of δ, and we get

$$\kappa_x = \kappa^2 - \delta \operatorname{sech}^2\left(\frac{x}{\sqrt{2}}\right).$$

We record for later use that the reversible contact-defect solution $\kappa_*(x)$ then satisfies

$$-\sqrt{2}\delta \le \kappa_*(x) \le 0, \qquad x \ge 0,$$

with $\kappa_*(0) = 0$ and $\kappa_*(x) = K_1/x^2 + \mathrm{O}(1/x^3)$ as $x \to \infty$ for some $K_1 \le 0$. The reduced PDE eigenvalue problem (5.24) about the contact defect is again given by

$$\Lambda \begin{pmatrix} \phi \\ A \end{pmatrix} = \begin{pmatrix} d_0[\partial_{xx} - 2\kappa_*\partial_x] & 2\delta A_* \\ 0 & d_1[\partial_{xx} + 1 - A_*^2] \end{pmatrix} \begin{pmatrix} \phi \\ A \end{pmatrix}. \tag{5.32}$$

As shown in [39, Theorem 3], contact defects have generically a single simple eigenvalue at the origin. Furthermore, it is a consequence of the results in [39] that the only admissible eigenfunctions of (5.32) are those that decay algebraically as $x \to \pm\infty$. We will therefore focus on the decoupled eigenvalue problem

$$\phi_{xx} - 2\kappa_*(x)\phi_x = \frac{\Lambda\phi}{d_0} \tag{5.33}$$

for ϕ and prove that it has no spectrum in the closed right half-plane for $\delta > 0$. Eigenfunctions belonging to nonzero eigenvalues Λ of (5.33) in the closed right half-plane decay necessarily exponentially with nonzero rate $\sqrt{\Lambda}$ and, using the algebraic convergence $\kappa_*(x) = K_1/x^2 + \mathrm{O}(1/x^3)$ of the contact defect as $x \to \pm\infty$, we may therefore set

$$\Phi(x) := \exp\left(\int_\infty^x \kappa_*(y)\,\mathrm{d}y\right)\phi(x),$$

which transforms the eigenvalue problem for ϕ into the equivalent eigenvalue problem

$$\Phi_{xx} - \delta \operatorname{sech}^2\left(\frac{x}{\sqrt{2}}\right)\Phi = \frac{\Lambda\Phi}{d_0}$$

for Φ. For $\delta > 0$, there are no eigenvalues in the closed right half-plane, with the possible exception of the origin. We focus therefore on the eigenvalue problem (5.33) with $\Lambda = 0$, which is given by

$$\phi_{xx} - 2\kappa_*(x)\phi_x = 0. \tag{5.34}$$

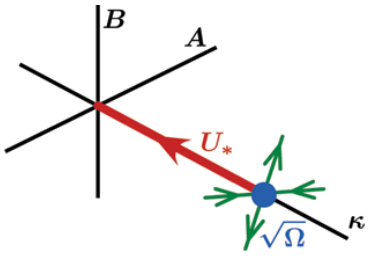


Figure 7. *The boundary sink U_* which accommodates Neumann boundary conditions and the period-one wave train with nonzero wave number $\sqrt{\Omega}$ is shown for $\Omega > 1/d$.*

For $\delta > 0$, the unique solution which decays algebraically as $x \to \infty$ is given by

$$\phi(x) = \int_\infty^x \exp\left(\int_\infty^y 2\kappa_*(z)\,\mathrm{d}z\right)\mathrm{d}y.$$

This solution is an eigenfunction provided it is odd as $\phi(x) = 1$ is the unique even solution of (5.34). Thus, we need $\phi_x(0) = 0$ but have

$$\phi_x(0) = \exp\left(\int_\infty^0 2\kappa_*(z)\,\mathrm{d}z\right) \neq 0,$$

which proves that $\Lambda = 0$ is not an eigenvalue of (5.33). Last, the perturbation from $(a, c, d) = 0$ can be dealt with by regular perturbation theory using the Evans-function construction in [39]. ■

5.4. Boundary sinks. Last, we investigate the existence and stability of boundary sinks for Neumann boundary conditions. In other words, we seek solutions $U(x)$ of

$$\begin{aligned} \kappa_x &= -\Omega + \kappa^2 + bA^2, \\ A_x &= B, \\ B_x &= [-(\mu + a\Omega) + (a + d)\kappa^2 + (1 + ab)A^2]A + c\kappa B \end{aligned} \tag{5.35}$$

for $x \leq 0$ so that $U(0)$ lies on the A-axis, corresponding to Neumann boundary conditions, and $U(x)$ converges to an equilibrium U_- of (5.35) with positive group velocity as $x \to -\infty$.

We focus on the equilibria $U_0 = (\kappa, A, b) = (\sqrt{\Omega}, 0, 0)$ which correspond to the period-one wave trains with nonzero wave number $\sqrt{\Omega}$ and group velocity $c_{\mathrm{g}}^0 = 2\kappa = 2\sqrt{\Omega} > 0$. In this case, the boundary sink is given explicitly by

$$U_*(x) = (\kappa, A, B)(x) = \left(-\sqrt{\Omega}\tanh(\sqrt{\Omega}x), 0, 0\right), \qquad x \leq 0;$$

see Figure 7. The PDE stability of the boundary sink U_* can be analyzed as follows. Evaluating (5.24) at $\epsilon = 0$, we find that the reduced eigenvalue problem associated with the boundary sink $U_* = (\kappa_*(x), 0, 0)$ is given by

$$\begin{aligned} \phi_{xx} - 2\kappa_*(x)\phi_x &= \frac{\Lambda}{d_0}\phi, \\ A_{xx} - c\kappa_*(x)A_x + [\mu + a\Omega - (a + d)\kappa_*^2(x)]A &= \frac{\Lambda}{d_1}A \end{aligned}$$

on $\mathbb{R}^-$ together with Neumann boundary conditions $\phi_x(0) = A_x(0) = 0$. The equation for ϕ decouples and coincides with the eigenvalue problem of Lax shocks of the Burgers equation; in particular, there are no point eigenvalues in the closed right half-plane, and the essential spectrum consists of the curve $\Lambda/d_0 = -k^2 - 2\sqrt{\Omega}ik$ for $k \in \mathbb{R}$; see, for instance, [6, Lemma 8.2]. It remains to analyze the equation for A given by

$$A_{xx} + c\sqrt{\Omega}\tanh(\sqrt{\Omega}x)A_x + \left[\mu - d\Omega - (a+d)\Omega\,\mathrm{sech}^2(\sqrt{\Omega}x)\right]A = \frac{\Lambda}{d_1}A, \qquad x < 0, \tag{5.36}$$
$$A_x(0) = 0.$$

The essential spectrum of (5.36) is given by

$$\frac{\Lambda_{\mathrm{ess}}(k)}{d_1} = \mu - \Omega d - k^2 - c\sqrt{\Omega}ik, \tag{5.37}$$

and we denote by

$$\frac{\Lambda_{\mathrm{bp}}}{d_1} = \mu - \Omega\left(d + \frac{c^2}{4}\right)$$

the branch point of the linear dispersion relation Λ_{ess}. The point spectrum of (5.36) can also be calculated explicitly: Using the independent variable $z = \tanh(\sqrt{\Omega}x)$, real-valued solutions to (5.37) are given in terms of Ferrers functions which are appropriate linear combinations of the associated Legendre functions [25, section 5]. Using the results in [25, sections 5.12 and 5.15], we find that the point spectrum of (5.36) consists precisely of the points Λ_n given by

$$\frac{\Lambda_n}{d_1} = \mu - \Omega d + \frac{\Omega}{4}\left(\left[\sqrt{(c-1)^2 + 4(a+d)} - 1 - 4n\right]^2 - c^2\right)$$

for those integers $n \geq 0$ for which

$$\sqrt{(c-1)^2 + 4(a+d)} \geq 1 + 4n. \tag{5.38}$$

In particular, the rightmost point eigenvalue Λ_0 is given by

$$\frac{\Lambda_0}{d_1} = \mu - \Omega d + \frac{\Omega}{4}\left(\left[\sqrt{(c-1)^2 + 4(a+d)} - 1\right]^2 - c^2\right), \tag{5.39}$$

assuming that the term in the square brackets is positive.

We shall assume from now on that the group velocity c of the period-doubling mode is negative so that $c < 0$. The period-one wave train U_- undergoes a pitchfork bifurcation at $\mu = \Omega d$ which, as outlined in section 5.2, leads to the period-two wave train U_{pd} which has a nonzero A-component. We discuss now how this bifurcation, which occurs when the essential spectrum Λ_{ess} crosses the imaginary axis, interacts with the bifurcation of boundary sinks, which occurs when the eigenvalue Λ_0 destabilizes. There are three relevant cases:

(1) Λ_0 does not exist; that is, (5.38) is not met for $n = 0$.
(2) $\Lambda_{\mathrm{bp}} < \Lambda_0 < \Lambda_{\mathrm{ess}}$.
(3) $\Lambda_{\mathrm{ess}} < \Lambda_0$.

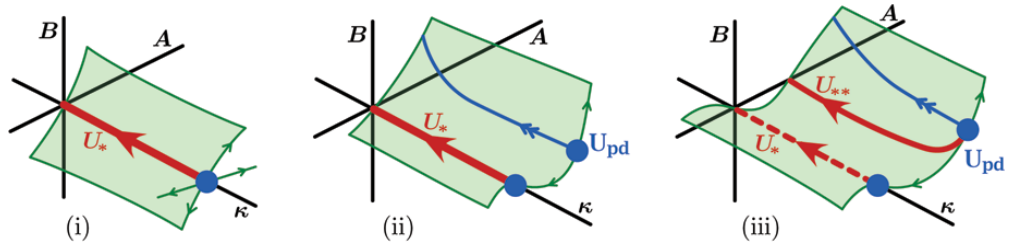


Figure 8. *We illustrate case* (2) *($c < 0$ and $\Lambda_{bp} < \Lambda_0 < \Lambda_{ess}$) for increasing μ under the assumption that the bifurcation associated with Λ_0 is supercritical. The first instability occurs when $\Lambda_{ess} = 0$, leading in* (ii) *to a stable period-two wave train U_{pd} which does not persist under Neumann boundary conditions as there is no boundary sink available. When Λ_0 destabilizes, the stable boundary sink U_{**} bifurcates from U_*, and the period-two wave train U_{pd} persists now under Neumann boundary conditions due to the presence of U_{**}.*

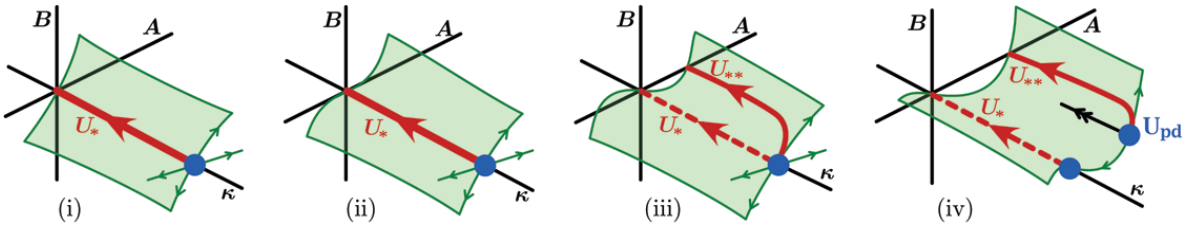


Figure 9. *Case* (3) *($c < 0$ and $\Lambda_{ess} < \Lambda_0$) is illustrated for increasing μ: The boundary sink U_* destabilizes in* (ii) *when the eigenvalue Λ_0 crosses the imaginary axis. This leads in* (iii) *to the existence of a stable period-doubled boundary sink U_{**}. The essential instability which occurs when $\Lambda_{ess} = 0$ leads then in* (iv) *to a stable period-two wave train U_{pd} which persists under Neumann boundary conditions due to the presence of U_{**}.*

Using $c \le 0$, we see that the last case occurs for $a + d > 0$, while the eigenvalue Λ_0 disappears in the branch point Λ_{bp} when the term in the square brackets in (5.39) becomes zero. Since case (1) has already been discussed in section 3.4, and case (2) is similar to (3), we concentrate in the following on (3) and refer the reader to Figure 9 for an illustration of case (2).

Thus, assume that $c < 0$ and $\Lambda_{ess} < \Lambda_0$: Upon increasing μ, the boundary sink U_* destabilizes when $\Lambda_0 = 0$. In terms of the spatial ODE (5.35), this bifurcation manifests itself as a tangency of the unstable manifold of the equilibrium U_- as indicated in Figure 9(ii). We show in Lemma 5.1 below that this bifurcation can be supercritical, thus leading to a stable boundary sink U_{**} which connects U_- to the A-axis as illustrated in Figure 9(iii). Since the A-component of U_{**} is not zero, the boundary sink U_{**} will have period two, even though the period-two wave train U_{pd} has not yet bifurcated from U_-. A further increase of μ then leads to the period-two wave trains U_{pd} which persist under Neumann conditions thanks to the boundary sink U_{**} as indicated in Figure 9(iv). The characteristic feature of scenario (3) is therefore that the period-doubling sets in first at the boundary, where it is also most pronounced during the entire bifurcation sequence. It remains to prove that the pitchfork bifurcation of the boundary sink is supercritical.

Lemma 5.1. *Assume that $b \le 0$, $c < 0$, $a + d > 0$, and $1 + ab \ge 0$; then the pitchfork bifurcation of the boundary sink U_*, which occurs when $\Lambda_0 = 0$, is supercritical.*

Proof. We need to prove that the part of the unstable manifold of U_- which lies in $A > 0$ has $B \ge 0$. Using the projective coordinate $w = B/A$, we arrive at the system

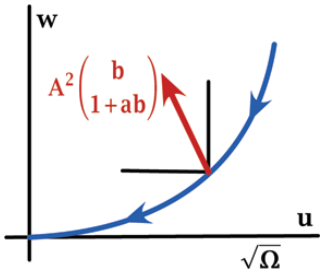


Figure 10. *Proof of Lemma 5.1: If the tangent space angle $w(x)$ decreases monotonically, then the unstable manifold lies above the tangent space for $b < 0$ and $1 + ab > 0$.*

$$\begin{aligned} u_x &= -\Omega + u^2 + bA^2, \\ A_x &= Aw, \\ w_x &= -(\mu + a\Omega) + (a + d)u^2 + (1 + ab)A^2 - cuw - w^2. \end{aligned}$$

If the solution that corresponds to the tangent space of the unstable manifold of U_- evaluated along the boundary sink U_* decreases monotonically for $x \in \mathbb{R}^-$, then the bifurcation will indeed be supercritical for $b < 0$ and $1 + ab > 0$ as outlined in Figure 10 since the nonlinear terms involving A point in the right direction. The tangent space of the unstable manifold of U_- evaluated along the boundary sink U_* satisfies the linearized equation

$$w_x = -(\mu + a\Omega) + (a + d)u_*^2(x) - cu_*(x)w - w^2.$$

We claim that $w_x < 0$ for all x for the solution that converges as $x \to -\infty$ to the tangent space of the unstable manifold of U_*. First, for u near $\sqrt{\Omega}$, we write $u = \sqrt{\Omega} - h$ and $w(x) = w_* + W(x)$, where w_* is the unique positive solution of

$$-(\mu + a\Omega) + (a + d)\Omega - c\sqrt{\Omega}w_* - w_*^2 = 0, \tag{5.40}$$

which corresponds to the unstable eigenvector of the linearization of (5.35) about U_-. The resulting system for W is

$$W_x = (c\sqrt{\Omega} - 2w_*)W - \frac{h}{\sqrt{\Omega}}[2\Omega(a + d) + c\sqrt{\Omega}w_*] + \mathrm{O}(h^2 + W^2).$$

Substituting (5.40), we get

$$W_x = (c\sqrt{\Omega} - 2w_*)W - \frac{h}{\sqrt{\Omega}}[\Omega(a + d) + \mu + a\Omega + w_*^2] + \mathrm{O}(h^2 + W^2),$$

which means that $W_x < 0$ for $h > 0$ since the term in the square brackets turns out to be positive when $\Lambda_0 = 0$. A similar argument shows that the solution $w(x)$ satisfies $w_{xx}(x) < 0$, whenever $w_x(x) = 0$, which completes the proof. ■

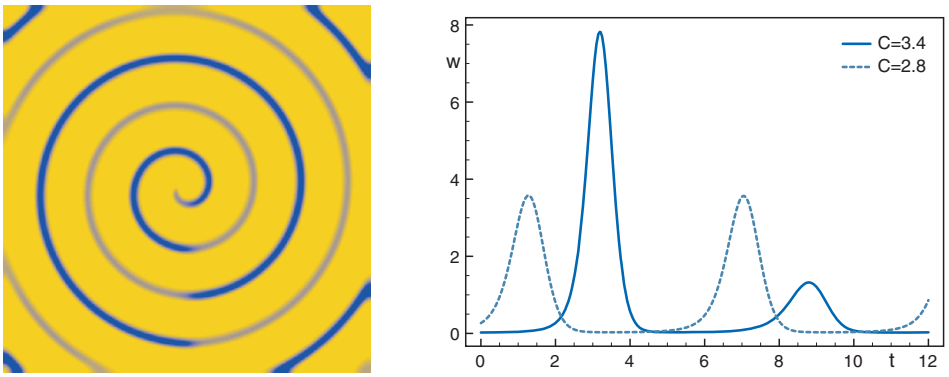


Figure 11. *A contour plot of the w-component of the period-doubled spiral wave is plotted in the left figure for $C = 3.4$. To the right, the w-components of the spatially homogeneous oscillations are plotted as functions of time.*

6. Period-doubling of spirals in the Rössler system: A case study. In this section, we apply our findings to the planar 3-component PDE

$$\begin{aligned} u_t &= 0.4\,\Delta u - v - w, \\ v_t &= 0.4\,\Delta v + u + 0.2\,v, \\ w_t &= 0.4\,\Delta w + uw - Cw + 0.2, \end{aligned} \tag{6.1}$$

written abstractly as

$$U_t = 0.4\,\Delta U + f(U, C), \tag{6.2}$$

on a square $(x, y) \in (0, L)^2 \subset \mathbb{R}^2$ with Neumann boundary conditions. Goryachev, Chaté, and Kapral [14] observed period-doubled spiral waves for (6.1) when changing the parameter C in the interval $(2.8, 3.4)$; see Figures 1(ii) and 11.

Spatially homogeneous solutions of (6.1) satisfy the Rössler equation

$$\begin{aligned} u_t &= -v - w, \\ v_t &= u + 0.2\,v, \\ w_t &= uw - Cw + 0.2, \end{aligned} \tag{6.3}$$

which is known to exhibit periodic solutions which undergo a period-doubling sequence beginning at $C = 2.83$; see Figure 11. The periodic solutions of (6.3) are accompanied by 1D wave trains $U(kx - \omega t)$ of (6.2) with nonzero wave number k which can be found as 2π-periodic solutions of the traveling-wave ODE

$$0.4\,k^2 U_{xx} + \omega U_x + f(U, C) = 0, \qquad x \in \mathbb{R}. \tag{6.4}$$

In the remainder of this section, we report on numerical computations for (6.1) and (6.4). We used Barkley's finite-difference code EZSPIRAL [3] for direct numerical simulations of spiral-wave solutions to (6.1), typically with $L = 250$, and the boundary-value solver AUTO97 [5] for

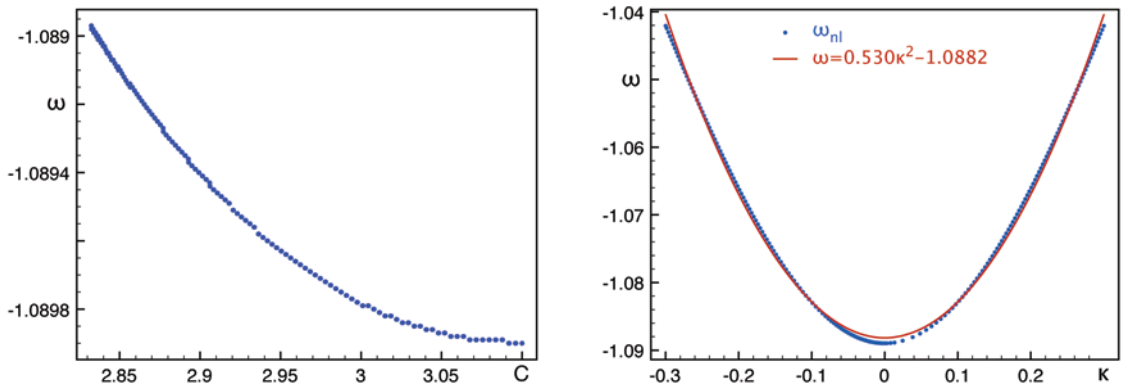


Figure 12. *In the left figure, we plot the nonlinear dispersion relation $\omega_{\rm nl}(k)$ of the 1D wave trains of (6.4) for $C = 2.8324$, i.e., at the period-doubling bifurcation. To the right, the temporal frequency ω of the period-doubled spatially homogeneous oscillations is plotted against the parameter C. Since onset occurs at $C = 2.834$, the bifurcation is supercritical, whence $\hat{b}_3 > 0$ in (6.5). Since the frequency decreases with the period-doubling amplitude, we see that the coefficient b_0 in (6.5) is negative.*

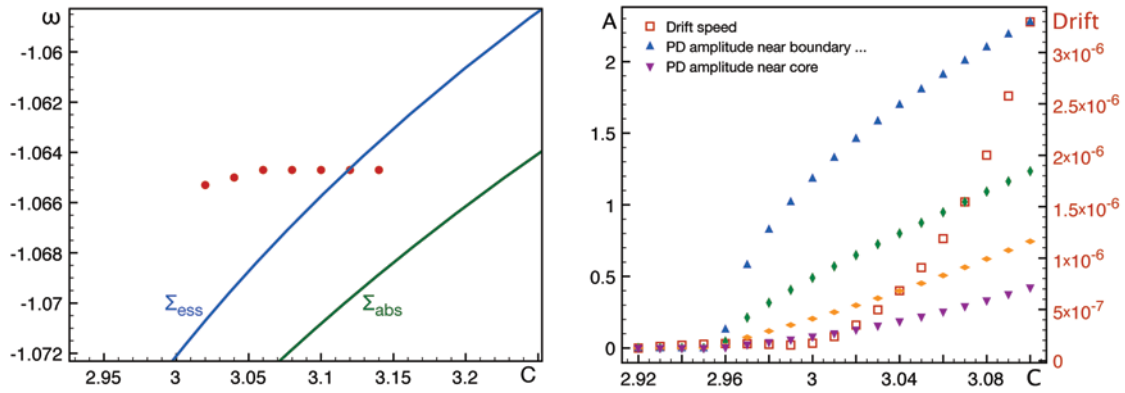


Figure 13. *The left plot shows the onset of absolute and essential period-doubling instabilities of 1D wave trains with temporal frequency ω as a function of C, and it also contains the frequencies selected by the 2D spiral waves of (6.1). The right figure shows the drift velocity of the spiral tip (right y-axis) and the period-doubling amplitude A of the spirals (left y-axis) evaluated at different points along a ray from the core to the boundary as functions of the parameter C (see text for details).*

all computations relating to the traveling-wave ODE (6.4). In particular, the absolute and essential spectra of wave trains are computed with AUTO97 using the algorithms outlined in [31, 35].

The nonlinear dispersion relation $\omega = \omega_{\rm nl}(k)$ of the wave trains of (6.4) is shown in Figure 12. Note that their phase velocity $c_{\rm p} = \omega/k$ and their group velocity $c_{\rm g} = {\rm d}\omega/{\rm d}k$ have opposite signs: Since the 2D spiral waves select the wave trains with positive group velocity, the wave trains in the far field of the 2D spirals travel toward the core rather than toward the boundary.

Next, we plot in Figure 13 the curves where the essential and absolute spectra of the 1D wave trains with frequency ω cross the imaginary axis. These instabilities are caused

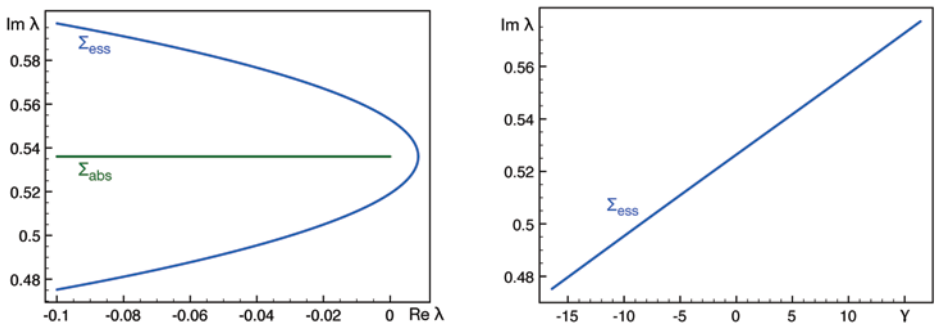


Figure 14. *Absolute and essential period-doubling spectra of the spiral waves are shown (left). The right plot shows the linear period-doubling dispersion relation with* $\operatorname{Im}\lambda_{\mathrm{pd}}$ *plotted against the associated wave number* y*: The linear group velocity* $c_{\mathrm{g}}^{\mathrm{pd}}$ *is therefore negative.*

by period-doubling modes with negative linear group velocity $c_{\mathrm{g}}^{\mathrm{pd}} < 0$; see Figure 14. Direct numerical simulations of (6.1) allow us to determine the temporal frequencies selected by spiral waves for different values of C, which are also shown in Figure 13. The associated spatial wave number of the wave trains in the far field is $k \approx 0.2$, which is close to zero in line with the observation that the period-doubling bifurcations are organized by spatially homogeneous oscillations.

The closeness to spatially homogeneous oscillations allows us to investigate the nature of the line defect that is visible in Figure 11 by applying the results of section 5.3 about coherent structures: Figure 12 shows that the coefficient b arising in (5.26) is negative, and Theorem 10 then implies that the line defect in Figure 11 is a source, rather than a contact defect. The analysis in section 5.3 predicts a $\tanh(x)$ profile of the period-doubling mode across the line defect which has indeed been measured in [14, equation (2) and Figure 3] based on numerical simulations of (6.1). We refer the reader to [48] for an analysis of line defects based on interpreting spirals as a field of coupled oscillators.

To determine when and how the spiral waves destabilize upon increasing C, we fix points (x_0, y_0) in the domain and record the time series $w_*(x_0, y_0, t)$ of the spiral wave. We then compute the difference between consecutive maxima of the time series which we use as a measure for the period-doubling amplitude. This computation is done for five points (x_j, y_j) which are spaced equidistantly on a ray that connects the spiral core to the boundary and avoids the line defect. Since our theoretical results predict that period-doubled spirals ought to drift, we also computed the spiral tip and its drift velocity. The results are shown in Figure 13. The indications are therefore that period-doubling of spirals sets in at $C \approx 2.96$. The instability appears to be most visible at the boundary, with a square-root type behavior reminiscent of pitchfork and Hopf bifurcations, and less pronounced toward the core. The spiral does begin to drift, but the drift velocity is very small, and we could not identify a clearly defined transition point to drift.

We now discuss the different possible mechanisms outlined in section 4 that may be responsible for the observed period-doubling in the Rössler system. First, we plot in Figure 14 representative absolute and essential spectra of the asymptotic 1D wave trains. Lemma 2.3 asserts that the absolute spectrum arising due to period-doubling bifurcations of wave trains

near spatially homogeneous oscillations has to lie on the symmetry line $\operatorname{Im}\lambda = \omega/2$, and this is indeed what happens here for the spatial wave numbers $k \approx 0.2$ selected by the spirals. Figure 14 also shows that the linear group velocity c_g^{pd} of the period-doubling modes is negative. However, both absolute and essential spectra are still in the left half-plane when the period-doubling sets in at $C = 2.96$. Furthermore, due to $c_g^{\mathrm{pd}} < 0$, the absolute eigenmodes decay toward the boundary, which appears to contradict Figure 13 which seems to imply that period-doubling is more pronounced at the boundary. Thus, the bifurcation does not seem to be caused directly by the absolute spectrum.

The second possibility is that the instability is caused by point eigenvalues that emanate from the branch point located at the edge of the absolute spectrum due to curvature effects of the Laplacian; see Remark 4.2. We have evaluated numerically the criterion derived in [41, section IV] using the algorithm described there and found that, in the notation of [41], $\Phi = \pi$, which means that no point eigenvalues arise near the branch point.

This leaves the last option, namely, that period-doubling is caused by point eigenvalues of the boundary sink. We have discussed this case in section 5.4 for boundary sinks in the near-spatially homogeneous case and shown that these sinks can indeed possess isolated point eigenvalues that are in resonance with $\mathrm{i}\omega/2$. In particular, the scenario described in Figure 9 is consistent with the numerical observations reported in Figure 13 and therefore provides the likeliest explanation for the occurrence of period-doubling in the Rössler system: As seen from Figure 9(iii), the period-doubling amplitude is most visible at the boundary, while decreasing toward the core. Since the adjoint eigenfunction associated with the translational eigenmodes of the spiral wave decreases exponentially toward the boundary, we expect that the drift coefficient is exponentially small in the domain diameter which may explain the slow drift observed in Figure 13.

To further corroborate this conclusion, we exploit that the wave trains selected by the spiral waves have wave numbers near zero and are therefore close to spatially homogeneous oscillations. Thus, if we can determine the coefficients appearing in the reduced eigenvalue problem (5.36) of the boundary sinks, then we can calculate the approximate location of the rightmost eigenvalue Λ_0 given in (5.39) and the expected onset of period-doubling. The unscaled version (5.5) of the amplitude equations is given by

$$\begin{aligned} d_0\kappa_x &= -\bar{\omega} + \frac{1}{2}\omega_{\mathrm{nl}}''(0)\kappa^2 + b_0 A^2, \\ d_1 A_{xx} &= [-\mu + \hat{b}_1\kappa_x + \hat{b}_2\kappa^2 + \hat{b}_3 A^2]A + b_4\kappa A_x. \end{aligned} \tag{6.5}$$

We remark that the coefficients b_0 and $\hat{b}_3$ do not enter into the calculations presented in section 5.4, but Figure 12 indicates that $\hat{b}_3$ is positive as required. The coefficient b_0 is relevant for the line defect of the period-doubled spiral and has already been discussed above. The parameter μ will be replaced later by an appropriate expression in C.

First, we note that $d_0 = d_1 = 0.4$ are equal to the diffusion coefficient in (6.1) since the diffusion matrix is a multiple of the identity. The fit to the nonlinear dispersion relation presented in Figure 12 gives $\frac{1}{2}\omega_{\mathrm{nl}}''(0) = 0.530$. The coefficient $\hat{b}_1$ can be computed numerically by evaluating (5.15): Since the diffusion matrix in (6.1) is a multiple of the identity, it follows from (5.13) that $u_\omega = 0$. After calculating the adjoint solution ψ_{pd} of the linearization of (6.3)

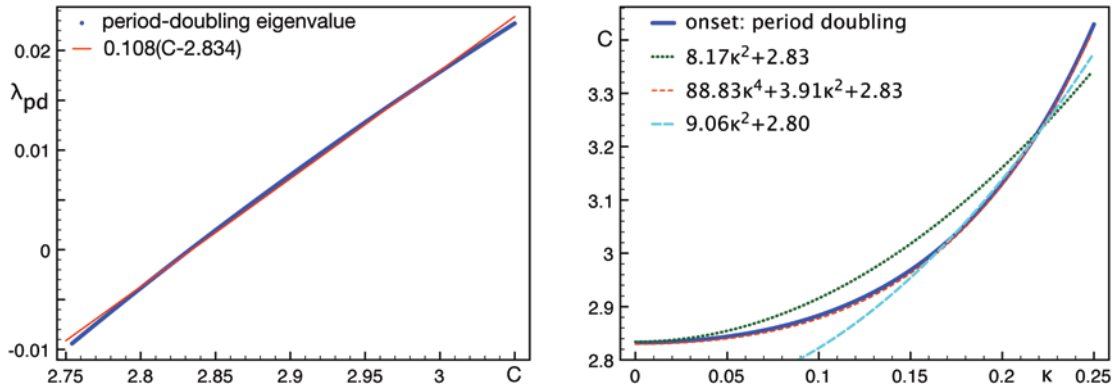


Figure 15. *In the left figure, we plot the period-doubling Floquet exponent λ_{pd} of the spatially homogeneous wave trains of (6.3) as a function of the parameter C. The right figure shows the period-doubling bifurcation curve of wave trains to (6.4) with wave number k together with various curve fits.*

about the homogeneous oscillation, we obtain $\hat{b}_1 = 0.925$. The coefficient b_4 is equal to the slope of the linear group velocity of the period-doubling mode, considered as a function of the wave number k of the underlying wave train: our computation of this slope gives $b_4 = -0.74$; see [31] for the relevant algorithms. Next, we need to express the bifurcation parameter μ in (6.5) by an appropriate expression in C: To this end, we calculated in Figure 15 the period-doubling Floquet exponent of the homogeneous oscillations as a function of C. A least-square fit gives $\lambda_{\mathrm{pd}} = 0.108(C - 2.834)$, and therefore $\mu = 0.108(C - 2.834)$ since μ in (6.5) and λ_{pd} in the reduced eigenvalue problem appear with the same coefficient. Summarizing the results obtained so far, we arrive at the equation

$$
\begin{aligned}
0.4\kappa_x &= -\bar{\omega} + 0.530\kappa^2, \\
0.4A_{xx} &= [-0.108(C - 2.834) + 0.925\kappa_x + \hat{b}_2\kappa^2]A - 0.74\kappa A_x.
\end{aligned} \tag{6.6}
$$

It remains to determine the coefficient $\hat{b}_2$, which measures the dependence of the onset of period-doubling on the wave number k of the underlying wave train of (6.4). This relation, together with various curve fits, is plotted in Figure 15. We recall that the spirals select wave numbers of around $k = 0.203$. In this region, however, there is unfortunately no accurate fit of the required form $C - 0.2834 = \hat{b}_2 k^2$. Thus, we do not seem to be in the region where the approximation by (6.5) is valid. We therefore proceed as follows: First, we may take $\hat{b}_2 = 9.06$, which corresponds to a quadratic fit in the interval $k \in (0.15, 0.25)$. Alternatively, we replace the $\hat{b}_2\kappa^2$ term in (6.6) by $88.8\kappa^4 + 3.91\kappa^2 + 2.83$, which is an excellent fit of Figure 15. In the latter case, our formula (5.39) for the rightmost boundary-sink eigenvalue is no longer valid, and we determine this eigenvalue numerically using finite differences. Both approximations result in an eigenvalue Λ_0 that destabilizes prior to the essential spectrum. The predicted parameter values for the onset of the point instability are $C = 2.99$, when taking $\hat{b}_2 = 9.06$, and $C = 3.06$ for the approximation by the quartic polynomial, compared with the value $C = 2.96$ indicated by the direct simulations from Figure 13. Thus, while the predicted and measured values for onset disagree, the amplitude equation does predict that the instability is caused by a point eigenvalue of the boundary sink rather than by the absolute spectrum.

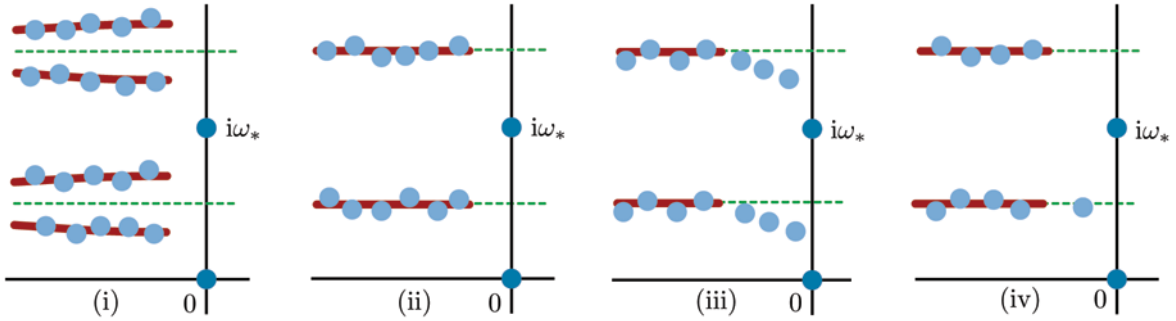


Figure 16. *Schematic plots of spiral spectra on large disks are shown in the complex plane under the assumption that the asymptotic wave trains undergo a period-doubling bifurcation: Solid lines represent the part of the absolute spectrum associated with period-doubling modes, solid bullets mark eigenvalues, and the dotted lines indicate where* $\operatorname{Im}\lambda = \omega_*/2$. *Also plotted are the rotation and translation eigenvalues at* $\lambda \approx 0, \pm i\omega_*$ *which persist on large disks. Note that absolute spiral spectra are periodic along the imaginary axis with period* $i\omega_*$. *The different cases, and their meanings, are discussed further in section 7.*

7. Discussion.

Period-doubling of spiral waves. In this paper, we investigated period-doubling bifurcations of 1D sources and 2D spiral waves. We started from the observation that period-doubling of sources and spirals must actually be a Hopf bifurcation at precisely half the temporal frequency ω_* of these patterns. This observation led to a prediction and a puzzle: The prediction is that period-doubled spirals should drift due to the 2:1 resonance of Hopf and rotation frequency, and we indeed found numerical evidence that period-doubled spirals in the Rössler system drift, albeit slowly.

The puzzle, and indeed the central theme of this paper, is the nature of the mechanism that causes the imaginary parts of the Hopf eigenvalues to lock to $\omega_*/2$ in a robust fashion. Indeed, with no extra structure present, temporal Hopf bifurcations at isolated point eigenvalues λ_H with $\operatorname{Im}\lambda_H = \omega_*/2$ have codimension two and not one, as observed experimentally and numerically. The resolution to this issue, put forward in this paper, is as follows. The spiral waves emerging at the locked Hopf bifurcation exhibit a period-doubled spatial structure in the far field. This suggests that the locked Hopf bifurcation is caused by period-doubling of the asymptotic wave trains in the far field of the spiral wave. Thus, we assumed that the asymptotic wave trains undergo a period-doubling bifurcation, as solutions to the traveling-wave ODE or to the reaction-diffusion system on the real line, and investigated the implications for spiral spectra on large bounded domains. We found that there are four generic cases for how the critical, neutrally stable parts of spiral spectra may look, of which three are consistent with spatio-temporal period-doubling of spiral waves. These four cases are illustrated in Figure 16, and we discuss them now in more detail.

First, we recall that the absolute spectrum associated with the period-doubling modes of the asymptotic wave trains is invariant under reflections across the line $\operatorname{Im}\lambda = \omega_*/2$. Hence, branches of the absolute spectrum are either symmetric, i.e., contained in the line of reflection (Figure 16(ii)), or asymmetric, i.e., not intersecting this line (Figure 16(i)); both cases are generic, and symmetric spectrum leads to robust Hopf bifurcations with frequency $\omega_*/2$, while asymmetric spectrum does not. We also showed that the critical part of the absolute spectrum

is always symmetric if the wave trains have sufficiently small wave number, that is, are close to being spatially homogeneous. We stress that the absolute spectrum repeats itself periodically in the vertical direction with period $i\omega_*$, so that there are infinitely many branches of absolute spectrum that cross the imaginary axis simultaneously.

Besides the PDE eigenvalues that lie near the absolute spectrum and accumulate onto it as the domain size increases, other isolated eigenvalues may arise in three different ways:

(1) Discrete eigenvalues and resonance poles of the planar spiral wave persist on large domains.
(2) A family of discrete eigenvalues may emerge from the edges of the absolute spectrum as indicated in Figure 16(iii), depending on a certain sign condition derived in [41].
(3) Discrete eigenvalues and resonance poles of the boundary sink that accommodates the boundary conditions persist on large domains.

As argued above, the resulting eigenvalues from (1), which can be associated with the core of the spiral wave, typically have an imaginary part different from $\omega_*/2$. The eigenvalues from (2) emerge from the edge of each of the infinitely many branches of the absolute spectrum and are due to the $1/r$ curvature terms in the Laplacian: they have an imaginary part close to $\omega_*/2$ and lead therefore to spatio-temporal period-doubling. Last, Neumann boundary sinks may, as shown in section 5.4, have a simple eigenvalue near the line $\operatorname{Im}\lambda = \omega_*/2$ for wave trains that are almost spatially homogeneous; see Figure 16(iv). For the Rössler system with Neumann conditions, the evidence presented in section 6 indicates that period-doubling of spirals is caused by the boundary sink as illustrated in Figure 16(iv).

The conclusions on the different possible spectral instability scenarios presented above are valid for generic boundary conditions. We also expect them to hold when domain boundaries are replaced by interfaces to other patterns. One example is period-doubling of spiral waves in domains with periodic boundary conditions which has been reported in [15, Figure 1a]. The effective boundary for each spiral wave is then given by the Lax-shock interface that is formed between wave trains that propagate toward each other.

A key ingredient to realizing spatio-temporal period-doubling of spirals in a given reaction-diffusion system is therefore that the associated traveling-wave ODE exhibits period-doubling bifurcations of wave trains or periodic orbits. Besides the Rössler system (1.3) and the related three-component Willamoski–Rössler system (see [16] and references therein), we are not aware of any systems that exhibit period-doubling bifurcations of wave trains, although these have been observed in lattice maps [1].

While our spectral analysis is rigorous, large parts of our nonlinear analysis are only formal. For instance, the prediction of drift for truncated sources is based on the reduced equation (3.8) on an appropriate center manifold: It is not clear whether the center-manifold reduction is valid in a uniform region near the source, and not even whether the Taylor jet of the reduced vector field has a limit as the domain diameter goes to infinity. Similarly, we are currently not able to analyze the nonlinear bifurcation of spiral waves on the plane or on large bounded disks. The situation for spirals is worse, compared to that of 1D sources, as infinitely many copies of essential and absolute spectra cross the imaginary axis simultaneously at $\lambda = i\omega_*/2 + i\omega_*\ell$ with $\ell \in \mathbb{Z}$.

Line defects. The line defect visible in Figure 1(ii) is another interesting aspect of the spatio-temporal period-doubling of spiral waves. Previous work on line defects (see [16, 29, 48])

Figure 17. *A snapshot of a period-doubled spiral wave with five interacting line defects is plotted for the Rössler system* (6.1) *with* $C = 3.4$.

modeled line defects using kinetic theory. We clarified the structure of line defects near onset in systems where the asymptotic wave trains are close to spatially homogeneous by showing that they are sources or reversible contact defects. The spatial asymptotics of spiral eigenmodes in (4.11), which reflects the periodicity of spiral spectra in the complex plane, indicate that multiple line defects are possible as higher harmonics of a single line defect, and Figure 17 shows that several line defects can indeed be excited near onset. Neighboring line defects typically attract each other, which eventually leads to pairwise annihilation. Preliminary analyses show that the time scales of this interaction depend strongly on whether the line defects are of source or contact type.

Besides straight line defects similar to those shown in Figure 1(ii), other types of defects have been observed in experiments. Examples are line defects[2] which curl around and form spirals. For the Belousov–Zhabotinsky reaction, stationary defects of this type were first found in [27], while meandering curved line defects were reported in [28]. In [16, 27], line-defect turbulence, consisting of turbulent states that are mediated by curved line defects which interact in a complicated way, was observed in experiments and numerical simulations.

Period-doubling cascades. The Rössler ODE exhibits a period-doubling cascade, and one may therefore expect to see a cascade of period-doubling bifurcations of spirals in the Rössler PDE (1.3). Higher-order period-doubling of spiral waves has indeed been observed experimentally and numerically in [14, 27]. The conclusions of our spectral analysis remain valid for these bifurcations as they rely only on the spectral properties of the asymptotic wave trains. As discussed in [14, 16], higher-order period-doubling engenders different types of line defects which mediate between different individual wave trains in a period-k orbit. We refer the reader to [14, Figure 6] for an illustration and to [16] for an extensive discussion of the different types of line defects that occur in higher-order period-doubling bifurcations of spiral waves.

Appendix A. Spectra of sources on large bounded domains. We outline the proof of Theorem 3. Suppose that $u_*(x,t)$ is a source on $(-L,L)$ obtained from Theorem 1 as the

[2]We continue to refer to curved defects that accommodate phase differences between adjacent wave trains as "line" defects even though they are not straight lines.

concatenation of a source on $\mathbb{R}$ and two boundary sinks. The linearization of (3.4) about u_* is given by

$$\begin{aligned} v_t &= Dv_{xx} + f_u(u_*(x,t);\mu)v, \qquad x \in (-L,L),\\ 0 &= v_x(\pm L,t), \end{aligned}$$

and we denote its evolution by Φ'_t. Floquet multipliers ρ can be found by seeking nontrivial solutions v_0 to $\Phi'_T v_0 = \rho v_0$, where $T = 2\pi/\omega_*$ denotes the temporal period of the source u_*. Writing

$$v(x,t) = \mathrm{e}^{\Lambda t}u(x,t)$$

for $v(x,t) = \Phi'_t v_0$, we see that $\rho = \mathrm{e}^{\Lambda T}$ is a Floquet multiplier if and only if $u(x,t)$ satisfies

$$\begin{aligned} u_t &= Du_{xx} + f_u(u_*(x,t);\mu)u - \Lambda u, \qquad x \in (-L,L),\\ 0 &= u_x(\pm L,t), \end{aligned} \tag{A.1}$$

with $u(x,t)$ being T-periodic in t. As in [38, section 4.1], we write (A.1) as

$$\begin{pmatrix} u_x \\ v_x \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ D^{-1}[\partial_t - f_u(u_*(x,t);\mu) + \Lambda] & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}, \tag{A.2}$$

with $\mathbf{u} = (u,v) \in \mathcal{X} := H^{1/2}_{\mathrm{per}}(0,T) \times L^2_{\mathrm{per}}(0,T)$ for all x, together with the boundary conditions $\mathbf{u}(\pm L) \in H^{1/2}_{\mathrm{per}}(0,T) \times \{0\}$.

We want to prove that the Floquet spectrum of the truncated source u_* is the union of two disjoint sets: One of these approaches the absolute spectrum of the asymptotic wave trains $u_{\mathrm{wt}}(k_*x - \omega_* t)$ in the symmetric Hausdorff distance as $L \to \infty$, while the other one converges to the union of the extended point spectra of the source on $\mathbb{R}$ and the two boundary sinks. This issue has previously been addressed in [34] in the case where the linearized problem (A.2) is an ODE.

The convergence proof for the absolute spectrum in [34, section 5.3] involves only exponential dichotomies and Lyapunov–Schmidt reduction and therefore carries over immediately to (A.2) once the absolute spectrum of the wave trains u_{wt} is identified. For constant-coefficient problems

$$\mathbf{u}_x = \mathcal{A}(\Lambda)\mathbf{u}, \qquad \mathbf{u} \in \mathbb{R}^{2n},$$

the absolute spectrum is given by

$$\Sigma_{\mathrm{abs}} = \{\Lambda \in \mathbb{C};\ \mathrm{Re}\,\nu_n = \mathrm{Re}\,\nu_{n+1}\},$$

where $\nu_j = \nu_j(\Lambda)$ with $j = 1,\dots,2n$ are the eigenvalues of the matrix $\mathcal{A}(\Lambda)$, ordered with increasing real part. The corresponding definition for (A.2) uses spatial Floquet exponents instead of eigenvalues. We consider the asymptotic $2\pi/k_*$-periodic system

$$\mathbf{u}_x = \begin{pmatrix} 0 & 1 \\ D^{-1}[\partial_t - f_u(u_{\mathrm{wt}}(k_*x - \omega_* t);\mu) + \Lambda] & 0 \end{pmatrix} \mathbf{u}$$

whose spatial Floquet exponents ν are found by seeking solutions $\mathbf{u} \in \mathcal{X}$ of the form

$$\mathbf{u}(x,t) = \mathrm{e}^{\nu x}\mathbf{u}_0(k_* x - \omega_* t),$$

where $\mathbf{u}_0$ is 2π-periodic in its argument. As shown in [37, Proposition 2.10 and section 4] or [38, section 3.4], there are infinitely many spatial Floquet exponents $\nu_j(\Lambda)$ for each fixed Λ which, alternatively, can also be found as roots ν of the function $\mathcal{D}(\Lambda, \nu)$. Ordering the resulting roots ν_j by increasing real part, we end up with the absolute spectrum (2.19) of the wave trains in the laboratory frame. With this identification, the proofs given in [34, section 5.3] for the absolute spectrum carry over to (A.2).

It remains to prove that the remaining spectrum converges to the union of the extended point spectra of the source on $\mathbb{R}$ and the boundary sinks. There are two different proofs that give this result: First, we may invoke [32], where the spectrum of concatenated multi-pulses was investigated, using again only exponential dichotomies and Lyapunov–Schmidt reduction. An alternative proof uses the same topological winding-number arguments based on Evans functions as in [34, section 4.3] but now applied to a finite-dimensional Galerkin approximation of (A.2). It is a consequence of the results proved in [21, 37] that a sufficiently high-dimensional Galerkin approximation captures all eigenvalues of the truncated source.

REFERENCES

[1] R. E. Amritkar and P. M. Gade, *Wavelength doubling bifurcations in coupled map lattices*, Phys. Rev. Lett., 70 (1993), pp. 3408–3411.

[2] D. Barkley, *Euclidean symmetry and the dynamics of rotating spiral waves*, Phys. Rev. Lett., 72 (1994), pp. 164–167.

[3] D. Barkley, ezspiral*: A Code for Simulating Spiral Waves*, University of Warwick, Coventry, UK, 2002.

[4] J. Davidsen, R. Erichsen, R. Kapral, and H. Chaté, *From ballistic to Brownian vortex motion in complex oscillatory media*, Phys. Rev. Lett., 93 (2004), 018305.

[5] E. Doedel, A. R. Champneys, T. F. Fairgrieve, Y. A. Kuznetsov, B. Sandstede, and X. Wang, auto97*: Continuation and Bifurcation Software for Ordinary Differential Equations (with* homcont*)*, Technical report, Concordia University, Montreal, Quebec, Canada, 1997.

[6] A. Doelman, B. Sandstede, A. Scheel, and G. Schneider, *The dynamics of modulated wave trains*, Mem. Amer. Math. Soc., to appear.

[7] B. Fiedler, B. Sandstede, A. Scheel, and C. Wulff, *Bifurcations from relative equilibria of non-compact group actions: Skew products, meanders, and drifts*, Doc. Math., 1 (1996), pp. 479–505.

[8] R. A. Gardner, *On the structure of the spectra of periodic travelling waves*, J. Math. Pures Appl. (9), 72 (1993), pp. 415–439.

[9] E. A. Glasman, A. A. Golovin, and A. A. Nepomnyashchy, *Instabilities of wavy patterns governed by coupled Burgers equations*, SIAM J. Appl. Math., 65 (2004), pp. 230–251.

[10] A. A. Golovin, B. J. Matkowsky, A. Bayliss, and A. A. Nepomnyashchy, *Coupled KS–CGL and coupled Burgers–CGL equations for flames governed by a sequential reaction*, Phys. D, 129 (1999), pp. 253–298.

[11] A. A. Golovin, A. A. Nepomnyashchy, and B. J. Matkowsky, *Traveling and spiral waves for sequential flames with translation symmetry: Coupled CGL–Burgers equations*, Phys. D, 160 (2001), pp. 1–28.

[12] M. Golubitsky, V. LeBlanc, and I. Melbourne, *Meandering of the spiral tip—an alternative approach*, J. Nonlinear Sci., 7 (1997), pp. 557–586.

[13] M. Golubitsky, I. Stewart, and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory* II, Springer-Verlag, New York, 1988.

[14] A. Goryachev, H. Chaté, and R. Kapral, *Synchronization defects and broken symmetry in spiral waves*, Phys. Rev. Lett., 80 (1998), pp. 873–876.
[15] A. Goryachev, H. Chaté, and R. Kapral, *Transitions to line-defect turbulence in complex oscillatory media*, Phys. Rev. Lett., 83 (1999), pp. 1878–1881.
[16] A. Goryachev, H. Chaté, and R. Kapral, *Synchronization defect lines*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 10 (2000), pp. 1537–1564.
[17] W. Jahnke, W. E. Skaggs, and A. T. Winfree, *Chemical vortex dynamics in the Belousov-Zhabotinsky reaction and in the two-variable Oregonator mode*, J. Chem. Phys., 93 (1989), pp. 740–749.
[18] T. Kapitula and B. Sandstede, *Stability of bright solitary-wave solutions to perturbed nonlinear Schrödinger equations*, Phys. D, 124 (1998), pp. 58–103.
[19] H. Kielhöfer, *Bifurcation Theory: An Introduction with Applications to PDEs*, Springer-Verlag, New York, 2004.
[20] G. Li, Q. Ouyang, V. Petrov, and H. L. Swinney, *Transition from simple rotating chemical spirals to meandering and travelling spirals*, Phys. Rev. Lett., 77 (1996), pp. 2105–2108.
[21] G. J. Lord, D. Peterhof, B. Sandstede, and A. Scheel, *Numerical computation of solitary waves in infinite cylindrical domains*, SIAM J. Numer. Anal., 37 (2000), pp. 1420–1454.
[22] P. C. Matthews and S. M. Cox, *Pattern formation with a conservation law*, Nonlinearity, 13 (2000), pp. 1293–1320.
[23] A. Mielke, *A spatial center manifold approach to steady bifurcations from spatially periodic patterns*, in Dynamics in Dissipative Systems: Reductions, Bifurcations and Stability, Pitman Res. Notes Math. Ser. 352, G. Dangelmayr, B. Fiedler, K. Kirchgässner, and A. Mielke, eds., Pitman, Boston, 1996, pp. 209–262.
[24] S. Nettesheim, A. von Oertzen, H. H. Rotermund, and G. Ertl, *Reaction diffusion patterns in the catalytic CO-oxidation on Pt*(110)*-front propagation and spiral waves*, J. Chem. Phys., 98 (1993), pp. 9977–9985.
[25] F. W. J. Olver, *Asymptotics and Special Functions*, Academic Press, New York, 1974.
[26] Q. Ouyang and J. M. Flesselles, *Transition from spirals to defect turbulence driven by a convective instability*, Nature, 379 (1996), pp. 143–146.
[27] J.-S. Park and K. J. Lee, *Complex periodic spirals and line-defect turbulence in a chemical system*, Phys. Rev. Lett., 83 (1999), pp. 5393–5396.
[28] J.-S. Park and K. J. Lee, *Formation of spiraling line-defect and its meandering transition in a period-2 medium*, Phys. Rev. Lett., 88 (2002), 224501.
[29] J.-S. Park, S.-J. Woo, and K. J. Lee, *Transverse instability of line defects of period-2 spiral waves*, Phys. Rev. Lett., 93 (2004), 098302.
[30] D. Peterhof, B. Sandstede, and A. Scheel, *Exponential dichotomies for solitary wave solutions of semilinear elliptic equations on infinite cylinders*, J. Differential Equations, 140 (1997), pp. 266–308.
[31] J. D. M. Rademacher, B. Sandstede, and A. Scheel, *Computing Absolute and Essential Spectra Using Continuation*, preprint, 2005.
[32] B. Sandstede, *Stability of multiple-pulse solutions*, Trans. Amer. Math. Soc., 350 (1998), pp. 429–472.
[33] B. Sandstede and A. Scheel, *Essential instability of pulses and bifurcations to modulated travelling waves*, Proc. Roy. Soc. Edinburgh Sect. A, 129 (1999), pp. 1263–1290.
[34] B. Sandstede and A. Scheel, *Absolute and convective instabilities of waves on unbounded and large bounded domains*, Phys. D, 145 (2000), pp. 233–277.
[35] B. Sandstede and A. Scheel, *Absolute versus convective instability of spiral waves*, Phys. Rev. E (3), 62 (2000), pp. 7708–7714.
[36] B. Sandstede and A. Scheel, *Essential instabilities of fronts: Bifurcation and bifurcation failure*, Dyn. Syst., 16 (2001), pp. 1–28.
[37] B. Sandstede and A. Scheel, *On the structure of spectra of modulated travelling waves*, Math. Nachr., 232 (2001), pp. 39–93.
[38] B. Sandstede and A. Scheel, *Defects in oscillatory media: Toward a classification*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 1–68.
[39] B. Sandstede and A. Scheel, *Evans function and blow-up methods in critical eigenvalue problems*, Discrete Contin. Dyn. Syst., 10 (2004), pp. 941–964.

[40] B. Sandstede and A. Scheel, *Basin boundaries and bifurcations near convective instabilities: A case study*, J. Differential Equations, 208 (2005), pp. 176–193.
[41] B. Sandstede and A. Scheel, *Curvature effects on spiral spectra: Generation of point eigenvalues near branch points*, Phys. Rev. E (3), 73 (2006), 016217.
[42] B. Sandstede and A. Scheel, *Dynamics of Spiral Waves*, in preparation.
[43] B. Sandstede, A. Scheel, and C. Wulff, *Dynamics of spiral waves on unbounded domains using center-manifold reduction*, J. Differential Equations, 141 (1997), pp. 122–149.
[44] B. Sandstede, A. Scheel, and C. Wulff, *Bifurcations and dynamics of spiral waves*, J. Nonlinear Sci., 9 (1999), pp. 439–478.
[45] P. Wheeler and D. Barkley, *Computation of spiral spectra*, SIAM J. Appl. Dyn. Syst., 5 (2006), pp. 157–177.
[46] A. Winfree, *Varieties of spiral wave behavior: An experimentalist's approach to the theory of excitable media*, Chaos, 1 (1991), pp. 303–334.
[47] M. Yoneyama, A. Fujii, and S. Maeda, *Wavelength-doubled spiral fragments in photosensitive monolayers*, J. Amer. Chem. Soc., 117 (1995), pp. 8188–9191.
[48] M. Zhan and R. Kapral, *Model for line defects in complex-oscillatory spiral waves*, Phys Rev. E (3), 72 (2005), 046221.
[49] L. Q. Zhou and Q. Ouyang, *Spiral instabilities in a reaction-diffusion system*, J. Phys. Chem. A, 105 (2001), pp. 112–118.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 3, pp. 549–575

Symmetry and Symmetry-Breaking for a Wave Equation with Feedback*

Prashant G. Mehta†, Gregory Hagen‡, and Andrzej Banaszuk§

Abstract. This paper is concerned with model-independent approaches for the analysis of inception and suppression of oscillations in certain feedback interconnections arising in aerospace and industrial applications. One of the subsystems in the interconnection is known and assumed here to be the wave equation on the circle. The dynamic model of the other subsystem is uncertain, and the approach assumes only its structure, namely its symmetry properties. We show that *only* the structure (skew-symmetry) of the feedback can be used to explain the instability, and manipulation of the structure (mistuning) can be used to suppress the instability.

Key words. dynamical systems, jet engines, wave equation, symmetry, bifurcation, design of dynamics, mistuning

AMS subject classifications. 74H60

DOI. 10.1137/060666044

1. Introduction.

In several applications, feedback interconnections between dynamical subsystems can lead to large sensitivity, instability, and bifurcation to (often) detrimental behavior. Examples include thermoacoustic (combustion) instabilities where feedback coupling arises between acoustics and combustion [23], aeroacoustic instabilities with feedback coupling between acoustics and flow [58], and flutter instabilities with fluid-structure coupling [26]. In these examples, the dynamic model of one of the subsystems (for, e.g., combustion and fluid dynamics) in the feedback loop is complex and very often CFD-based. As a result, analysis, let alone suppression and control, of instability becomes a challenge.

This paper considers the problem of a feedback interconnection, shown in Figure 1, where the *simpler* known subsystem is a wave equation on the circle. The motivation comes from several examples of oscillatory phenomena in aerospace and industrial applications with circular geometries; Figure 2 depicts a few key examples, such as fan blade flutter, rotating stall, aeroacoustic, and thermoacoustic instabilities in jet engines. Of particular interest is the thermoacoustic instability where the wave equation is a simplified model of acoustics and the feedback subsystem models the effects due to combustion: turbulent reacting flow, multiphase fuel transport, chemical kinetics, and distributed heat release [12]. Needless to say, even the CFD-based models of combustion are at best approximations and often computationally

*Received by the editors July 25, 2006; accepted for publication (in revised form) by R. Murray March 14, 2007; published electronically July 11, 2007. This research was supported by United Technologies Research Center and AFOSR (grant FA9550-04-C-0042).
http://www.siam.org/journals/siads/6-3/66604.html
†Department of Mechanical Science & Engineering and the Coordinated Sciences Laboratory, University of Illinois at Urbana-Champaign, 1206 W. Green Street, Urbana, IL 61801 (mehtapg@uiuc.edu).
‡Systems Department, United Technologies Research Center, 411 Silver Lane, East Hartford, CT 06040 (hagengs@utrc.utc.com).
§Systems Department, United Technologies Research Center, 411 Silver Lane, East Hartford, CT 06040 (banasza@utrc.utc.com).

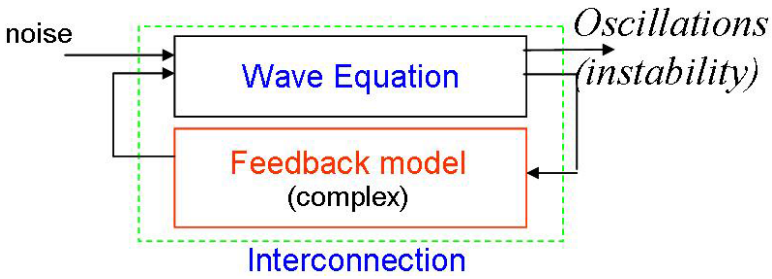


Figure 1. *Interconnection of a wave equation on a circle in feedback with a complex dynamical system.*

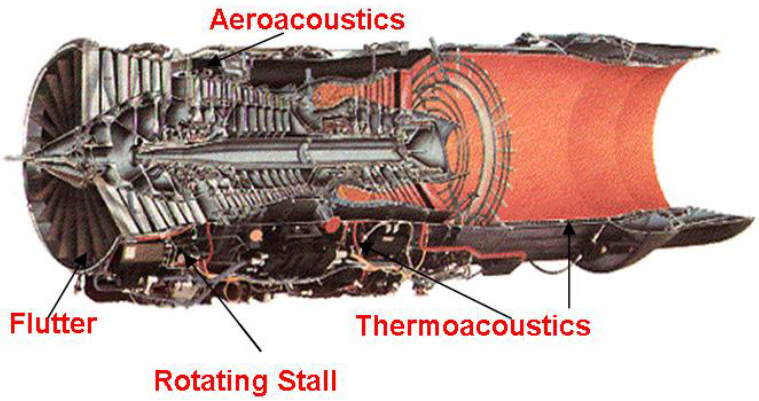


Figure 2. *A jet engine with cylindrical geometry: the dynamics can often be approximately described using a wave equation on a circle with dynamic feedback.*

intractable for realistic operating conditions and geometries. Analysis approaches for instabilities range from computational methods [57, 46] to the use of simplified ad hoc models for predicting qualitative trends and control [4, 15]. Both the approaches typically suffer from large uncertainty in models, do not provide robust explanation of the observed instabilities and the resulting dynamic behavior, and are not applicable to design methods for suppression of instabilities.

Current practical remedies for thermoacoustic oscillations include (a) operating in design space where these instabilities are mild [11], (b) using diagnostics such as monitoring for blade fatigue [24], (c) passive fixes such as tuned Helmholtz resonators and liners that serve to dissipate acoustic energy in the case of combustion instability [31, 9, 54], and (d) active feedback control approaches [17, 50, 8]. These passive approaches suffer from increased weight, cost, and performance penalties. Active control approaches require accurate models for controller design, and most importantly suffer from fundamental performance limitations due to physical factors such as delay, actuator bandwidth, etc. [16, 7].

In this paper, we propose a symmetry/structure-based approach for the analysis of interconnection shown in Figure 1. Our two main contributions are that (a) the symmetry properties alone of the feedback can be used to explain the inception of instability, and (b) manipulation (mistuning) of certain structural aspects can be used to suppress the instability. The key innovation is that *both* the analysis and control are feedback *model-independent*:

it assumes only the structural properties such as symmetry of the feedback subsystem and not its particular dynamic model. This makes the approach robust to the uncertainty of the complex feedback model, and thus suitable for practical applications; cf. [41].

In addition to the present focus on thermoacoustic oscillations, recent work on symmetry-breaking has focused on the application of minimizing fan blade flutter. Studies concerning stability properties of turbine blade flutter after introduction of spatial nonuniformities has appeared in [10, 47] among many others. Optimal mistuning in arrays of bladed disks has appeared in [45, 53]. In these studies, mistuning of the stiffness of the individual blades was employed to reduce vibrations. The optimal mistuning pattern was determined from large scale optimization problems, where the symmetry properties of the system were used to reduce the size of the optimization problem. Such optimization problems often result in model-specific nonrobust solutions. Furthermore, the role of the symmetry properties of the feedback on instability and its suppression was not addressed.

In analysis of the instability, we use the equivariance (symmetry) properties of the wave equations on the circle. Without explicitly assuming dynamics of the feedback subsystem, the role of its symmetry on instability is investigated. In particular, it is shown that a certain skew-symmetry in feedback can lead to instability in a *robust* manner, irrespective of the dynamics. A bifurcation result is then presented to show that such an instability can lead to a rotating wave with a preferential direction of rotation. This is indeed consistent with the experimental evidence of rotating wave instabilities seen in jet engines with cylindrical geometries. The analysis results are useful because they provide for a model-independent explanation of the instability. In particular, the sign of the feedback gain is not important as long as it has certain skew-symmetric structure. This nonintuitive result arises because of the assumed $O(2)$-equivariance of the wave equation. In fact, the results are relevant to the general case of $O(2)$-equivariant dynamics with feedback. Other well-known examples of such systems where instability analysis is important are also noted.

In suppression of the instability, the idea is to modify the structural aspects of the interconnection model. This is accomplished by introducing precise spatial variations (mistuning) in the mean properties such as wave speed of the wave equation. It is shown that just as skew-symmetry in feedback provides a dynamic-independent explanation of the instability, mistuning provides for its robust suppression by reversing the effect of skew-symmetry irrespective of the feedback dynamics. A physical interpretation in terms of damping augmentation is provided, whereby damped system modes are shown to stabilize the unstable modes. For a given skew-symmetric feedback, there is a fundamental limitation on the efficacy of mistuning beyond which it becomes ineffective. This limit is used to establish guidelines for optimal design modification.

Symmetry-based methods have long been used for representation and analysis of equivariant (symmetric) dynamical systems. Group representation theory [52] provides for a simpler (diagonal) representation of *any* linear dynamical system that commutes with a group action. For nonlinear dynamical systems, methods of equivariant bifurcation theory are by now well developed for the study of (local) bifurcations; cf. [30]. These methods have also proved useful in the study of nonlinear partial differential equations (pde) with symmetry; cf. [49, 35, 29]. In the context of this paper, equivariant bifurcation theory has been used to provide an elegant explanation of wave solutions seen in numerous physical systems with spatial translation

($SO(2)$ or Z^n) and reflection (Z^2) symmetry. Examples include rotating waves in coupled identical oscillator systems [1], traveling waves in translation invariant fluid systems with planar [13] and circular [14] geometries, rotating waves in elastic rods [34] and pipes [2], spiral waves in reaction-diffusion [22], and in pattern-formation systems [19]. General results for the local bifurcation analysis of dynamical systems with $O(2)$ spatial symmetry appear in [25, 28, 30, 18, 21, 29].

Symmetry-based methods are powerful because they provide for *model-independent* conclusions on dynamics. However, these methods only delineate the possibilities in a model-independent fashion. The *actual* solution behavior is a function of dynamics (gain and phase, damping, and frequencies) present in the problem. By studying feedback interconnection of Figure 1 with a nominally symmetric wave equation, we not only study an important application but also use symmetry methods to obtain results on (feedback) model-independent but dynamically robust behavior. We further combine this with symmetry-breaking ideas, an area that, to the best of our knowledge, has not been investigated within the equivariant bifurcation literature.

The outline of this paper is as follows. In section 2, we describe the interconnection models considered in this paper, formulate the problem, and summarize a few example problems where these models and problems arise. In section 3, we study the role of symmetry in sensitivity of the feedback loop, inception of instability, and bifurcation. In section 4, we present symmetry-breaking as a method for stabilization of a class of instabilities, relate it to the symmetry properties of the interconnected system, study its fundamental limitations, and comment on the robustness of both the instability analysis and the approach proposed for its suppression. Finally, we summarize the conclusions in section 5.

2. Problem statement.

2.1. Wave-equation with feedback.

Figure 1 depicts the feedback interconnection of an azimuthally distributed one-dimensional wave equation in feedback with another dynamic subsystem. The interconnection is modeled as the pde

$$\frac{\partial^2 p}{\partial t^2} - a^2(x)\frac{\partial^2 p}{\partial x^2} = \epsilon Q\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial x}\right](p), \tag{2.1}$$

where x denotes the azimuthal angle coordinate, t denotes the time, a is the wave-speed, and ϵ is a real-valued parameter. Q denotes a general nonlinear, dynamic, and distributed feedback with $Q(0) = 0$. $p(x) \equiv 0$ is thus the trivial solution of the interconnection (2.1). Define $L \doteq D_p Q(0)$ to be the linearization of the dynamic feedback at this trivial solution; D_p denotes the derivative with respect to p. Using this notation, the linearization of the interconnection is given by

$$\frac{\partial^2 p}{\partial t^2} - a^2(x)\frac{\partial^2 p}{\partial x^2} = \epsilon L\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial x}\right](p). \tag{2.2}$$

The parameter ϵ will be used to bookkeep $O(\epsilon)$ perturbations in the eigenvalues of the wave equation as a result of the feedback.

In several applications of interest, p represents acoustic pressure, $p_x \doteq \frac{\partial p}{\partial x}$ may be scaled to get acoustic velocity, and Q represents the effect of combustion, fluid, or structural coupling. Specific simple and well-known examples for the feedback term Q include

$$\text{Mean flow:} \quad Q(p) = m \cdot p_x, \tag{2.3}$$

$$\text{Damping:} \quad Q(p) = -b \cdot p_t, \tag{2.4}$$

$$\text{Delay model of combustion:} \quad Q(p) = \frac{\partial q}{\partial t}(p_x(t-\tau)), \tag{2.5}$$

where $\frac{\partial q}{\partial t}$ is the so-called heat release rate, which in the delay model is simply taken to be some function of the acoustic velocity p_x with a delay τ. The mean flow may arise due to swirl, the damping term is used to model acoustic losses on the boundaries due to acoustic liners, and delay arises due to finite speed of fuel-air propagation in combustion models. Note that in applications a reduced order model for the feedback may be a combination of (2.3)–(2.5).

Crucial to the subject of this paper is the symmetry and the structure of the interconnected models. In the absence of any feedback ($Q = 0$), and with $a^2(x) = a_0^2$ a homogeneous constant, the wave equation (2.1) is equivariant with respect to the spatial symmetry group $O(2)$:

$$\begin{aligned} SO(2) &: \tau_\theta[p(x)] = p(x+\theta) \quad \text{for } \theta \in [0, 2\pi], \\ Z^2 &: \sigma[p(x)] = p(-x). \end{aligned} \tag{2.6}$$

The translation symmetry $SO(2)$ arises because (2.1) is spatially homogeneous in x. To see that the equations are Z^2-equivariant, note that

$$\sigma[p_{tt} - a_0^2 p_{xx}] = \left(\frac{\partial^2}{\partial t^2} - a_0^2 \frac{\partial^2}{\partial x^2}\right) \sigma[p]. \tag{2.7}$$

The dynamic feedback Q in general destroys the $O(2)$ equivariance for the interconnected feedback system, and symmetry is not thought to be relevant to many practical applications. Motivated by rotating wave instabilities seen in numerous physical applications arising in jet engines, we use the wave equation with $O(2)$ symmetry as the "organizing center" for our analysis; the term in quotes has been used by Golubitsky to highlight the relevance of symmetry methods in physical problems [27].

In order to study the simplest and perhaps the most relevant case first, we make the additional assumption that the feedback subsystem Q is itself $SO(2)$-equivariant. In particular, this rules out explicit x-dependence in coefficients of Q. The assumption is motivated by the presence of identical components in many feedback subsystems of interest [3]. For example, combustion feedback is due to the presence of many identical combustor elements: flameholders in bluff body stabilized augmentors [5], or swirlers in swirl stabilized combustion chambers [55, 38, 37]. Two special classes of the $SO(2)$-equivariant feedback subsystem of interest include (spatially) local and integral models:

1. Local feedback:

$$Q(p)(x,t) = Q\left[\frac{\partial}{\partial t}, e^{-s\tau}\right] p_x(x,t), \tag{2.8}$$

where Q is a function consisting of time-derivative $\frac{\partial}{\partial t}$ and delay $e^{-s\tau}$ operator. In fact, a delay with static nonlinearity is often a simple model of the combustion feedback [4, 7].

2. Integral feedback:

$$Q(p)(x,t) = \int_0^{2\pi} K(x-\theta)p_t(\theta,t)d\theta, \tag{2.9}$$

where the special form of kernel K arises due to the assumption on the $SO(2)$-equivariance of the operator.

For an $SO(2)$-equivariant Q, it follows that the linearization L, taken about the fully symmetric trivial solution, also commutes with $SO(2)$.

In general, the assumption of $SO(2)$-equivariance is only an idealization. Although the physics in circular geometries such as jet engines is often translation invariant, there are two mechanisms by which symmetry is broken. One, the feedback subsystems such as combustors or blades are often composed of discrete elements. The resulting rotation symmetry group is thus finite: $SO(2)$ is replaced by the discrete rotation group Z^N with N large. For such a case, one can simply repeat the analysis of this paper under the assumption of a discrete symmetry group D^N instead of the continuous group $O(2)$ in (2.6). As sketched in section 3.4, the conclusions of this paper are not expected to be different, provided that the spatial length scale of the instability is much larger than the smallest length scale of discrete symmetry. The second mechanism for breaking symmetry is where the wave equation (the left-hand side of (2.1) or (2.2)) is derived with a spatially inhomogeneous mean. This often introduces parametric inhomogeneities, such as x-dependent wave speed $a(x)$. Influence of some of these asymmetries on instability will be studied in section 4.

There are three topics of interest in this paper:

1. role of symmetry of the dynamic feedback subsystem on the sensitivity of the feedback loop, inception of instability, and possible bifurcations;
2. role of parametric inhomogeneities in the suppression of instability;
3. robustness of these results with respect to the uncertainty in the dynamics of the feedback subsystem.

2.2. Other examples. Other than the wave equation in circular geometries, the $O(2)$ symmetry of subsystems in feedback interconnections is also relevant to other physically important examples where instability is important. Two such examples, the nonlinear Moore–Greitzer model of compressor instability and the planar Couette flow equation, are depicted in Figure 3. The nonlinear Moore–Greitzer model describes the dynamics of the axial flow through the compressor, which are determined by the compressor blade forcing and the overall pressure differential [44]. The model is used for analysis of the so-called rotating stall and surge instabilities observed in axial compressors. The planar Couette flow problem refers to the incompressible two-dimensional flow between two planes that move with a constant velocity relative to each other [51]. The role of instabilities in the transition to turbulence observed in this flow represents a longstanding problem of historical interest. In both these examples, the basic (laminar) solution state is $O(2)$-symmetric, and thus the linearization (of the appropriate equations) taken about this solution commutes with the group action. Figure 3 depicts the symmetry in these two example problems. Since we use the $O(2)$ symmetry

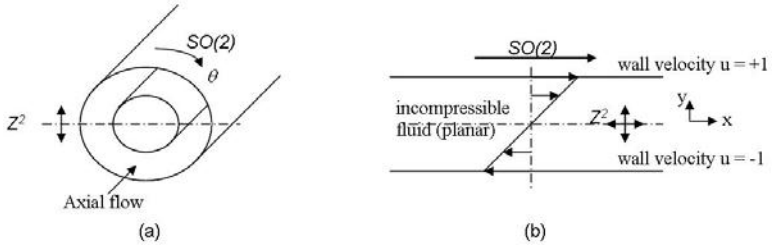


Figure 3. *Schematics of* (a) *the Moore–Greitzer problem where the laminar solution is symmetric with respect to rotations* $\theta \to \theta + c$ *and reflection* $\theta \to -\theta$ *and* (b) *the planar Couette flow problem where the laminar solution is symmetric with respect to translations* $x \to x + c$ *and reflection* $(x, y) \to (-x, -y)$.

as the organizing center for our analysis, the results of this paper are also relevant to these important example problems.

3. Symmetry, instability, and bifurcation.

3.1. Symmetry and eigenvalues.

The eigenvalues of *any* $O(2)$-equivariant dynamic model are double. This is because the so-called irreducible representation of the abstract $O(2)$ group is two-dimensional. For the homogeneous wave equation on the circle, the kth eigenvalue pair is

$$\lambda_k = \pm ika_0, \tag{3.1}$$

and the eigenvectors $\{e^{ik(x\pm a_0 t)}, e^{-ik(x\pm a_0 t)}\}$ span a four-dimensional real vector space. Likewise, the linearized Moore–Greitzer model has double real eigenvalues [33], and the linearization of the planar Couette flow has double real and complex eigenvalues that lie in the complex left half-plane (LHP) for all values of the Reynolds number [42].

Since the eigenvalues are double due to the symmetry of the problem, their double nature will persist, irrespective of the dynamics of the feedback subsystem, as long as the feedback subsystem also shares the $O(2)$-equivariance. In the following section, we study the influence of *only* the symmetry properties of the feedback on these eigenvalues.

3.2. Decomposition of linear feedback dynamics.

We assume that the feedback dynamics are $SO(2)$-equivariant:

$$L\tau_\theta = \tau_\theta L. \tag{3.2}$$

As already noted, if the feedback dynamics are additionally Z^2-equivariant, then the double nature of the eigenvalues persists. In order to study the more general situation, we consider the following decomposition of dynamics. Define

$$L^* = \sigma L \sigma, \tag{3.3}$$

where σ denotes the reflection subgroup as defined in (2.6). It is easy to see that if L is Z^2-equivariant, then $L^* = \sigma L \sigma = \sigma^2 L = L$. Use L^* to consider a decomposition

$$\begin{aligned} L &= \frac{L + L^*}{2} + \frac{L - L^*}{2} \\ &\qquad \text{symmetric} + \text{skew-symmetric}, \end{aligned} \tag{3.4}$$

where the terminology symmetric and skew-symmetric are used because

$$\left(\frac{L+L^*}{2}\right)\sigma = \frac{L\sigma+\sigma L}{2} = \sigma\left(\frac{L+L^*}{2}\right), \tag{3.5}$$

$$\left(\frac{L-L^*}{2}\right)\sigma = \frac{L\sigma-\sigma L}{2} = -\sigma\left(\frac{L-L^*}{2}\right). \tag{3.6}$$

Example 3.1.

1.
$$\left(\frac{\partial}{\partial t}\right)^* = \frac{\partial}{\partial t}, \qquad \left(\frac{\partial}{\partial x}\right)^* = -\frac{\partial}{\partial x} \tag{3.7}$$
are examples of symmetric and skew-symmetric dynamics, respectively. ■

2. Consider the linear integral dynamics
$$L(p)(x) = \int_0^{2\pi} K(x-\theta)p_t(\theta,t)d\theta. \tag{3.8}$$
A simple calculation shows that
$$L^*(p)(x) = \int_0^{2\pi} K(-x+\theta)p_t(\theta,t)d\theta. \tag{3.9}$$
L is symmetric if $K(\theta)$ is an even function of θ and is skew-symmetric if K is an odd function. In general, the decomposition here is given by
(3.10)
$$L(p)(x) = \int \frac{K(x-\theta)+K(-x+\theta)}{2}p_t(\theta,t)d\theta + \int \frac{K(x-\theta)-K(-x+\theta)}{2}p_t(\theta,t)d\theta,$$
which corresponds to decomposition of the kernel K as a sum of even and odd functions. ■

In the following section, we obtain symmetry-dependent conclusions on instability for the linear interconnection (2.2) with $a(x) = a_0$:

$$\frac{\partial^2 p}{\partial t^2} - a_0^2\frac{\partial^2 p}{\partial x^2} = \epsilon L(p). \tag{3.11}$$

For a closed-loop stable interconnection, we will sometimes include an additive input disturbance d on the right-hand side of the equation.

3.3. Symmetry and instability. Equation (2.6) describes a particular *representation* of the abstract groups $SO(2)$ and Z^2 for the vector space $L^2(0,2\pi)$. From group representation theory, a representation defines a unique isotypic decomposition—into the so-called symmetry coordinates—of the vector space [52]. Isotypic decomposition is convenient because it serves to diagonalize any commuting linear operator. For the representation corresponding to $SO(2)$ or $O(2)$, the isotypic decomposition of $L^2(0,2\pi)$ is given by the Fourier modes:

$$L^2 = \bigoplus\nolimits_k V_k, \tag{3.12}$$

where $\bigoplus$ is used to denote the direct sum and where the kth component

$$V_k = \{s_k \sin(kx) + c_k \cos(kx) \ : \ c_k, s_k \in \mathbb{R}\} \tag{3.13}$$

is a two-dimensional vector space in L^2; cf. [29]. If L in (3.11) commutes with the representation of $SO(2)$—as in (3.2)—the diagonalization means that $L : V_k \to V_k$. Taking a Fourier series expansion of the pressure in these symmetry coordinates,

$$p(x) = c_0 + \sum_k s_k \sin(kx) + c_k \cos(kx), \tag{3.14}$$

we simplify the analysis of linear pde (3.11). For the representation (2.6), the induced group action for the Fourier coordinate (s_k, c_k) is given by

$$\begin{aligned} SO(2) : \tau_\theta \begin{bmatrix} s_k \\ c_k \end{bmatrix} &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} s_k \\ c_k \end{bmatrix}, \\ Z^2 : \sigma \begin{bmatrix} s_k \\ c_k \end{bmatrix} &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_k \\ c_k \end{bmatrix}. \end{aligned} \tag{3.15}$$

We denote the restriction as $L_k \doteq L|_{V_k}$. By the property of isotypic decomposition, $L_k : V_k \to V_k$. After taking the Laplace transform, $L_k(s)$ is simply a 2×2 matrix defined with respect to the coordinate (s_k, c_k) of the two-dimensional vector space V_k.

The induced action given by (3.15) together with $SO(2)$-commutativity in (3.2) can be further used to simplify $L_k(s)$. A simple calculation shows that any 2×2 matrix that commutes with $\left[\begin{smallmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{smallmatrix}\right]$ must be of the form

$$L_k(s) = \begin{bmatrix} A_k(s) & -B_k(s) \\ B_k(s) & A_k(s) \end{bmatrix}. \tag{3.16}$$

Finally, we use $\sigma = \left[\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ in (3.15) to construct the decomposition of (3.4):

$$\begin{aligned} L_k(s) \ &= \ A_k(s)I \ + \ B_k(s)J \\ &\qquad \text{symmetric} + \text{skew-symmetric}, \end{aligned} \tag{3.17}$$

where $I = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $J = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$. Substituting this into (3.11), the linear dynamics in subspace V_k are given by the matrix equation

$$(s^2 + a_0^2 k^2) I \begin{bmatrix} s_k \\ c_k \end{bmatrix} = \epsilon (A_k(s) I + B_k(s) J) \begin{bmatrix} s_k \\ c_k \end{bmatrix} + \begin{bmatrix} d_s \\ d_c \end{bmatrix}, \tag{3.18}$$

where d_s and d_c are the Fourier components of the input disturbance d. $A_k(s)$ and $B_k(s)$ model the symmetric and skew-symmetric components, respectively, of the linearization of any general $SO(2)$-equivariant feedback subsystem Q.

Example 3.2.

1. For a symmetric $L = -\frac{\partial}{\partial t}$, $A_k(s) = -s$ and $B_k(s) = 0$.
2. For a skew-symmetric $L = -\frac{\partial}{\partial x}$, $A_k(s) = 0$ and $B_k(s) = 1$.

3. For $L(p)(x) = \int_0^{2\pi} K(x-\theta)p_t(\theta,t)d\theta$, let $K(x) = K_0 + \frac{1}{\pi}\sum_k a_k \cos(kx) + b_k \sin(kx)$; then a simple calculation shows that $A_k(s) = a_k \cdot s$ and $B_k = -b_k \cdot s$. For an even kernel K the integral is symmetric and $B_k = 0$, and for an odd kernel the integral is skew-symmetric and $A_k = 0$. ■

In the following, we study perturbations of the *nominal* eigenvalues $\lambda_k = \pm i a_0 k$ of the homogeneous wave equation (see (3.1)) as a result of the feedback $A_k(s)I + B_k(s)J$. In particular, destabilization of these eigenvalues is of interest. The eigenvalues of the closed loop (3.11), referred to as the *perturbed* kth eigenvalues, are given by roots of the determinant equation

$$\begin{vmatrix} s^2 + a_0^2k^2 - \epsilon A_k(s) & \epsilon B_k(s) \\ -\epsilon B_k(s) & s^2 + a_0^2k^2 - \epsilon A_k(s) \end{vmatrix} = 0, \tag{3.19}$$

from which follow

$$s^2 + a_0^2k^2 = \epsilon(A_k(s) + iB_k(s)), \tag{3.20}$$

$$s^2 + a_0^2k^2 = \epsilon(A_k(s) - iB_k(s)). \tag{3.21}$$

Equations (3.20) and (3.21) have complex conjugate roots; i.e., s is a root of (3.20) if and only if $\bar{s}$ is a root of (3.21). As a result, it suffices to evaluate roots of (3.20) alone. Since perturbation to the nominal eigenvalues $\pm i a_0 k$ is of special interest, we express these roots as regular perturbations in the small parameter ϵ:

$$\begin{aligned} s_+ &= \epsilon r_+ + i(a_0k + \epsilon q_+) + o(\epsilon), \\ s_- &= \epsilon r_- + i(-a_0k + \epsilon q_-) + o(\epsilon). \end{aligned} \tag{3.22}$$

The real parts $r_\pm$ correspond to the roots $s_\pm$ of (3.20). We note that in general $s_\pm$ need not be complex conjugate.

In the absence of skew-symmetric feedback ($B_k(s) = 0$), (3.20) has real coefficients and its roots are complex conjugate. The eigenvalues of (3.19) are thus double and move as a pair: either both the eigenvalues move into the left half-plane or they both move into the right half-plane. This depends upon the particular dynamic model $A_k(s)$. For example, $A_k(s) = s$ is destabilizing while $A_k(s) = -s$ is stabilizing. The following theorem summarizes the conclusion on the eigenvalue movement on account of symmetric and skew-symmetric feedbacks.

Theorem 3.3. *Consider the eigenvalue problem for the linear pde (3.11), where the linear feedback L commutes with the representation of $SO(2)$ in (2.6). In the limit as $\epsilon \to 0$,*

1. *the sum of the real part of the kth eigenvalues*

$$r_+ + r_- = \frac{\mathrm{Imag}(A_k(ia_0k))}{a_0k}; \tag{3.23}$$

2. *the difference of the real part of the kth eigenvalues*

$$r_+ - r_- = \frac{\mathrm{Real}(B_k(ia_0k))}{a_0k}. \tag{3.24}$$

Proof. Substituting the regular perturbation (3.22) into (3.20), the $O(\epsilon)$ balance yields

$$O(\epsilon) : 2r_+ a_0 k = \text{Real}(B_k(ia_0k)) + \text{Imag}(A_k(ia_0k)), \tag{3.25}$$

$$-2r_- a_0 k = \text{Real}(B_k(-ia_0k)) - \text{Imag}(A_k(ia_0k)). \tag{3.26}$$

Equations (3.23)–(3.24) follow by subtracting and adding the two expressions, respectively. Thus, the cumulative damping of the system is fixed by symmetric feedback $A_k(s)$, while the skew-symmetric feedback merely serves to exchange damping. ∎

For a purely skew-symmetric feedback with $A_k(s) = 0$,

$$r_+ + r_- = 0, \tag{3.27}$$

i.e., to $O(\epsilon)$ the sum of real parts of the roots due to an *arbitrary* skew-symmetric feedback is 0. As a result, the eigenvalues of (3.19) split: if one of the eigenvalue pair moves into the LHP, then the other must move into the right half-plane.

Equations (3.23)–(3.24) can also be derived by considering a broadband approximation of the feedback, where $A_k(s) \approx A_k(\pm ia_0k)$ and $B_k(s) \approx B_k(\pm ia_0k)$ are used for the computations of the kth eigenvalues. Such a broadband approximation is often used in determination of eigenvalues for cases where the phase of the feedback subsystem rolls off "slowly" near the natural frequency.

Remark 3.4. Depending upon the sign of $\text{Imag}(A_k(ia_0k))$ in (3.23), the symmetric feedback can either stabilize or destabilize the interconnection's eigenvalues. In contrast, the skew-symmetric feedback *always* causes destabilization—movement of one eigenvalue pair into the right half-plane—for the generic case of $\text{Real}(B_k(ia_0k)) \neq 0$.

Example 3.5. Consider the wave equation

$$\frac{\partial^2 p}{\partial t^2} - a_0^2 \frac{\partial^2 p}{\partial x^2} = \epsilon_a \frac{\partial p}{\partial t} + \epsilon_b \frac{\partial p}{\partial x} \tag{3.28}$$

$$\text{symmetric} + \text{skew-symmetric}.$$

Figure 4(a) depicts the nominal eigenvalues $\pm ia_0k$ without any feedback ($\epsilon_a = \epsilon_b = 0$). In the presence of feedback, $A_k(s) = \epsilon_a s$, $B_k(s) = \epsilon_b$, and the kth eigenvalue is determined by the roots of

$$(s^2 + a_0^2 k^2) I = \epsilon_a s \cdot I + \epsilon_b \cdot J. \tag{3.29}$$

As shown in Figure 4(b), the symmetric feedback can destabilize both eigenvalues, and the sum of the real part of the eigenvalues is in fact given by

$$r_+ + r_- = \epsilon_a = \frac{\text{Imag}(A_k(ia_0k))}{a_0 k}. \tag{3.30}$$

As shown in Figure 4(c), the presence of the skew-symmetric feedback splits the nominal eigenvalues such that

$$r_+ - r_- \approx \frac{\epsilon_b}{a_0 k} = \frac{\text{Real}(B_k(ia_0k))}{a_0 k}. \tag{3.31}$$

∎

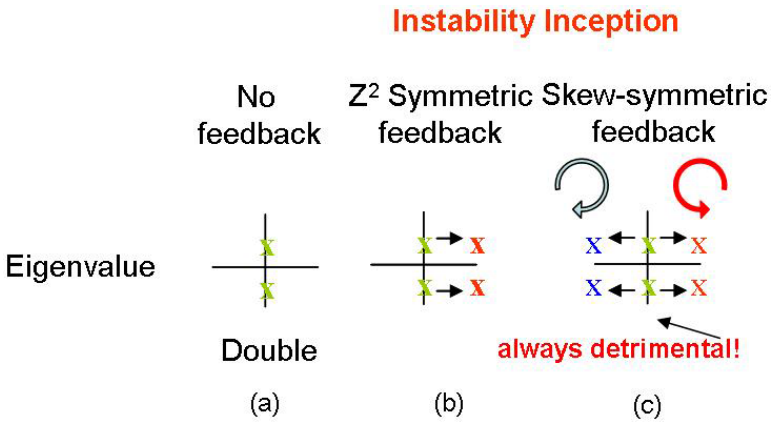


Figure 4. *The kth eigenvalues for* (a) *the nominal homogeneous wave equation,* (b) *with symmetric destabilizing feedback, and* (c) *with arbitrary skew-symmetric feedback.*

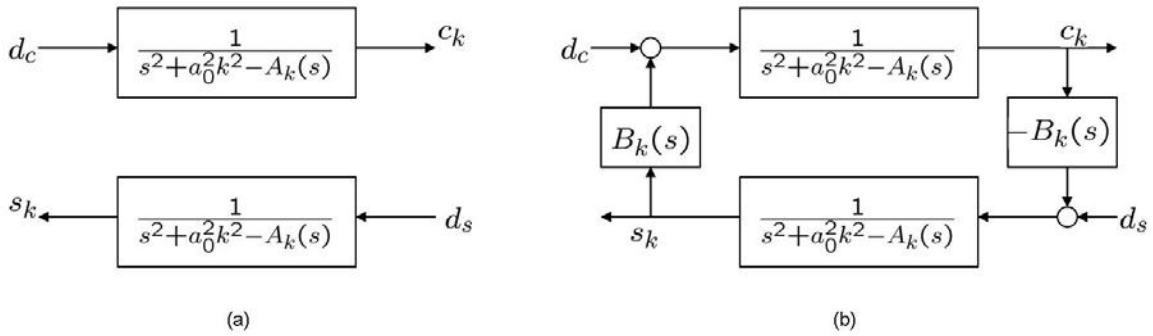


Figure 5. *The closed-loop dynamics in V_k:* (a) *with symmetric feedback $A_k(s)$ alone, and* (b) *together with skew-symmetric feedback $B_k(s)$. d_c and d_s are the components of the input disturbance (right-hand side of (3.11)) in V_k.*

Figure 5 provides another interpretation of the role of symmetric and skew-symmetric feedback in the kth Fourier subspace V_k. We begin by noting that the $SO(2)$ symmetry implies that there is no coupling between dynamics in distinct Fourier subspaces. With symmetric feedback, there are two identical uncoupled closed-loop dynamical systems corresponding to the two Fourier modes $\{\cos(kx), \sin(kx)\}$ in V_k. The individual systems capture the feedback coupling of the homogeneous wave equation with $A_k(s)$. The presence of skew-symmetric feedback $B_k(s)$ serves to couple these two systems. Although the strength of the coupling depends upon $B_k(s)$, it *always* destabilizes one of the two Fourier modes. We refer to such a destabilization (instability) as *robust* because the conclusion is feedback model-independent. It depends only upon the (skew-symmetric) structure of the feedback. Figure 5 depicts the dynamical systems in V_k together with input disturbance. Such a model is applicable for description of oscillations in a stable regime [7]. The role of the skew-symmetric feedback here is to make the closed loop sensitive by moving an eigenvalue pair close to the imaginary axis. Sensitivity of the loop can then lead to large oscillations in the presence of disturbance.

3.4. Discrete symmetry. In jet engine applications, the feedback subsystems are often composed of discrete elements (combustors, blades, etc.). The resulting rotation symmetry group is thus finite. $SO(2)$ is replaced by the discrete rotation group Z^N:

$$Z^N : \tau_{\theta_s}[p(x)] = p(x + \theta_s), \tag{3.32}$$

where $\theta_s = \frac{2\pi s}{N}$ for $s = 0, 1, \ldots, N-1$, and N denotes the number of discrete elements (typically large). Equation (3.2) is replaced by the condition of Z^N-equivariance:

$$L\tau_{\theta_s} = \tau_{\theta_s} L. \tag{3.33}$$

The model-independent conclusions of the preceding section depended upon the isotypic decomposition of $L^2[0, 2\pi]$ into Fourier subspace V_k. As a result of this decomposition, closed-loop dynamics can easily be analyzed in decoupled subspaces (see Figure 5). With discrete symmetry, such a fine decomposition no longer holds, but the conclusion of Theorem 3.3 and Figure 5 is still valid, provided that the wave number k of instability is small compared to N. This is because of the following lemma, which clarifies the nature of coupling between the Fourier subspaces.

Lemma 3.6. *Consider the dynamics of the closed-loop (3.11), where the linear feedback L commutes with the representation of Z^N in (3.32). A necessary condition for the kth Fourier subspace V_k to couple, via feedback, with the mth Fourier subspace V_m is*

$$|k \pm m| = 0 \mod N. \tag{3.34}$$

Proof. For $p(x) \in V_k(x)$, consider a Fourier expansion $L(p)(x) = \sum_m a_m e^{imx}$. Using the commutativity (3.33), one obtains

$$a_m e^{i(m-k)\theta_s} = a_m \quad \text{for } s = 0, 1, \ldots, N-1, \tag{3.35}$$

which implies the condition in (3.34) or $a_m = 0$. Since the nominal wave equation is $SO(2)$-equivariant, any coupling arises due to the feedback alone. The result follows. ∎

In order to analyze the effect of discrete symmetry, we assume that $k \ll N$. Such a condition is consistent with applications where the number of discrete elements (N) is typically large compared to the wave number (k) of the global instability mode. Using the result of Lemma 3.6, the feedback serves to couple the kth Fourier mode with the $(k \pm mN)$th for $m = 1, \ldots$. Using the wave equation, the nominal eigenvalues of these modes are $\pm i a_0 k$ and $\pm i a_0 (k \pm mN)$, respectively. For $k \ll N$, the difference in these eigenvalues is large ($> a_0 N$), and the insensitivity of the eigenvalue locations to such higher mode couplings then follows from root locus arguments.

3.5. Bifurcation. Post linear destabilization, a nonlinearity Q in (2.1) can result in a limit cycle. In the context of the preceding analysis, there are two kinds of local bifurcation results corresponding to the symmetric and skew-symmetric cases. In this section, we summarize the Hopf bifurcation result for the skew-symmetric case. In order to facilitate the bifurcation analysis, we consider a damped version of the interconnection:

$$\frac{\partial^2 p}{\partial t^2} + b\frac{\partial p}{\partial t} - a_0^2 \frac{\partial^2 p}{\partial x^2} = \epsilon Q\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial x}\right](p), \tag{3.36}$$

where $b > 0$ and ϵQ denotes a nonlinear feedback term. As before, L denotes the linearization of Q taken about the trivial solution. It is assumed to be skew-symmetric. We further assume that the resulting linearization of (3.36),

$$\frac{\partial^2 p}{\partial t^2} + b\frac{\partial p}{\partial t} - a_0^2\frac{\partial^2 p}{\partial x^2} = \epsilon L\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial x}\right](p), \tag{3.37}$$

yields a Fredholm operator of index 0 (the dimension of its eigenspace is finite-dimensional), and that the kth eigenvalue pair crosses the imaginary axis with a nonzero speed at the critical parameter value of $\epsilon = \epsilon_0$. The general bifurcation result then follows as a Hopf bifurcation for pdes via the method of Lyapunov and Schmidt.

Theorem 3.7. *Consider the nonlinear wave equation* (3.36) *together with its linearization* (3.37)*, where L is assumed to be skew-symmetric. Equation* (3.37) *is assumed to be Fredholm. Let $s(\epsilon)$ denote the kth complex eigenvalue pair and $r(\epsilon)$ its real part. Assume that*

1. *$s(\epsilon_0)$ is purely imaginary,*
2. *no other eigenvalues of* (3.37) *lie on the imaginary axis, and*
3. *the speed of the real part of eigenvalue $r'(\epsilon_0)$ is nonzero.*

Then there is a unique branch of nontrivial rotating wave solutions $p(x - at)$ bifurcating at the critical value of parameter $\epsilon = \epsilon_0$. $a(\epsilon)$ is the wave-speed, such that $a(\epsilon_0)$ is either a_0 or $-a_0$. To leading order, the rotating wave solution is given as

$$p = Ce^{ik(x-at)} + h.o.t., \tag{3.38}$$

where $e^{ik(x-at)}$ is the eigenfunction of the linearized (3.37) *at $\epsilon = \epsilon_0$, "h.o.t." stands for higher order terms, and the amplitude C satisfies one of the two normal forms*

$$\begin{aligned}(\epsilon - \epsilon_0)C + C^3 &= 0,\\ (\epsilon - \epsilon_0)C - C^3 &= 0,\end{aligned} \tag{3.39}$$

where the sign depends upon the nonlinearity Q.

The proof follows from the general Hopf bifurcation theorem for nonlinear pdes with a Fredholm linearization; cf. [36]. The bifurcation theorem provides a nonlinear counterpart to the linear destabilization result because of skew-symmetric feedback.

3.6. Couette flow example. The nonlinear equations of motion for planar Couette flow are $O(2)$-equivariant with a certain representation of $O(2)$, as indicated in Figure 3. The linearization taken about the fully symmetric laminar solution thus commutes with the group action. As a result, the eigenvalues of the planar Couette flow's linearization are double. Furthermore, these eigenvalues are known to reside in the LHP for all values of the Reynolds number [48].

As an example of destabilization caused by introduction of a skew-symmetric feedback, consider a perturbation of the planar Couette flow due to the planar Poiseuille flow (with a parabolic velocity profile). Figure 6(a) depicts the schematic of the planar Poiseuille–Couette flow. The Poiseuille flow is spatially homogeneous and preserves the translation ($SO(2)$) symmetry but destroys the reflection (Z^2) symmetry. Figure 6(b) summarizes the effect on the eigenvalues: the double eigenvalues split on account of the skew-symmetric component

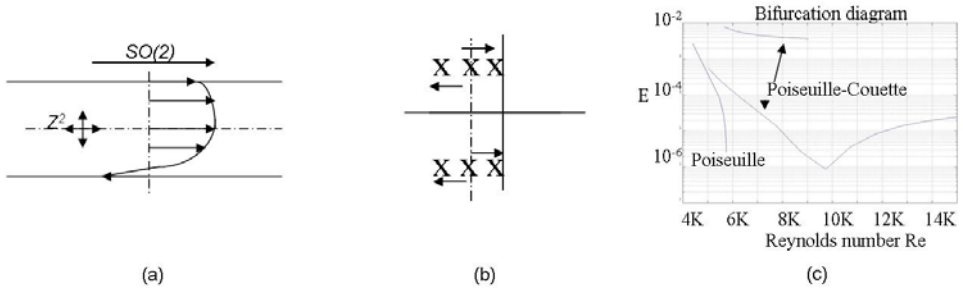


Figure 6. (a) *Schematic of the planar Poiseuille–Couette flow,* (b) *double eigenvalues of the planar Couette flow split as a result of the Poiseuille perturbation, and* (c) *two-dimensional traveling wave solution branches for the Poiseuille–Couette flow.*

of the feedback. The planar Couette flow is destabilized for a large enough perturbation (feedback) due to Poiseuille flow, which leads to traveling wave solutions for the nonlinear problem. Figure 6(c) depicts such a solution branch obtained using the continuation code AUTO. The kinetic energy $E \propto \int |u|^2$ is used as a solution measure (u denotes the two-dimensional velocity of the perturbation).

Planar Couette flow is thus another $O(2)$-equivariant example where introduction of a (linear) skew-symmetric perturbation leads to destabilization via splitting of the nominally double eigenvalues. In the presence of nonlinearities, here due to the convective term in the Navier–Stokes equation, such a destabilization leads to traveling waves, as shown in Figure 6. These solutions suggest nonlinear sensitivity of the planar Couette flow to (skew-symmetric) perturbations. However, the two-dimensional solutions of the perturbed problem are not believed to be relevant in explaining the experimentally observed three-dimensional vortex "streaks" seen in the Couette flow experiments; cf. [51]. Additional details on the discretization, nonlinear bifurcation analysis, computations, role of symmetry and perturbations to the planar Couette flow problem appear in the thesis [39] and the papers [40, 42].

4. Symmetry-breaking and suppression.

4.1. Wave-speed mistuning.

In this section, we study a *mistuned* version of the linearized wave equation with feedback:

$$p_{tt} - a^2(x)p_{xx} = \epsilon L\left[\frac{\partial}{\partial t}, \frac{\partial}{\partial x}\right](p). \tag{4.1}$$

$\epsilon L(p)$ denotes an arbitrary $SO(2)$-equivariant linear feedback as before, and

$$a^2(x) = a_0^2 + \epsilon\alpha(x), \tag{4.2}$$

where $\alpha(x)$ models the spatial inhomogeneity in wave speed. It is assumed to be some (periodic) function defined on a circle $[0, 2\pi]$. The objective is to study the effects of a "small" amount of mistuning on the eigenvalues of the linearized problem. In particular, it is of interest to use mistuning to ameliorate the detrimental effects of the destabilizing skew-symmetric feedback. This is the reason for scaling both the mistuning and feedback L by parameter ϵ.

We note that an inhomogeneity $\alpha(x)$ will in general break the $SO(2)$-equivariance of the wave equation, so the methods of the previous section need to be suitably modified. To compute the eigenvalues, we take the Laplace transform and consider a boundary value problem along the coordinate x:

$$a_0^2 p_{xx} - s^2 p + \epsilon\alpha(x) p_{xx} + \epsilon L\left[s, \frac{\partial}{\partial x}\right](p) = 0, \tag{4.3}$$

with periodic boundary conditions

$$\begin{aligned} p(0) &= p(2\pi), \\ p_x(0) &= p_x(2\pi). \end{aligned} \tag{4.4}$$

The term $\epsilon\alpha(x)p_{xx}$ models the effect of wave-speed parameter mistuning, and $\epsilon L(s,\cdot)p$ represents the feedback term after taking the Laplace transform. We are interested in the eigenvalues of (4.3)–(4.4), i.e., values of s for which a periodic solution $p(x)$ exists. To compute these, we use a perturbation method expressing the eigenfunction and eigenvalue in a series form:

$$p(x) = p_0(x) + \epsilon p_1(x) + o(\epsilon), \tag{4.5}$$

$$s = s_0 + \epsilon s_1 + o(\epsilon). \tag{4.6}$$

We note that ϵs_1 denotes the perturbation to the nominal eigenvalue s_0 as a result of (a) feedback ϵL (this case has already been studied in section 3), and (b) mistuning. Substituting (4.6) into (4.3),

$$O(1): \quad a_0^2 (p_0)_{xx} - s_0^2 p_0 = 0, \tag{4.7}$$

whose periodic eigensolution is given by

$$p_0 = c_k \cos(kx) + d_k \sin(kx), \tag{4.8}$$

$$s_0^2 = -a_0^2 k^2, \tag{4.9}$$

where k is any integer and $\{c_k, s_k\}$ are arbitrary real constants. Next,

$$\begin{aligned} O(\epsilon): \quad a_0^2 (p_1)_{xx} - s_0^2 p_1 &= 2 s_0 s_1 p_0 - \alpha(x)(p_0)_{xx} - L\left(s_0, \frac{\partial}{\partial x}\right)(p_0) \\ &\doteq R. \end{aligned} \tag{4.10}$$

Substituting $s_0^2 = -a_0^2 k^2$ on the left-hand side leads to a resonance condition for the right-hand-side term, denoted by R. In particular, for a solution p_1 to exist, R must lie in the range space of the linear operator

$$\left(a_0^2 \frac{\partial^2}{\partial x^2} + a_0^2 k^2\right). \tag{4.11}$$

For the self-adjoint operator, the range space is given by the complement of its null space $V_k = \text{span}\{\sin(kx), \cos(kx)\}$. This gives the resonance conditions

$$\begin{aligned} \frac{1}{\pi}\langle R, \sin(kx)\rangle &= 0, \\ \frac{1}{\pi}\langle R, \cos(kx)\rangle &= 0, \end{aligned} \tag{4.12}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product in $L^2(0, 2\pi)$. Explicitly, this leads to a matrix equation,

$$\begin{bmatrix} 2s_0s_1 - A_k(s_0) & B_k(s_0) \\ -B_k(s_0) & 2s_0s_1 - A_k(s_0) \end{bmatrix} \begin{bmatrix} s_k \\ c_k \end{bmatrix} - \frac{1}{\pi} \begin{bmatrix} \langle \alpha(x)(p_0)_{xx}, \sin(kx)\rangle \\ \langle \alpha(x)(p_0)_{xx}, \cos(kx)\rangle \end{bmatrix} = 0. \tag{4.13}$$

Without mistuning ($\alpha(x) \equiv 0$), this gives an approximation of the determinant equation (3.19),

$$\begin{vmatrix} 2s_0s_1 - A_k(s_0) & B_k(s_0) \\ -B_k(s_0) & 2s_0s_1 - A_k(s_0) \end{vmatrix} = 0, \tag{4.14}$$

where $s_0 = \pm i a_0 k$; note $2\epsilon s_0 s_1 = s^2 + a_0^2 k^2 + O(\epsilon^2)$. Its solution s_1 provides an $O(\epsilon)$ correction, on account of the feedback ϵL alone, to the eigenvalue $s_0 = \pm i a_0 k$. Let r_+ and r_- be the real parts corresponding to the two roots $s_1 = s_+$ and s_-, respectively. By taking trace and determinant, it follows that

$$\begin{aligned} r_+ + r_- &= \frac{\text{Imag}(A_k(ia_0k))}{a_0 k}, \\ r_+ - r_- &= \frac{\text{Real}(B_k(ia_0k))}{a_0 k}. \end{aligned} \tag{4.15}$$

These are in fact also (3.23)–(3.24). In the following, we evaluate the effect of parametric mistuning on these two quantities.

The resonance term on account of the wave-speed mistuning is given by

$$\begin{aligned} -\frac{1}{\pi}\langle \alpha(x)(p_0)_{xx}, \sin(kx)\rangle &= \frac{k^2}{\pi}\left(I_{cs} \cdot c_k + I_{ss} \cdot s_k\right), \\ -\frac{1}{\pi}\langle \alpha(x)(p_0)_{xx}, \cos(kx)\rangle &= \frac{k^2}{\pi}\left(I_{cc} \cdot c_k + I_{cs} \cdot s_k\right), \end{aligned} \tag{4.16}$$

where

$$\begin{aligned} I_{cs} &= \int_0^{2\pi} \alpha(x)\cos(kx)\sin(kx)dx, \\ I_{ss} &= \int_0^{2\pi} \alpha(x)\sin(kx)\sin(kx)dx, \\ I_{cc} &= \int_0^{2\pi} \alpha(x)\cos(kx)\cos(kx)dx. \end{aligned} \tag{4.17}$$

Evaluating these integrals gives

$$\frac{1}{\pi}\begin{bmatrix} \langle \alpha(x)(p_0)_{xx}, \sin(kx)\rangle \\ \langle \alpha(x)(p_0)_{xx}, \cos(kx)\rangle \end{bmatrix} = \begin{bmatrix} \alpha_0 - \alpha_c & \alpha_s \\ \alpha_s & \alpha_0 + \alpha_c \end{bmatrix}\begin{bmatrix} s_k \\ c_k \end{bmatrix}, \tag{4.18}$$

where

$$\begin{aligned} \alpha_0 &= \frac{k^2}{2\pi}\int_0^{2\pi} \alpha(x)dx, \\ \alpha_c &= \frac{k^2}{2\pi}\int_0^{2\pi} \alpha(x)\cos(2kx)dx, \\ \alpha_s &= \frac{k^2}{2\pi}\int_0^{2\pi} \alpha(x)\sin(2kx)dx. \end{aligned} \tag{4.19}$$

We denote by

$$\alpha_2^2 \doteq \alpha_c^2 + \alpha_s^2 \tag{4.20}$$

the "amplitude" of the perturbation corresponding to a $2k$ spatial harmonic. The effect of the resonant term due to the wave-speed mistuning (in (4.18)) on the eigenvalues is summarized in the following theorem. The simplest case, where $B_k(\pm i a_0 k)$ is real-valued, is treated first. We note that this represents the most important case because the instability results from a split in the real part of the eigenvalues and the difference $r_+ - r_-$ depends *only* on the real part of $B_k(i a_0 k)$, as shown in (4.15).

Theorem 4.1. *Consider the parametric wave-speed mistuning with nonzero $\alpha(x)$ in the linearized pde (4.1). Additionally, assume that $B_k(i a_0 k)$ is real-valued. In the limit as $\epsilon \to 0$,*

1. *the sum of the real part of the kth eigenvalues is unchanged by mistuning, i.e.,*

$$r_+ + r_- = \frac{\mathrm{Imag}(A_k(i a_0 k))}{a_0 k}, \tag{4.21}$$

2. *the difference of the real part of the kth eigenvalues*

$$r_+ - r_- = \begin{cases} \dfrac{1}{a_0 k}\,\mathrm{Imag}\left(\alpha_2^2 - B_k(i a_0 k)^2\right)^{\frac{1}{2}} & \text{for } \alpha_2 < |B_k(i a_0 k)|, \\ 0 & \text{otherwise.} \end{cases} \tag{4.22}$$

Proof. With the term in (4.1) due to the wave-speed mistuning, the resonance condition is modified to

$$\begin{bmatrix} 2s_0 s_1 - A_k(s_0) - \alpha_0 - \alpha_c & B_k(s_0) + \alpha_s \\ -B_k(s_0) + \alpha_s & 2s_0 s_1 - A_k(s_0) - \alpha_0 + \alpha_c \end{bmatrix}\begin{bmatrix} s_k \\ c_k \end{bmatrix} = 0. \tag{4.23}$$

The roots are determined by the determinant

$$\begin{vmatrix} 2s_0 s_1 - A_k(s_0) - \alpha_0 - \alpha_c & B_k(s_0) + \alpha_s \\ -B_k(s_0) + \alpha_s & 2s_0 s_1 - A_k(s_0) - \alpha_0 + \alpha_c \end{vmatrix} = 0, \tag{4.24}$$

where $s_0 = \pm i a_0 k$. Let r_+ and r_- be the real parts corresponding to the two roots $s_1 = s_+$ and s_-, respectively. By taking the trace with $s_0 = i a_0 k$, it follows that

$$r_+ + r_- = \frac{\text{Imag}(A_k(i a_0 k))}{a_0 k}; \tag{4.25}$$

i.e., the net amount of damping in the system is not modified by introducing mistuning. With $s_0 = -i a_0 k$, one obtains the same conclusion, and the eigenvalues $s_0 + \epsilon s_1$ for the choice of $s_0 = \pm i a_0 k$ arise as complex conjugates. This also shows (1) in the proof. To show (2), we evaluate the determinant which leads to

$$2 s_0 s_1 = A_k(s_0) + \alpha_0 \pm \left(\alpha_2^2 - B_k(s_0)^2\right)^{\frac{1}{2}}. \tag{4.26}$$

This gives

$$r_+ - r_- = \frac{1}{a_0 k} \text{Imag}\left(\alpha_2^2 - B_k(s_0)^2\right)^{\frac{1}{2}}. \tag{4.27}$$

Since $B_k(i a_0 k)$ is assumed to be real, the right-hand side is zero for values of $\alpha_2 > |B_k(i a_0 k)|$. Thus, mistuning serves to reverse the effect of skew-symmetry by bringing the eigenvalues closer. A critical amount of mistuning $\alpha_2 = |B_k(i a_0 k)|$ completely cancels the effect of skew-symmetry. Additional mistuning does not add additional damping. ■

Although the expression in (4.21)–(4.22) relating mistuning to symmetric and skew-symmetric feedback is particularly simple for real-valued $B_k(i a_0 k)$, a similar result holds for the general case of complex $B_k(i a_0 k)$ too. In particular, (4.21) remains unaffected, and the net amount of system damping is a function of symmetric feedback A_k alone. For a complex-valued B_k, we still have (4.27) and

$$r_+ - r_- = \frac{1}{a_0 k} \text{Imag}\left(\alpha_2^2 - B_k(s_0)^2\right)^{\frac{1}{2}}. \tag{4.28}$$

If $B_k(s_0)$ is purely imaginary, then (3.24) implies that $r_+ - r_- = 0$, and there is no instability on account of skew-symmetry. It is also clear from (4.28) that the mistuning is not needed for such a case. For a complex-valued $B_k(i a_0 k)$, the mistuning causes $|r_+ - r_-|$ to decrease and is thus *always* beneficial, but the relative benefit becomes smaller for larger mistuning levels. A good rule of thumb is to define optimal mistuning as

$$\alpha_2^o = |\text{Real}(B_k(i a_0 k))|, \tag{4.29}$$

the amount needed to completely cancel the effect of $\text{Real}(B_k(i a_0 k))$.

Remark 4.2. Mistuning does not affect the net amount of damping in the system, which is a function of the symmetric feedback $\text{Imag}(A_k(i a_0 k))$ as in (4.21). The mistuning ameliorates the detrimental effect of skew-symmetric feedback, and both the destabilization because of skew-symmetry and stabilization due to mistuning are model-independent.

We provide two interpretations for stabilization due to mistuning. One is in terms of 1:2 spatial resonance of the kth instability mode with $2k$th spatial harmonic of the mistuning. Such a "parametric resonance" effect is well studied in nonlinear oscillator models with

parametric excitation, the Mathieu's equation being one well-known example [56]. There is one important difference here because the circular geometry fixes the spatial frequencies to a discrete set. Thus, the 1:2 spatial resonance is exact here—the kth instability mode and the $2k$th spatial harmonic of the mistuning. In models of distributed systems with extended domains, such as planar Couette flow, thicker "Arnold tongues" with beneficial regions of stabilization due to spatial resonance are expected. Such is the case in the mistuning literature on structures where these regions arise as the so-called stop-bands [10].

The other interpretation is in terms of the dynamic model of interconnection in the subspace V_k (see Figure 5). It has already been pointed out that the presence of skew-symmetric feedback $B_k(s)$ serves to detrimentally couple the two dynamical systems corresponding to the two Fourier modes $\{\cos(kx), \sin(kx)\}$ in the subspace V_k. Although the strength of the coupling depends upon $B_k(s)$, it *always* destabilizes one of the two Fourier modes at the expense of the other. Similar, but now beneficial, coupling occurs in the presence of mistuning: the two kth Fourier modes are coupled due to the $2k$th spatial mistuning term. Mistuning, as the name suggests, causes the frequencies (imaginary part) of any $O(2)$-symmetric double eigenvalues to split. Such a split of the frequencies effectively desensitizes the feedback loop to a certain maximal amount of skew-symmetric feedback.

4.2. Application. We consider a model of thermoacoustic oscillations on a circular domain similar to the one described in [32, 5]. It is used to model counter-rotating transverse acoustic modes coupled via simple models of damping and an unsteady heat-release process. The system is of the form of (2.3), (2.4), with wave-speed mistuning $\alpha(x)$:

$$\begin{aligned} \frac{\partial^2 p}{\partial t^2} - a_0^2 \frac{\partial^2 p}{\partial x^2} = -b\left(1 - \frac{\partial^2}{\partial x^2}\right)\frac{\partial p}{\partial t} + \quad & m\frac{\partial p}{\partial x} \quad + \quad \alpha(x)\frac{\partial^2 p}{\partial x^2} \\ \text{symmetric} + \text{skew-symmetric} \quad & + \quad \text{mistuning}. \end{aligned} \tag{4.30}$$

We illustrate the stability properties of the system with respect to variations in the feedback components b and m and wave-speed mistuning $\alpha(x)$. The homogeneous case with $\alpha(x) = 0$ is considered first to compare the effects of the symmetric and skew-symmetric feedback components. Figure 7 shows the maximum real part of the closed-loop eigenvalues as a function of b and m. Note that the stability of the least stable eigenvalue depends directly on the sign of b but is independent of the sign of m. This is in agreement with the schematic shown in Figure 4 because the acoustic damping term $-b(1-\frac{\partial^2}{\partial x^2})\frac{\partial p}{\partial t}$ is Z^2-equivariant, and the skew-symmetric feedback term $m\frac{\partial p}{\partial x}$ is always destabilizing. This implies that the phase-response of the symmetric feedback directly affects the stability of the system. If the phase-response of b is 180° out of phase with the acoustic pressure, the effect will actually destabilize the system, while if the phase-response of b is in phase with the acoustic pressure, the effect will stabilize the system. In the case of a symmetric heat-release response, this is equivalent to the familiar Rayleigh criterion (see, e.g., [20]). Furthermore, this effect must also be considered for the application of acoustic damping using tuned resonators, as in [31]. On the other hand, the skew-symmetric coupling through the transverse acoustic velocity always has a destabilizing effect on the system, regardless of its phase-response.

We next consider the stabilizing effect of wave-speed mistuning for the system with skew-symmetric feedback. Figure 8 shows the contours of $p(x,t)$ resulting from a simulation of

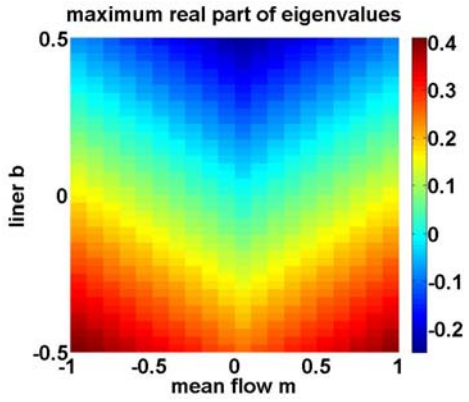


Figure 7. *Contours of the least stable eigenvalue for varying skew-symmetric (mean-flow) and symmetric (liner) feedback components. The eigenvalues become less stable for increasing values of $|m|$ and more stable for increasing values of b. The stability properties depend on the sign of b but are independent of the sign of m.*

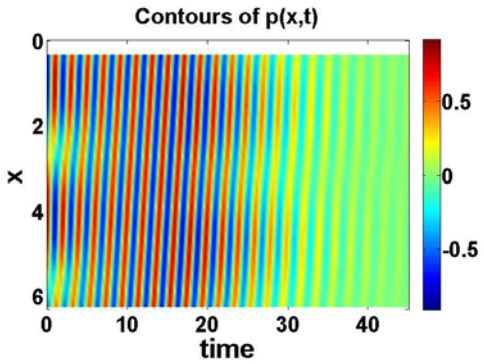


Figure 8. *Contours of $p(x,t)$ for the wave equation with symmetric and skew-symmetric feedback components. The wave-speed mistuning is initially set to zero, and the skew-symmetric feedback causes the instability of a rotating wave (indicated by the slanted contours between time $= 5$ and 15). At time $= 15$, the wave-speed mistuning is applied and the rotating wave is stabilized.*

the full wave equation with wave-speed mistuning initially set to zero. The skew-symmetric feedback causes the instability of a rotating wave (indicated by the slanted contours between time $= 5$ and 15. At time $= 15$, the wave-speed mistuning is applied and the rotating wave is stabilized.

In thermoacoustic applications, the precise value and sign of the parameter m describing the unsteady coupling between the transverse velocity and pressure is difficult to predict, as it may generally depend on a number of physical parameters such as the combustion chamber geometry, the inlet flow conditions, and the unsteady combustion process. If the sign of m changes, the result is that the instability wave rotates in the opposite direction. Therefore it is of interest to suppress acoustic waves rotating in either direction. Figure 9 shows the contours of $p(x,t)$ resulting from a simulation of the full wave equation with the sign of m opposite of what is shown in Figure 8. Notice that the wave rotates in the opposite direction (the contours are slanted in the opposite direction between time $= 5$ and time $= 15$). At

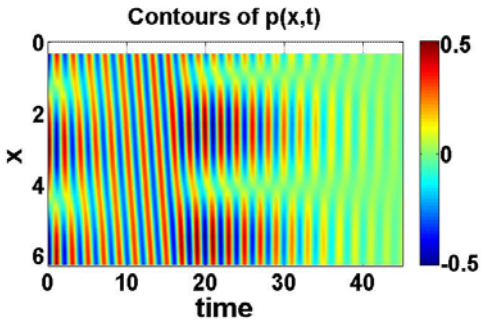


Figure 9. *Contours of $p(x,t)$ for the wave equation with symmetric and skew-symmetric feedback components. The skew-symmetric feedback is of opposite sign to that shown in Figure 8, and the wave rotates in the opposite direction. The wave-speed mistuning is initially set to zero, and the skew-symmetric feedback causes the instability of a rotating wave (indicated by the slanted contours between time $= 5$ and 15). At time $= 15$, the wave-speed mistuning is applied and the rotating wave is stabilized.*

time $= 15$ the same wave-speed mistuning is applied, and the rotating wave decays over time. This shows that the wave-speed mistuning stabilizes waves with either direction of rotation.

4.3. Fundamental limitation. The total amount of damping in the system is set by the symmetric feedback $A_k(s)$. Both the skew-symmetry and mistuning merely serve to exchange damping between the modes. Skew-symmetric feedback destabilizes one mode at the expense of the other, while mistuning reverses this. For a given skew-symmetric feedback (split of eigenvalues), there is an optimal amount of mistuning that reverses the destabilizing effect of skew-symmetric feedback. The optimal mistuning corresponds to the $2k$th spatial harmonic of amplitude

$$\alpha_2 = |\text{Real}(B_k(ia_0k))|. \tag{4.31}$$

Let Δr denote the split in the real eigenvalues $|r_+ - r_-|$ without any mistuning ($\alpha(x) \equiv 0$). Using (3.24), a useful formula for the optimal level of mistuning expressed only in terms of the eigenvalue split is given by

$$\alpha_2^0 = \Delta r \cdot a_0 k. \tag{4.32}$$

This optimal amount corresponds to the eigenvalue diagram where the kth eigenvalue pair is the closest. Decreasing the amount of mistuning from the optimal amount causes one of the modes to become more damped at the expense of the other mode, which becomes less damped. On the other hand, increasing the mistuning beyond the optimal amount causes the frequencies of the two modes to shift without any additional damping augmentation. This can be seen in Figure 10, which shows the first eigenvalues of the system (4.30) with nonzero b and m. The nominal eigenvalues of the system with symmetric damping and skew-symmetric feedback components are shown as red diamonds. Notice that one eigenvalue is stable, while the other eigenvalue is unstable. The wave-speed mistuning causes the eigenvalues to exchange stability, effectively stabilizing the unstable eigenvalue while destabilizing the stable eigenvalue. At the optimal mistuning amount, the eigenvalues meet at the black triangle. Further addition of wave-speed mistuning causes the eigenvalues to split in frequency to the blue squares.

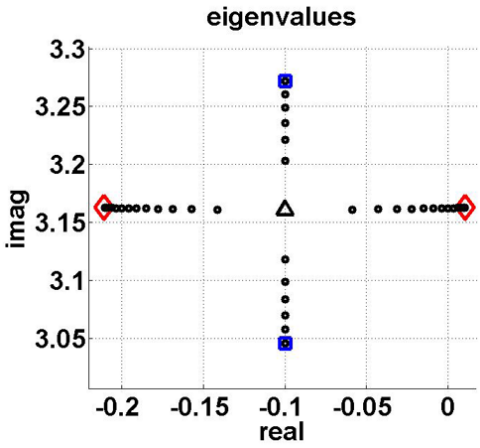


Figure 10. *The first eigenvalues of the system (4.30) with nonzero b and m and increasing amounts of wave-speed mistuning α_2. The nominal eigenvalues are shown as red diamonds. The wave-speed mistuning exchanges damping between the two modes until the optimal mistuning is reached, at which the eigenvalues meet at the black triangle. Further mistuning causes the eigenvalues to split in frequency to the blue squares.*

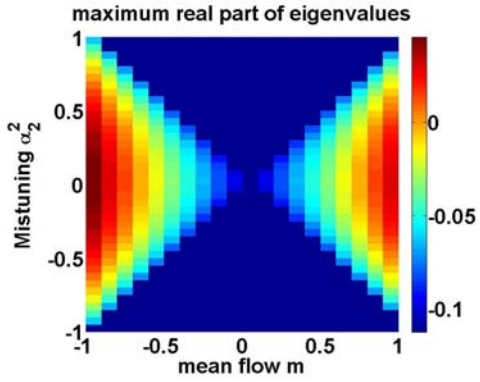


Figure 11. *Contours of the least stable eigenvalue for varying skew-symmetric (mean-flow) feedback and wave-speed mistuning. The eigenvalues become less stable for increasing values of $|m|$ and more stable for increasing values of $|\alpha_2^2|$. The stability properties are independent of the signs of these parameters.*

4.4. Robustness. Skew-symmetry gives a feedback model-independent conclusion on the inception of instability, while mistuning of the mean properties yields model-independent suppression of the instability. Figure 11 shows a model-independent stability boundary as a function of skew-symmetry and wave-speed mistuning. It is applicable to quite general dynamics or even physics of the feedback model. Moreover, the unmodeled dynamics can be lumped as part of the dynamic feedback. This makes the approach suitable to practical applications, where it has been used for the following:

1. developing a set of computational tools to identify and analyze time-series instability data from engine experiments. In particular, we formulate the feedback model in terms of its symmetric and skew-symmetric components and seek to identify these from the data.

2. designing changes on one of the engines. The optimal formula in (4.32) is particularly useful because the optimal value can be readily identified from experimental data. Furthermore, even smaller than optimal values of mistuning help the system stability. Hence, mistuning provides engineers with a robust and flexible design option.

A distinct advantage of using mistuning is that it constitutes a passive design change that is generally preferred in aerospace applications. A liner is a passive solution to augment the net damping—it works by stabilizing *both* the modes. In contrast, mistuning is a passive solution that uses the more damped system mode to stabilize the less damped one.

One of the disadvantages of using passive liners such as narrow-band tuned resonators is that if the frequency of oscillations changes, the performance deteriorates sharply [6]. In contrast, this approach is robust to many unmodeled physical effects, such as changes in frequency, as long as the modal structure of the problem is approximately preserved. Quantitative amounts of mistuning necessary for stabilization depend only upon the mean-flow effects. Such models are typically more reliable than their dynamic counterparts. Finally, the precise amount of mistuning can be "switched on" via gain scheduling or adaptation perhaps, only over a small portion of engine operating conditions where the instability is present.

5. Conclusion. In this paper, we outlined an approach for robust analysis and suppression of instability that utilizes the structure in the problem, namely its equivariance or symmetry properties. We obtained the results for a wave equation on a circle in feedback with a dynamic model, whose structure is assumed to be known. The two main ideas presented in this paper are that (a) the structure of the feedback can be used to explain the instability, and (b) manipulation of the structure can be used to control the instability.

The first idea, based on the methods of equivariant bifurcation theory, is to use only the symmetry properties of the feedback model in Figure 1 to explain the instability. The homogeneous wave equation has the so-called spatial symmetry group $O(2)$, whereby the pde is equivariant with respect to rotations and reflections. As a result of this, the individual eigenvalues are double. Moreover, because of physical considerations, these eigenvalues are lightly damped and close to the imaginary axis. The double eigenvalues correspond to the fact that a clockwise-rotating eigenmode is accompanied by its counter-clockwise-rotating symmetric counterpart. The instability phenomenon is related to the migration of one of these eigenvalues into the right half complex plane because of the dynamics of the feedback model.

To derive results on stability, it was shown that a feedback model can be decomposed as a sum of a symmetric and a skew-symmetric feedback. Conceptually, the symmetric feedback corresponds to dynamics that have reflection (about centerline) symmetry, while the skew-symmetry is a result of local asymmetry in feedback. Figure 4 shows the impact of symmetric and skew-symmetric heat release feedback on any double eigenvalues of the acoustics. The symmetric feedback causes the two eigenvalues to move as a pair in the same directions. It can either stabilize or destabilize, depending upon the feedback model. The skew-symmetric feedback, on the other hand, is always detrimental regardless of the feedback model. It splits the eigenvalues, causing one rotating mode to gain damping while causing the other rotating mode to lose the same amount of damping.

The second idea was to modify the structural aspects of the model in order to control the instability. This was accomplished by introducing precise spatial variations (mistuning)

in the mean properties such as wave-speed of the wave equation. While the skew-symmetric feedback causes the two eigenvalues to move apart, mistuning causes the eigenvalues to move closer, as shown in Figure 10. In either case, the net amount of damping in the system remains the same. This net damping depends upon the net symmetric feedback, due to the presence of liner perhaps, and is not affected by spatial variation in the mean. In effect, the mistuning utilizes the more heavily damped system modes to augment the damping of the lightly damped modes.

For a given skew-symmetric feedback (split of eigenvalues), there is an optimal amount of mean variation that reverses the detrimental effect of skew-symmetric feedback. This optimal amount corresponds to the eigenvalue diagram where the nominally double eigenvalues are the closest. Decreasing the amount of mistuning from the optimal amount causes one of the modes to become more damped at the expense of the other mode, which becomes less damped. On the other hand, increasing the mistuning beyond the optimal amount causes the frequencies of the two counter-rotating modes to shift without any additional damping augmentation.

The innovation lies in using only the symmetry structure of the feedback model to carry out both the analysis of the instability and design for its suppression. Skew-symmetry in feedback gives a feedback model-independent conclusion on stability, while mistuning of the mean properties yields model-independent suppression of the instability. Figure 11 shows a model-independent stability boundary as a function of skew-symmetry and mistuning.

Acknowledgments. The symmetry breaking idea in jet engines was inspired by extensive mistuning literature in structures [10], by thesis work of Shapiro [53], and by Mezic's discovery of the role of junk DNA in breaking symmetry in DNA molecules [43]. The authors are grateful to Michael Dellnitz, John Guckenheimer, Tim Healey, Mirko Hessel, and Jerry Marsden for numerous substantive conversations on symmetry and dynamical systems. We also acknowledge help and support from the UTRC dynamics and combustion control team, with special thanks to Jeff Cohen, Clas Jacobson, William Proscia, and Marios Soteriou. We are also grateful to the anonymous reviewers, whose comments helped improve this paper.

REFERENCES

[1] J. C. Alexander and G. Auchmuty, *Global bifurcations of phase locked oscillators*, Arch. Rational Mech. Anal., 93 (1986), pp. 253–270.

[2] A. K. Bajaj and P. R. Sethna, *Bifurcations in three-dimensional motions of articulated tubes*, Trans. ASME Ser. E J. Appl. Mech., 49 (1982), pp. 606–618.

[3] B. Bamieh, F. Paganini, and M. A. Dahleh, *Distributed control of spatially-invariant systems*, IEEE Trans. Automat. Control, 47 (2002), pp. 1091–1107.

[4] A. Banaszuk, K. Ariyur, M. Krstic, and C. A. Jacobson, *An adaptive algorithm for control of combustion instability*, Automatica J. IFAC, 40 (2004), pp. 1965–1972.

[5] A. Banaszuk, G. Hagen, P. Mehta, and J. Oppelstrup, *A linear model for control of thermoacoustic instabilities on an annular domain*, in Proceedings of the IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2003, Vol. 3, pp. 2346–2351.

[6] A. Banaszuk, P. G. Mehta, and G. Hagen, *The role of control in design: From fixing problems to the design of dynamics*, in ADCHEM: International Symposium on Advanced Control of Chemical Processes, 2006, pp. 913–928.

[7] A. Banaszuk, P. G. Mehta, C. A. Jacobson, and A. I. Khibnik, *Limits of achievable performance of controlled combustion processes*, IEEE Trans. Control Systems Technol., 14 (2006), pp. 881–895.

[8] A. Banaszuk, G. Rey, and D. Gysling, *Active control of flutter in turbomachinery using off blade actuators sensors. Part* II: *Control algorithm*, in Proceedings of the 41st IEEE Conference on Decision and Control, Vol. 4, IEEE Press, Piscataway, NJ, 2002, pp. 3704–3709.

[9] V. Bellucci, P. Rohr, C. O. Paschereit, and F. Magni, *On the use of Helmholtz resonators for damping acoustic pulsations in industrial gas turbines*, ASME J. Eng. Gas Turbines Power, 2 (2004), pp. 271–275.

[10] O. O. Bendiksen, *Localization phenomena in structural dynamics*, Chaos Solitons Fractals, 11 (2000), pp. 1621–1660.

[11] J. M. Bonell, R. L. Marshall, and G. T. Riecke, *Combustion Instability in Turbojet and Turbofan Augmentors*, AIAA-1971-698, 1971.

[12] S. Candel, *Combustion dynamics and control: Progress and challenges*, Proc. Combustion Inst., 29 (2002), pp. 1–28.

[13] T. S. Chen and D. D. Joseph, *Subcritical bifurcation in plane Poiseuille flow*, J. Fluid Mech., 58 (1973), pp. 337–351.

[14] P. Chossat and G. Iooss, *The Couette-Taylor Problem*, Appl. Math. Sci., Springer-Verlag, New York, 1994.

[15] Y. C. Chu, K. Glover, and A. P. Dowling, *Control of combustion oscillations via H-infinity loop-shaping, mu-analysis and integral quadratic constraints*, Automatica J. IFAC, 39 (2003), pp. 219–231.

[16] J. Cohen and A. Banaszuk, *Factors affecting the control of unstable combustors*, J. Propulsion Power, 19 (2003), pp. 811–821.

[17] J. M. Cohen, J. H. Stufflebeam, and W. Proscia, *The effect of fuel/air mixing on actuation authority in an active combustion instability control system*, ASME J. Eng. Gas Turbines Power, 23 (2001), pp. 537–542.

[18] J. D. Crawford and E. Knobloch, *On degenerate Hopf bifurcation with broken $o(2)$ symmetry*, Nonlinearity, 1 (1988), pp. 617–652.

[19] M. Cross and P. Hohenberg, *Pattern formation out of equilibrium*, Rev. Modern Phys., 65 (1998), pp. 851–1112.

[20] F. E. Culick, *A note on Rayleigh's criterion*, Combustion Sci. Technol., 56 (1987), pp. 159–166.

[21] G. Dangelmayr and E. Knobloch, *Hopf bifurcation with broken circular symmetry*, Nonlinearity, 4 (1991), pp. 399–427.

[22] M. Dellnitz, M. Golubitsky, A. Hohmann, and I. Stewart, *Spirals in scalar reaction-diffusion equations*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 5 (1995), pp. 1487–1501.

[23] A. P. Dowling and A. S. Morgans, *Feedback control of combustion oscillations*, in Annual Review of Fluid Mechanics, Annu. Rev. Fluid Mech. 37, Annual Reviews, Palo Alto, CA, 2005, pp. 151–182.

[24] M. Drumm, *Early warnings—Diagnostics and prognostics in aero engines*, Mechanical Engineering Magazine, January, 2005; available online at http://www.memagazine.org/contents/current/webonly/wex12405.html.

[25] B. Fiedler, *Global Bifurcation of Periodic Solutions with Symmetry*, Lecture Notes in Math., Springer-Verlag, Berlin, 1988.

[26] J. D. Goddard, *Material instabilities in complex fluids*, in Annual Review of Fluid Mechanics, Annu. Rev. Fluid Mech. 35, Annual Reviews, Palo Alto, CA, 2003, pp. 113–133.

[27] M. Golubitsky, *private communication*, 2005.

[28] M. Golubitsky and M. Roberts, *A classification of degenerate Hopf bifurcations with $o(2)$ symmetry*, J. Differential Equations, 69 (1987), pp. 216–264.

[29] M. Golubitsky and I. Stewart, *The Symmetry Perspective. From Equilibrium to Chaos in Phase Space and Physical Space*, Birkhäuser, Basel, 2002.

[30] M. Golubitsky, I. Stewart, and D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory*, Vol. II, Springer-Verlag, New York, 1988.

[31] D. L. Gysling, G. S. Copeland, D. C. McCormick, and W. M. Proscia, *Combustion system damping augmentation with Helmholtz resonators*, ASME J. Eng. Gas Turbines Power, 122 (2000), pp. 269–274.

[32] G. Hagen and A. Banaszuk, *Uncertainty propagation in a reduced order thermo-acoustic model*, in Proceedings of the 43rd IEEE Conference on Decision and Control, IEEE Press, Piscataway, NJ, 2004, Vol. 2, pp. 2250–2255.

[33] G. Hagen and P. G. Mehta, *On steady solutions of a pde model of compressor stall*, in Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, HI, Vol. 2, IEEE Press, Piscataway, NJ, 2003, pp. 1848–1853.
[34] T. J. Healey, *Large rotatory oscillations of transversely isotropic rods: Spatio-temporal symmetry-breaking bifurcation*, SIAM J. Appl. Math., 52 (1992), pp. 1120–1135.
[35] G. Iooss and M. Adelmeyer, *Topics in Bifurcation Theory and Applications*, Adv. Ser. Nonlinear Dynam., World Scientific, Singapore, 1992.
[36] H. Kielhofer, *Bifurcation Theory: An Introduction with Applications to Partial Differential Equations*, Appl. Math. Sci., Springer-Verlag, New York, 2004.
[37] W. Krebs, G. Walz, and S. Hoffmann, *Thermoacoustic Analysis of Annular Combustor*, AIAA Paper 99-1971, 1999.
[38] K. Kunze, C. Hirsch, and T. Sattelmayer, *Thermoacoustic modeling and control of multi burner combustion systems*, in Proceedings of the ASME Turbo Expo, ASME paper 2003-GT-38688, 2003 (CD-ROM).
[39] P. G. Mehta, *Nonbifurcating Solutions for Parallel Flows*, Ph.D. thesis, Department of Mathematics, Cornell University, Ithaca, NY, 2004.
[40] P. G. Mehta, *A unified well-posed computational approach to the 2d Orr-Sommerfeld problem*, J. Comput. Phys., 199 (2004), pp. 541–557.
[41] P. G. Mehta, G. Hagen, and A. Banaszuk, *Dynamical systems in jet engines: A symmetry based approach*, SIAM Dynam. Systems Magazine (DS Web), October 2005.
[42] P. G. Mehta and T. J. Healey, *On steady solutions of symmetry preserving perturbations of the 2d Couette flow problems*, Phys. Fluids, 17 (2005), paper 094108.
[43] I. Mezic, *Dynamics and control of large-scale molecular motion*, in Proceedings of the 2005 IFAC World Congress, Prague, 2005, Elsevier, New York, 2005.
[44] F. K. Moore and E. M. Greitzer, *A theory of post-stall transients in axial compression systems: Part 1—Development of equations*, ASME J. Eng. Gas Turbines Power, 108 (1986), pp. 68–76.
[45] E. P. Petrov, R. Vitali, and R. T. Haftka, *Optimization of Mistuned Bladed Discs Using Gradient-Based Response Surface Approximations*, AIAA-2000-1522, 2000.
[46] T. Poinsot and D. Veynante, *Theoretical and Numerical Combustion*, R. T. Edwards, Philadelphia, 2005.
[47] A. J. Rivas-Guerra and M. P. Mignolet, *Local/global effects of mistuning on the forced response of bladed disks*, ASME J. Eng. Gas Turbines Power, 125 (2003), pp. 1–11.
[48] V. A. Romanov, *Stability of plane-parallel Couette flow*, Funct. Anal. Appl., 7 (1973), pp. 137–146.
[49] D. H. Sattinger, *Bifurcation of periodic solutions of the Navier-Stokes equations*, Arch. Rational Mech. Anal., 41 (1971), pp. 66–80.
[50] S. S. Sattinger, Y. Neumeier, A. Nabi, B. T. Zinn, D. J. Amos, and D. D. Darling, *Sub-scale demonstration of the active feedback control of gas-turbine combustion instabilities*, ASME J. Eng. Gas Turbines Power, 122 (2000), pp. 262–268.
[51] P. J. Schmid and D. S. Henningson, *Stability and Transition in Shear Flows*, Springer, New York, 2001.
[52] J. P. Serre, *Linear Representations of Finite Groups*, Springer-Verlag, New York, 1977.
[53] B. Shapiro, *A symmetry approach to extension of flutter boundaries via mistuning*, J. Propulsion Power, 14 (1998), pp. 354–366.
[54] S. Stow and A. Dowling, *Modelling of circumferential modal coupling due to Helmoltz resonators*, in Proceedings of the ASME Turbo Expo, ASME paper 2003-GT-38168, 2003 (CD-ROM).
[55] S. Stow and A. Dowling, *Low-order modelling of thermoacoustic limit cycles*, in Proceedings of the ASME Turbo Expo, ASME paper GT2004-54245, 2004 (CD-ROM).
[56] A. Tondl, T. Ruijgrok, F. Verhulst, and R. Nabergoj, *Autoparametric Resonance in Mechanical Systems*, Cambridge University Press, Cambridge, UK, 2000.
[57] D. Veynante and T. Poinsot, *Large eddy simulation of combustion instabilities in turbulent premixed burners*, in Annual Research Briefs, Center for Turbulence Research, NASA Ames/Stanford University, Palo Alto, CA, 1997, pp. 253–274.
[58] M. Wang, J. B. Freund, and S. K. Lele, *Computational prediction of flow generated sound*, in Annual Review of Fluid Mechanics, Annu. Rev. Fluid Mech. 38, Annual Reviews, Palo Alto, CA, 2006, pp. 483–512.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 3, pp. 576–596

Multiple Gravity Assists, Capture, and Escape in the Restricted Three-Body Problem*

Shane D. Ross† and Daniel J. Scheeres‡

Dedicated to the memory of Kevin Granata and Liviu Librescu

Abstract. For low energy spacecraft trajectories such as multimoon orbiters for the Jupiter system, multiple gravity assists by moons could be used in conjunction with ballistic capture to drastically decrease fuel usage. In this paper, we investigate a special class of multiple gravity assists which can occur outside of the perturbing body's sphere of influence (the Hill sphere) and which is dynamically connected to orbits that get captured by the perturber and orbits which escape to infinity. We proceed by deriving a family of symplectic twist maps to approximate a particle's motion in the planar circular restricted three-body problem. The maps capture well the dynamics of the full equations of motion; the phase space contains a connected chaotic zone where intersections between unstable resonant orbit manifolds provide the template for lanes of fast migration between orbits of different semimajor axes. Within the chaotic zone, the concept of a set of reachable orbits is useful. This set can be considered bounded by, on one end, orbits leading to ballistic capture around the perturber, and on the other end, the orbits escaping to infinity or a bounding surface at finite distance.

Key words. three-body problem, celestial mechanics, symplectic maps, phase space transport

AMS subject classifications. 70F07, 70F15, 37J45, 70H33

DOI. 10.1137/060663374

1. Introduction.

In recent numerical constructions of multimoon orbiters for the Jupiter system, multiple gravity assists for a spacecraft in the exterior and interior Hill's regions were used to construct low energy transfers between moons [25, 35]. These gravity assists do not lend themselves to the patched-conic approach of astrodynamics as they occur even when the spacecraft remains outside of the perturbing moon's sphere of influence or Hill sphere.

In this paper, we investigate such gravity assists by the explicit construction of an energy kick function approximating the effect of the perturbing moon on a spacecraft's jovicentric orbit. We use Picard's method of successive approximations to generate a symplectic twist map for the planar restricted three-body problem which approximates a Poincaré map at the surface of section corresponding to the periapsis condition. Other authors [23, 6, 16, 36] have considered similar maps to study the long-time evolution of nearly parabolic comets.

The family of maps we develop are applicable to objects on near-Keplerian elliptical orbits of low, moderate, and high eccentricity. We are especially interested in the dynamics of objects

*Received by the editors October 20, 2006; accepted for publication (in revised form) by M. Dellnitz April 24, 2007; published electronically July 27, 2007.
http://www.siam.org/journals/siads/6-3/66337.html
†Engineering Science and Mechanics, Virginia Polytechnic Institute and State University, 224 Norris Hall, MC 0219, Blacksburg, VA 24061 (sdross@vt.edu). This author acknowledges the support of National Science Foundation grant DMS 0402842.
‡Aerospace Engineering, University of Michigan, 3048 FXB, Ann Arbor, MI 48109-2140 (scheeres@umich.edu).

whose periapse or apoapse (closest and furthest approach to the central body, respectively) grazes the Hill sphere of the perturbing body. The engineering application envisioned is the design of low energy trajectories [7, 2], specifically between moons in the satellite system of one of the giant planets [15, 12]. Multiple gravity assists in resonance with the perturber are a key physical mechanism which could be exploited in future scientific missions [25, 26]. For example, a trajectory sent from Earth to the Jovian system, just grazing the orbit of the outermost icy moon Callisto, can migrate using little or no fuel from orbits with large apoapses to smaller ones.

The advantage of considering an analytical two-dimensional map as opposed to full numerical integration of the restricted three-body equations of motion is that we can apply all the theoretical and computational machinery applicable to phase space transport in symplectic twist maps [24, 19, 9]. We also make connections with earlier work on the restricted three-body problem, particularly capture via tube dynamics [14]. The result is a fuller picture of the global dynamics in the restricted three-body problem.

The paper is organized as follows. In section 2, we write the Hamiltonian for the restricted three-body problem in a form appropriate for application of Picard's method and introduce the energy regime we are considering. In section 3, we apply Picard's method of successive approximations to determine the orbital changes over one orbit. In section 4, we develop a family of area-preserving twist map approximations to the Poincaré return map which take one periapsis passage to another for orbits exterior to the secondary mass. The dynamics of the maps are discussed in sections 5 and 6, in particular, the relationship of multiple gravity assist trajectories to capture around the secondary mass and escape to infinity. We discuss our results and indicate future directions in section 7.

2. The Hamiltonian. Consider the planar circular restricted three-body problem (PCR3BP) with a test particle P in the gravitational field of two primary masses, m_1 and m_2, which are on circular orbits about their common center of mass. For illustrative purposes, we take m_1 to be Jupiter, m_2 to be one of its moons, and the particle to be a natural object or spacecraft of insignificant mass.

We use the standard system of units; the m_1-m_2 distance is scaled to 1, as is their mean motion about the center of mass, and their mass ratio is $\mu = m_2/(m_1 + m_2) \ll 1$. The PCR3BP is a perturbation of the two-body Kepler problem, where the particle is assumed to be on a near-Keplerian orbit around the m_1-m_2 barycenter. We can write the Hamiltonian in a frame centered on the barycenter,

$$H_{\text{iner}} = \frac{1}{2}(p_x^2 + p_y^2) - \frac{1-\mu}{r_1} - \frac{\mu}{r_2}. \tag{2.1}$$

We can write r_1 and r_2 in terms of (r, θ), where $r \gg \mu$ is the distance between the particle and the barycenter, and θ, as shown in Figure 1, is the angle between the particle and m_2, measured from the barycenter. Using $r = \sqrt{r_1^2 - 2\mu r_1 \cos\theta + \mu^2} = r_1 - \mu\cos\theta + \mathcal{O}(\mu^2)$ we can write

$$\frac{1}{r_1} = \frac{1}{r} - \frac{\mu}{r^2}\cos\theta + \mathcal{O}(\mu^2). \tag{2.2}$$

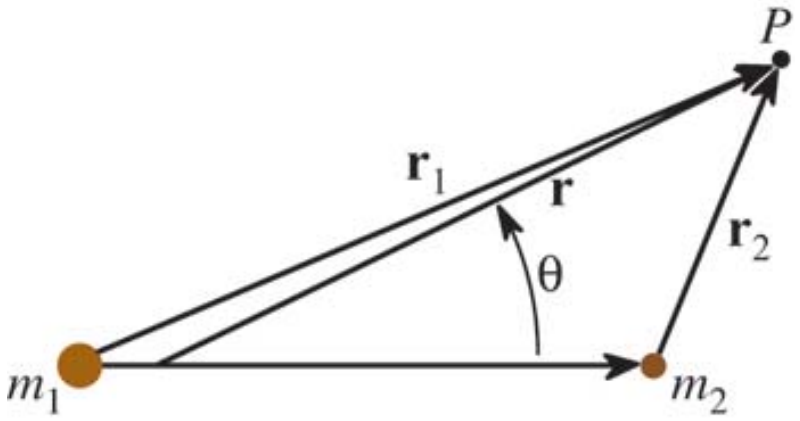


Figure 1. *The relationship between distances in the restricted three-body problem.*

Our Hamiltonian (2.1) can be rewritten as

$$\begin{aligned} H_{\text{iner}} &= \left(\frac{1}{2}(p_x^2+p_y^2)-\frac{1}{r}\right)+\mu\left(-\frac{1}{r_2}+\frac{\cos\theta}{r^2}+\frac{1}{r}\right)+\mathcal{O}(\mu^2) \\ &= K+\mu R+\mathcal{O}(\mu^2), \end{aligned} \tag{2.3}$$

where $K=-1/(2a)$ is the Keplerian part and μR the perturbing part to first order in μ. The particle's phase space position is given by its instantaneous semimajor axis a, eccentricity e, argument of periapsis ω, and true anomaly ν, which have geometric descriptions as given in Figure 2. These orbital elements are referred to as "osculating" or instantaneous orbital elements since they represent an instantaneous approximation of the motion of the object as a Keplerian orbit, which is the kind of orbit it would have if other perturbations were not present. In a celestial mechanics perturbation problem like the one we are considering, these elements are very useful.

For the perturbing function, we have

$$\begin{aligned} R &= -\frac{1}{r_2}+\frac{\cos\theta}{r^2}+\frac{1}{r} \\ &= -\frac{1}{\sqrt{1+r^2-2\,r\cos\theta}}+\frac{\cos\theta}{r^2}+\frac{1}{r}. \end{aligned} \tag{2.4}$$

The angle θ can be related to the traditional osculating elliptic elements of the particle as $\theta=\omega+\nu-t$, where $\nu=\nu(t)$ is the true anomaly of the particle and a function of time, ω is the particle's angle of periapsis, $r=p/(1+e\cos\nu)$, and $p=a(1-e^2)$.

In the frame corotating with m_2 and m_1 about their barycenter, the time-dependent Hamiltonian (2.3) can be rewritten in a time-independent form,

$$H_{\text{rot}}(a,e,\bar{\omega},\nu)=K(a)+\mu R(a,e,\bar{\omega},\nu)-G(a,e), \tag{2.5}$$

where $G=\sqrt{a(1-e^2)}=\sqrt{p}$ is the angular momentum of the particle's orbit and we drop the $\mathcal{O}(\mu^2)$ terms. In the rotating frame, the coordinate conjugate to G is $\bar{\omega}=\omega-t$, the angle

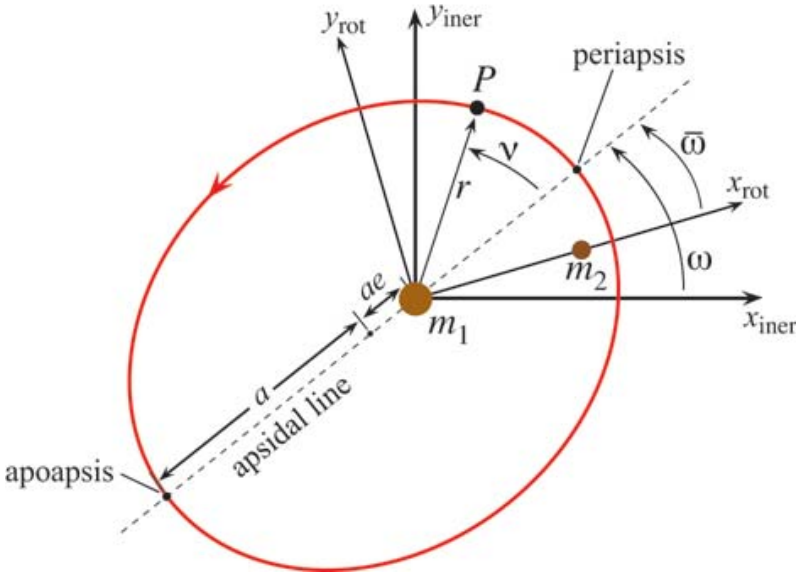


Figure 2. *The osculating or instantaneous orbital elements for a particle P in a near-Keplerian orbit about a massive central body m_1. The perturbing body, m_2, is in a circular orbit about m_1 of unit frequency, and the x axis of the rotating frame is defined as the line from m_1 to m_2, where the y axis of the rotating frame completes a right-handed coordinate system.*

of periapsis measured from the m_1-m_2 line. Hamiltonian (2.5) is time-independent [31] and therefore constant along particle trajectories. We refer to this constant as the Jacobi constant, $C_J = -2H_{\text{rot}}$. Although the invariance of the Hamiltonian in the rotating frame is useful, we continue to use ω (in the inertial frame) for the computations described in the next section.

The Jacobi constant allows us to obtain a coarse partition of accessible phase space for the particle. For some values of C_J, there are inaccessible regions in the rotating frame. The inaccessible regions whose boundaries are zero velocity curves divide the accessible regions, known historically as Hill's regions, naturally into three regions. For a given μ there are five basic cases of connectivity between the regions, corresponding to five intervals of C_J (see [14] for details). The cases are shown in Figure 3. The divisions between the cases are given by the Jacobi constant at the Lagrange points, i.e., $C_i = C_J(L_i)$.

We are focusing on particle motion which remains in the exterior region. According to the cases, this would mean $C_J > C_2$. In the Jupiter–Callisto system ($\mu = 5.667 \times 10^{-5}$), for example, we have $C_2 = 3.00618$. Even though motion from the exterior to the regions around m_2 and m_1 is possible for cases 3, 4, and 5 we find that for energies close to but below C_2, particle motion can remain in the exterior region for long times. Transit from the exterior region to the region around m_2 is possible for $C_J < C_2$, and the connection between multiple gravity assists and capture orbits will be discussed in section 6.

3. Changes in orbital elements over one orbit. To evaluate changes in the osculating orbital elements over one orbit, we use the first iteration of Picard's method of successive approximations, following [27] and [33].

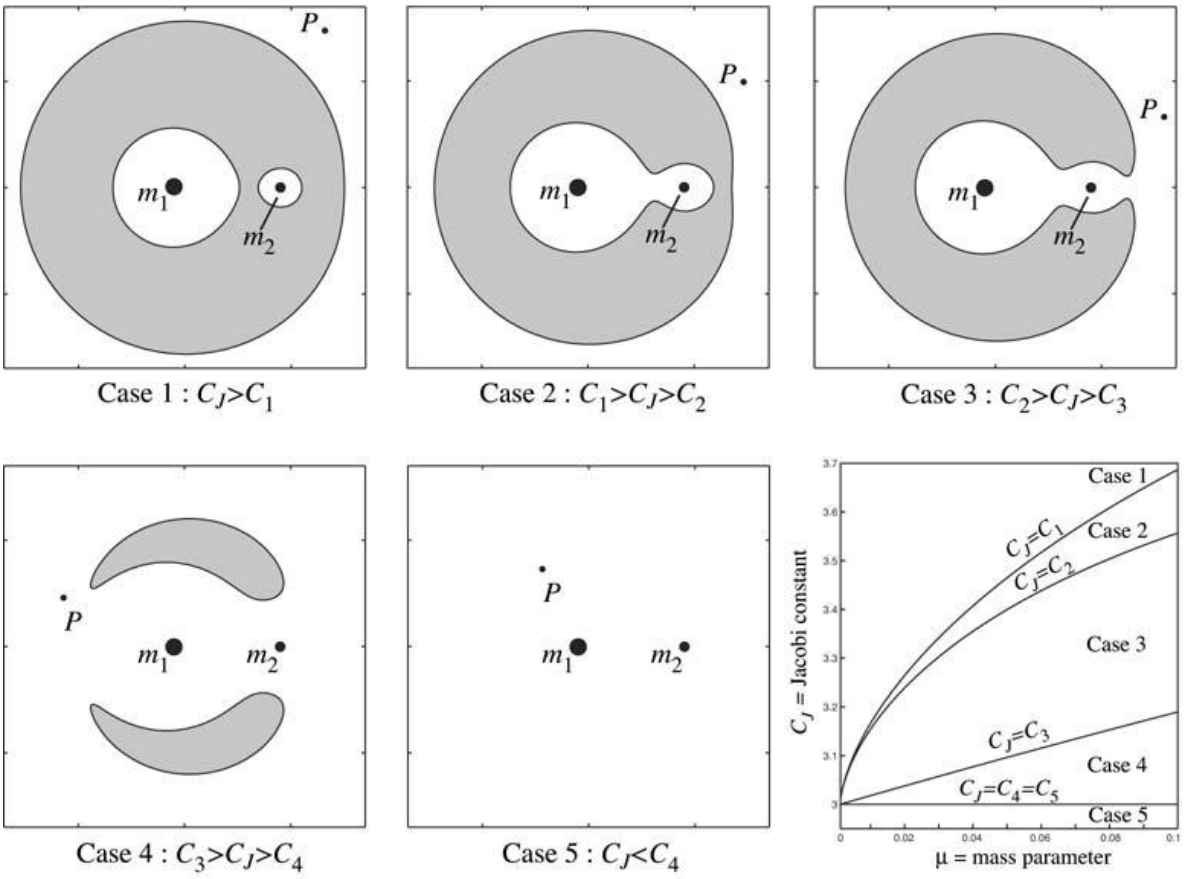


Figure 3. *Regions of possible motion. Zero velocity curves for five values of the Jacobi constant C_J, one in each of the cases, are shown on the x-y plane for $\mu = 0.3$. These curves bound the zone, in white, accessible by the particle P for a given C_J. The part of the x-y plane which is shaded is inaccessible for a given energy and is known as the forbidden region. The outermost accessible region, known as the exterior region, extends to infinity. In the fifth case, the forbidden region vanishes and motion over the entire x-y plane is possible. In the last panel, the (μ, C_J)-plane is partitioned into the five cases of Hill's regions.*

Picard iteration. First, we introduce Picard's method. Let $t \in \mathbf{R}$ be the time and $x(t)$ and $f(x,t)$ be functions with values in $\mathbf{R}^n$. Consider the problem of finding solutions for the dynamical system

$$\begin{aligned} \frac{dx}{dt} &= f(x,t), \\ x(t_0) &= x_0, \end{aligned} \tag{3.1}$$

where $t_0 \in \mathbf{R}$ and $x_0 \in \mathbf{R}^n$ are the initial time and configuration of the system.

The Picard iterate of a function $y(t)$ with initial condition x_0 is defined as

$$P_{x_0} y(t) = x_0 + \int_{t_0}^{t} f(y(\tau), \tau)\, d\tau. \tag{3.2}$$

Let $P_{x_0}^2 y(t) = P_{x_0}(P_{x_0} y(t))$, $P_{x_0}^3 y(t) = P_{x_0}(P_{x_0}^2 y(t))$, and so on. If f and $\frac{\partial f}{\partial t}$ are continuous, then it can be shown that given any continuous initial curve $y(t)$, its mth Picard iterate $P_{x_0}^m y(t)$ converges to the solution of (3.1) as $m \to \infty$ if t is in a suitable interval of values close to t_0 (see, for example, [4]).

Approximating changes in orbital elements. We use Picard's method as follows. Suppose $y(t)$ represents the osculating orbital elements of a particle in an orbit about the large primary body m_1. The function f includes the perturbation of the secondary body m_2 of mass μ. The unperturbed orbital elements are a constant function $y(t) = x_0$ over the time interval $[t_0, t_1]$, $t_1 > t_0$. The first iteration of Picard's method yields

$$P_{x_0} y(t) = x_0 + \int_{t_0}^{t} f(x_0, \tau)\, d\tau, \tag{3.3}$$

where the time variation in the integrand is due to the perturbation of the m_2 and the true anomaly ν of the particle orbit. Noting that $P_{x_0} y(t_0) = x_0$, and making the approximation $y(t_1) \approx P_{x_0} y(t_1)$, we derive the first order change in the orbital elements over one orbit as

$$\Delta y = \int_{t_0}^{t_1} f(x_0, \tau)\, d\tau, \tag{3.4}$$

where $T = t_1 - t_0$ is one period of the unperturbed particle orbit.

Perturbations to particles exterior to the orbit of the secondary mass. Consider a particle in the exterior realm, with Jacobi constant close to but above that of L_2; see case 2 in Figure 3. In this situation, the particle travels on a near-Keplerian orbit around the central mass m_1. The greatest perturbation occurs at periapsis, when the particle's orbit comes closest to the orbit of the perturbing mass. We therefore take the limits of integration to be from apoapsis to apoapsis, with $t_0 = t_* - T/2$, $t_1 = t_* + T/2$, where $T = 2\pi a^{3/2}$ is the unperturbed orbital period of the particle and t_* is the time of periapsis passage. We take $t_* = 0$ in general. Periapsis passage occurs at $\nu = t = 0$, and thus at the moment over which the perturbation is evaluated, the angles ω, $\bar{\omega}$, and θ coincide.

For our computations, it is useful to use the canonical form of the Lagrange planetary equations [5] which express the change in G as

$$\frac{dG}{dt} = -\mu \frac{\partial R}{\partial \omega}, \tag{3.5}$$

where, from (2.4), we calculate

$$\frac{\partial R}{\partial \omega} = \frac{r}{r_2^3} \sin(\omega + \nu - t) - \frac{1}{r^2} \sin(\omega + \nu - t). \tag{3.6}$$

The change in G over one orbit can be computed to first order in μ using (3.5) as the dynamical system for which we apply the approximation (3.4):

$$\begin{aligned}\Delta G &= -\mu \int_{-T/2}^{T/2} \frac{\partial R}{\partial \omega}\, dt \\ &= -\mu \int_{-T/2}^{T/2} \left[\frac{r}{r_2^3} \sin(\omega + \nu(t) - t) - \frac{1}{r^2} \sin(\omega + \nu(t) - t) \right] dt \\ &= -\frac{\mu}{\sqrt{p}} \int_{-\pi}^{\pi} \left[\left(\frac{r}{r_2} \right)^3 \sin(\omega + \nu - t(\nu)) - \sin(\omega + \nu - t(\nu)) \right] d\nu \\ &= -\frac{\mu}{\sqrt{p}} \left[\left(\int_{-\pi}^{\pi} \left(\frac{r}{r_2} \right)^3 \sin(\omega + \nu - t(\nu))\, d\nu \right) - \sin\omega \left(2 \int_0^{\pi} \cos(\nu - t(\nu))\, d\nu \right) \right]. \end{aligned} \tag{3.7}$$

If the first integral is expanded as a Fourier series in ω, the integrals can be expressed as functions of Hansen coefficients [28]. However, there is no significant advantage in this for the current application, and so the integrals are evaluated by quadratures in their current form.

Our goal is to compute ΔK, the change in Keplerian energy over one orbit. By the invariance of the Jacobi constant we have $\Delta H_{\text{rot}} = 0$ and, therefore, from (2.5),

$$\Delta K = \Delta G - \mu \Delta R, \tag{3.8}$$

where

$$\begin{aligned}\Delta R &= R(\nu = \pi) - R(\nu = -\pi), \\ &= \frac{1}{\sqrt{1 + Q^2 + 2\,Q\cos(\omega + \tau)}} - \frac{1}{\sqrt{1 + Q^2 + 2\,Q\cos(\omega - \tau)}} + \frac{2}{Q^2} \sin\omega \sin\tau, \end{aligned} \tag{3.9}$$

with $Q = a(1 + e)$ and $\tau = \pi a^{3/2}$ the apoapsis distance and half period of the unperturbed orbit, respectively.

ΔK is a function of ω, K, and e. The invariance of the Jacobi constant yields a relationship among these three variables, implying $\Delta K = \Delta K_{C_J}(\omega, K)$, where C_J is a parameter. The expression (3.8) can be written as $\Delta K_{C_J} = \mu f(\omega, K)$, where f is the *energy kick function* following the terminology of [16, 36, 21]. Physically, the energy kick ΔK_{C_J} between consecutive apoapsis passages can be approximated as a discrete event occurring at periapsis passage. Since $\omega = \bar{\omega}$ at the moment the kick takes place, we will drop the bar from now on but will consider ω to be the angle of periapsis as measured in the rotating frame, that is, with respect to the m_1-m_2 line; in other words, ω is the azimuthal separation of the particle and the perturbing body m_2 at the moment of periapsis passage and therefore at the moment of the kick.

For our application, $Q \gg q > 1$, and it is straightforward to show that ΔR is bounded:

$$|\Delta R| \leq \frac{4(Q^2 - \frac{1}{2})}{Q^2(Q^2 - 1)} \approx \frac{4}{Q^2}. \tag{3.10}$$

For values of K and C_J used in this study, the maximum contribution of $\mu\Delta R$ is much smaller than that of ΔG, so we ignore it for the remainder of the paper; i.e.,

$$f(\omega, K) = \Delta K_{C_J}(\omega, K)/\mu = \Delta G/\mu \tag{3.11}$$

is assumed.

In Figure 4(a), we plot f versus ω for $C_J = 3$ and an example value of semimajor axis a. Notice that the location and angular width of the greatest perturbations are anti-symmetric about the m_1-m_2 line ($\omega = 0$) and are independent of the size of the perturbation μ. As shown in Figure 4(c), particle orbits whose periapsis passages occur slightly ahead of m_2 in its orbit ($\omega > 0$) will decrease their energy, while those with periapsis passages just behind m_2 ($\omega < 0$) will increase their energy. The location and magnitude of the maximum kicks are plotted as a function of a in Figure 4(b).

We want to look at the cumulative effect of multiple passes near m_2. Such a trajectory has an invariant Jacobi constant, although its orbital elements a and e may change dramatically over time. In the next section, we consider ΔK_{C_J} where $C_J =$ constant. As the orbital energy K and thus the semimajor axis a changes according to ΔK_{C_J} for each kick, e changes to leave invariant (2.5), rewritten as

$$C_J = \frac{1}{a} + 2\sqrt{a(1-e^2)} + 2\mu\left(\frac{1}{r_2} - \frac{1}{r_1}\right). \tag{3.12}$$

For our application, the terms of $\mathcal{O}(\mu)$ are small and we are left with the Tisserand parameter from which we obtain the eccentricity.

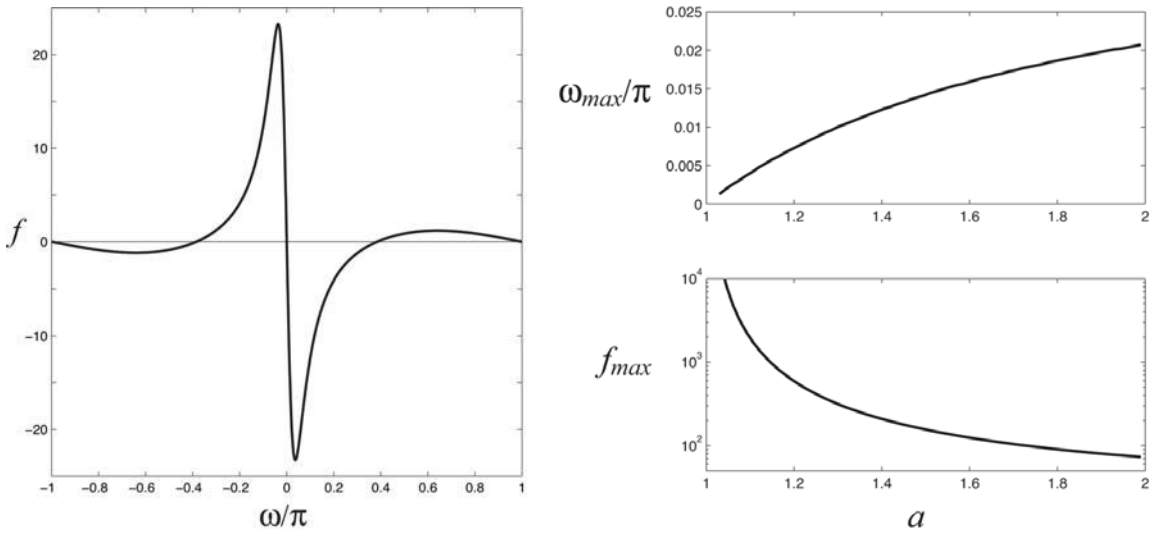
4. The Keplerian map derived. Consider the PCR3BP energy surface given by the pair (μ, C_J), where $\mu \ll 1$ and $C_J \approx 3$ is close to the Jacobi constant of L_2. We want to compute the sequence of pairs (ω_n, K_n), $n = 1, 2, 3, \ldots$, which result from an initial condition (ω_0, K_0). These pairs are the azimuthal separation of the particle and the perturbing mass and the particle's Keplerian energy at the nth periapsis passage.

In our approximation, (ω_n, K_n) represents the particle's orbit just before receiving an energy kick. Consider, for example, one of the trajectories in Figure 4(c) before the kick was received. Immediately following the kick, the orbit becomes (ω_{n+1}, K_{n+1}). The time until the next periapsis passage is now $\Delta t = 2\pi a_{n+1}^{3/2} = 2\pi(-2K_{n+1})^{-3/2}$. The change in the periapsis angle during this period is $\Delta\omega = -\Delta t$ modulo 2π. Note, we are neglecting the direct effect of the gravity interaction on the argument of periapsis, considering only its indirect effect through the change in Keplerian energy. We therefore obtain a two-dimensional update map $(\omega_{n+1}, K_{n+1}) = F(\omega_n, K_n)$ of the cylinder $\mathcal{A} = S^1 \times \mathbf{R}$ onto itself; i.e., $F : \mathcal{A} \to \mathcal{A}$, where

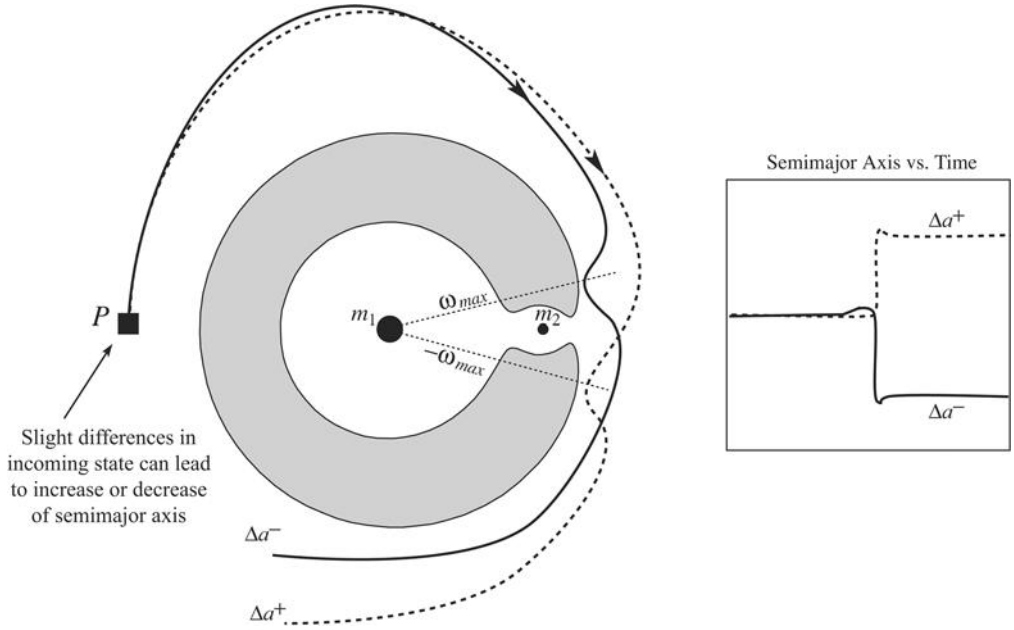
$$F\begin{pmatrix} \omega_n \\ K_n \end{pmatrix} = \begin{pmatrix} \omega_{n+1} \\ K_{n+1} \end{pmatrix} = \begin{pmatrix} \omega_n - 2\pi(-2K_{n+1})^{-3/2} \pmod{2\pi} \\ K_n + \mu f(\omega_n, K_n) \end{pmatrix}. \tag{4.1}$$

The Jacobian determinant of F is $1 + \mu\frac{\partial f}{\partial K}$. If we assume f is independent of K, then we have a symplectic (area-preserving) twist map, desirable for many known properties of such maps [19, 8]. For the remainder of this paper, we neglect the K dependence of f. Given a reference $\bar{K}$, we let $f(\omega) = f(\omega, \bar{K})$ and thereby make F area-preserving:

$$F\begin{pmatrix} \omega_n \\ K_n \end{pmatrix} = \begin{pmatrix} \omega_{n+1} \\ K_{n+1} \end{pmatrix} = \begin{pmatrix} \omega_n - 2\pi(-2K_{n+1})^{-3/2} \pmod{2\pi} \\ K_n + \mu f(\omega_n) \end{pmatrix}. \tag{4.2}$$



(a) (b)



(c)

Figure 4. (a) *The energy kick function f versus ω for $C_J = 3$, for $a = 15$. The plot is odd in ω, $f(-\omega) = -f(\omega)$.* (b) *The location and magnitude of the maximum kicks as a function of a. For f_{max}, the vertical axis is logarithmic. The largest positive kick is at $-\omega_{max}$, i.e., $f_{max} = f(-\omega_{max})$. The largest negative kick is at $-\omega_{max}$ with value $-f_{max}$. Notice that smaller a orbits yield larger maximum kicks.* (c) *Two trajectories with semimajor axis a^0 begin at the same position with slightly different velocities, shown here schematically in the rotating frame for energy case 3. The solid trajectory has its periapsis passage at ω_{max}, receives the largest negative energy kick, and drops in semimajor axis, shown in the time history on the right panel. The dashed trajectory has its periapsis passage at $-\omega_{max}$ and gets kicked to a larger semimajor axis. As shown schematically in the time history, the energy kick is nearly instantaneous.*

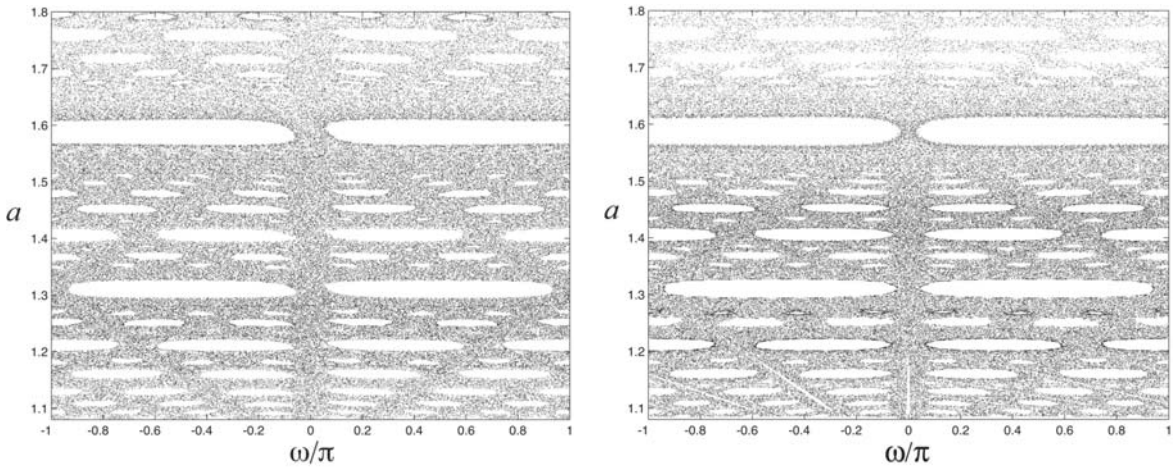


Figure 5. *Plot of a versus ω for $\mu = 5.667 \times 10^{-5}$, $C_J = 3$, $\bar{a} = -1/(2\bar{K}) = 1.35$. The left-hand plot shows trajectories computed using the Keplerian map. The right-hand plot shows trajectories computed via full integration of the circular restricted three-body problem, using a Poincaré surface of section at periapsis. The initial conditions for both were taken initially in the chaotic sea and followed for 10^4 iterates, thus producing the "Swiss cheese" appearance, where holes corresponding to stable resonant islands reside.*

This map has a twist to the right:

$$\left.\frac{\partial \omega_{n+1}}{\partial K_n}\right|_{\omega_n} = 3\pi \left(-2(K_n + \mu f(\omega_n))\right)^{-5/2} > 0 \tag{4.3}$$

for $K_n < -\mu f(\omega_n)$. From here on, we will understand F to be defined over the appropriate section of the cylinder $\mathcal{A}$ for which (4.3) holds.

The map (4.2) has been called the *Keplerian map* by earlier authors who derived it by other means for the case of near-parabolic orbits [23, 16]. Pan and Sari [21] considered the large a limit, referring to their map as the *eccentric mapping*. Our form is appropriate for elliptical orbits ($e < 1$), even those of low to moderate eccentricity. For a given Jacobi constant and reference energy $\bar{K}$ (where, say, $\bar{K} = K_0$), F is a mapping approximating the dynamics of the PCR3BP for orbits with nearby Keplerian energies, i.e., K_n close to $\bar{K}$. Specifically, F approximates the Poincaré return map of the fully integrated equations of motion where the surface of section is taken at periapsis, a map used recently in the study of the Hill problem [34, 22].

5. Dynamics of the Keplerian map. Other authors have considered similar maps to study the long-time evolution of nearly parabolic comets [6, 16, 36]. We apply our map to the identification of transfer trajectories applicable to spacecraft in a planet-moon system. For example, we can consider a spacecraft in the Jupiter–Callisto system ($\mu = 5.667 \times 10^{-5}$) with $C_J = 3$. Using semimajor axis as our vertical coordinate and applying the map (4.2) for several initial values in the (ω, a)-plane result in the left-hand plot of Figure 5. Throughout the paper we will reference a and K interchangeably; the context should make it clear which coordinate we are using.

For our map computations, we need only calculate $f(\omega)$ once from (3.11) for a grid of points $\omega \in [-\pi, \pi]$. Saving the results in a look-up table, we use interpolation to obtain f for arbitrary ω. The initial values for the left-hand side of Figure 5 were chosen in a connected chaotic sea (an irregular component in the sense of Birkhoff), avoiding the stable islands corresponding to stable mean motion resonances of the particle's orbit with Callisto's. As our phase space is the cylinder, the left and right sides of the plot ($\omega = \pm\pi$) are to be identified. The right-hand plot shows trajectories computed via full integration of the PCR3BP, using a Poincaré surface of section at periapsis.

The accuracy of the Keplerian map is demonstrated by the resemblance between the results of the map and fully integrated trajectories of the PCR3BP. The location and widths of the resonant islands appear to be in agreement. As discussed below, the geometrical features directing the motion of phase space regions are approximated well. The notable difference is the "warping" of trajectories of the map near $\omega = 0$, not seen in the full system. The full system Hamiltonian has a discrete time-reversal symmetry such that if (ω_0, a_0) maps to (ω_1, a_1) then $(-\omega_1, a_1)$ maps to $(-\omega_0, a_0)$. In our approximate map this symmetry is broken, albeit slightly.

Some remarks on the resonant structure of the map are in order. Transport in the map can be understood in terms of lobe dynamics and resonance zones [10, 25]. Lobes are parcels of phase space bounded by pieces of stable and unstable manifolds of hyperbolic points. The hyperbolic points of (4.2) occur at $s : r$ mean motion resonances, $a_{res} = (r/s)^{2/3}$, where in inertial space the moon orbits Jupiter in r complete circuits for every s particle orbits. In Figure 5, these appear as periodic points of period $r - s > 0$. These resonances are also known as being of order $r - s$. For every a_{res}, there is a band of at least $2(r-s)$ alternating elliptic and hyperbolic points, with stable islands (the holes in Figure 5) surrounding the elliptic points.

Let us consider the lowest order resonance in Figure 5, the first order 1:2 resonance. The period-one hyperbolic point (fixed point) corresponding to this resonance is located at $p_{res} = (\omega_{res}, a_{res})$, where $\omega_{res} = 0$ and $a_{res} = (2)^{2/3} \approx 1.587$. The stable and unstable manifolds for the hyperbolic point are shown in Figure 6. The shaded region is the resonance zone for this resonance, bounded by pieces of upper and lower branches of the stable and unstable manifolds, from the point p_{res} to a primary intersection point of the manifolds. The primary intersection points are also homoclinic orbits doubly asymptotic to p_{res}.

The unstable manifold is produced by first finding the unstable direction in the neighborhood of p_{res} and mapping forward a small seed of points along this direction using F. Linearizing F in the neighborhood of p_{res}, we obtain

$$\begin{pmatrix} \delta\omega_{n+1} \\ \delta K_{n+1} \end{pmatrix} = \begin{pmatrix} 1 + \mu\beta\gamma & -\gamma \\ -\mu\beta & 1 \end{pmatrix} \begin{pmatrix} \delta\omega_n \\ \delta K_n \end{pmatrix}, \tag{5.1}$$

where $\beta = -\frac{\partial f}{\partial \omega}|_{\omega=\omega_{res}}$ and $\gamma = 6\pi(a_{res})^{5/2}$. The unstable manifold is locally tangent to the eigenvector belonging to the maximum eigenvalue,

$$\lambda = \frac{1}{2}\left(2 + \mu\beta\gamma + \sqrt{\mu\beta\gamma(\mu\beta\gamma + 4)}\right). \tag{5.2}$$

The stable manifold is produced similarly, substituting F^{-1} for F in the above procedure.

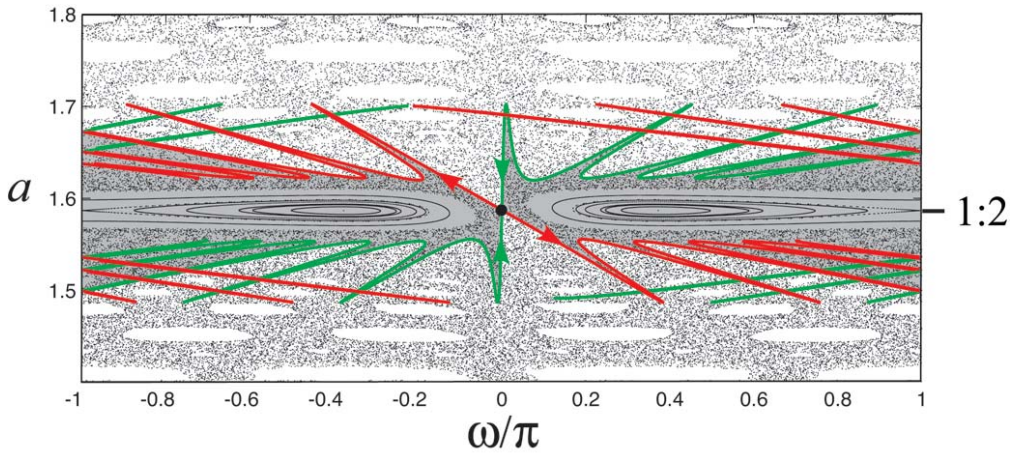


Figure 6. *Against the background of Figure 5, the stable and unstable manifolds of the central black point are plotted, the period-one hyperbolic fixed point* $p_{res} = (\omega_{res}, a_{res}) = (0, (2)^{2/3})$ *corresponding to the* 1:2 *mean motion resonance. The closed curves in the gray region are restricted to the stable island and disconnected from the connected chaotic sea.*

A similar picture of intersecting manifolds exists around each horizontal resonance zone encompassing the stable resonance holes and is similar to the resonance manifolds computed in the full equations [17]. The manifolds of different resonances intersect one another, providing the template for the migration of orbits through semimajor axis.

Finding the orbits which yield maximum change in semimajor axis. We consider the following problem: for a given $a_0 = -1/(2K_0)$, find the ω_0 which yields the maximum change $|a_n - a_0|$ after n periapses. Given our map F, a diffeomorphism of the cylinder to itself, we can consider iterates of the circle at a_0,

$$\Gamma_0 = \{(\omega_0, a_0) \in \mathcal{A} \mid \omega_0 \in S^1\}. \tag{5.3}$$

In Figure 7(a), we plot Γ_0 and its images $\Gamma_n = F^n(\Gamma_0)$, $n \geq 1$, in terms of the change in semimajor axis $\Delta a = a_n - a_0$ versus ω_0. The calculations are for a particle with $a_0 = 1.54$, $C_J = 3$ in the Jupiter–Callisto system, using $\bar{a} = 1.35$ for the map (4.2). The figure shows the effect of multiple periapses, using successive magnifications to reveal the complex self-similar structure as we follow the region with the greatest decrease in semimajor axis.

The function $\Delta a_n(\omega_0)$ gets very complex even for small n. For increasing n, $\max_{\omega_0} \Delta a_n$ and $\min_{\omega_0} \Delta a_n$ have increased magnitude, and the domains of the largest changes get thinner. We can estimate the size of these domains as follows. Suppose the local maxima and minima spike features at iterate n have a minimum width $\alpha(n)$ in ω. We can estimate $\alpha(1)$ from the kick function f and then assume that $\alpha(n) = [\alpha(1)]^n$. From Figure 4, it is reasonable to approximate $\alpha(1)$ as $2\omega_{max}$. For $\bar{a} = 1.35$, we have $\omega_{max} = 0.01\pi$; thus $\alpha(n) = (0.0628)^n$. From simulation, we find that this is a very conservative lower bound. Nevertheless, for computations to resolve the thin features at iterate n, we use adaptive refinement of a mesh of sample points, described elsewhere [9].

Figure 7(b) shows an example of a trajectory which quickly decreases semimajor axis over a duration of 25 orbits. This trajectory corresponds to an initial condition chosen such that it

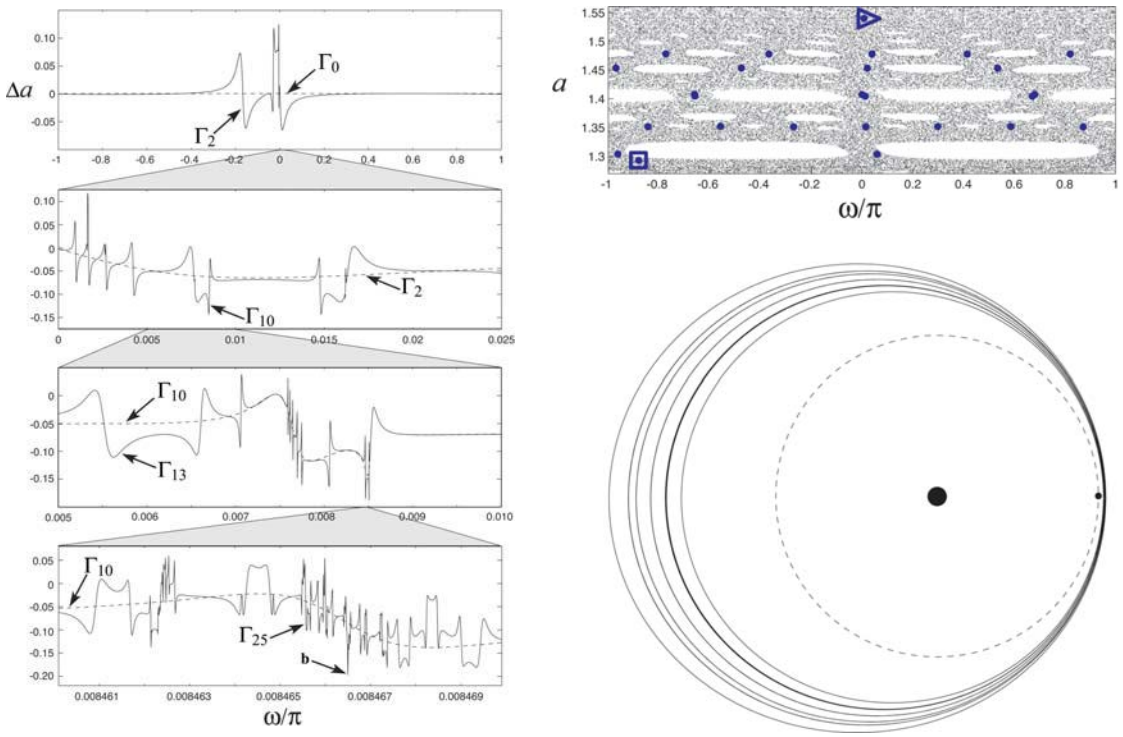
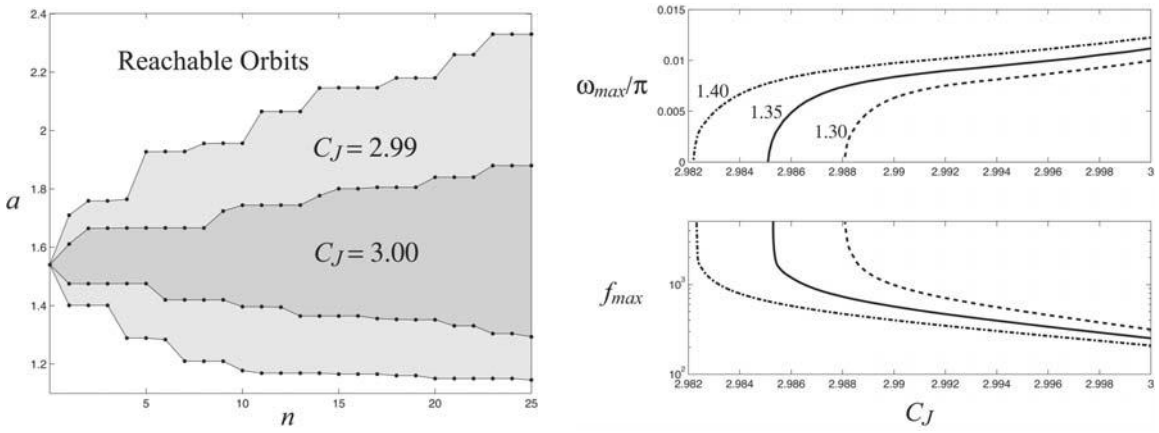


Figure 7. (a) *The total change in semimajor axis Δa after n periapsis passages is shown versus the initial angle, ω, where $(\mu, C_J, \bar{a}) = (5.667 \times 10^{-5}, 3, 1.35)$. In the top panel, the initial circle Γ_0 and its second image $\Gamma_2 = F^2(\Gamma_0)$ are shown. The second panel from the top shows a magnification and the tenth image, and so on. The bottom panel shows a portion of ω with Γ_{25}. We see small regions of significantly decreased semimajor axis after 25 periapsis passages.* (b) *Upper panel: The phase space trajectory of the trajectory marked as* **b** *in* (a). *The initial point is marked with a triangle and the final point with a square. Lower panel: The configuration space projections in an inertial frame for this trajectory. Jupiter and Callisto are shown at their initial positions, and Callisto's orbit is dashed. The particle migration is from larger to smaller semimajor axes, keeping the periapsis direction roughly constant in inertial space. Both the particle and Callisto orbit Jupiter in a counterclockwise sense.*

repeatedly experiences a periapsis kick near a minimum of the kick function; see Figure 4(a). For a randomly chosen initial condition, the effect of such kicks tends to average to zero. But trajectories like the one shown in Figure 7(b) can be found which exhibit large increases or decreases over small times—potentially useful trajectories for space missions.

Multiple gravity assists outside sphere of influence. We note that over the course of these multiple gravity assists, the particle does not come within the sphere of influence of the perturber. For the example in Figure 7(b), the particle at closest approach to the perturber is at a nondimensional distance of $r_{2,min} = 0.0341$, whereas the sphere of influence, approximated as the Hill's radius, is $r_h = (\mu/3)^{1/3} = 0.0266$. The phenomenon involved here is not the typical picture of a gravity assist from the patched-conic perspective, wherein a particle's path enters the sphere of influence of a perturber and can be approximated as a hyperbolic Keplerian trajectory with respect to the perturber [1].



(a) (b)

Figure 8. (a) *The maximum and minimum semimajor axes reached as function of the number of orbits for a trajectory starting at* $a_0 = 1.54$. *For* $C_J = 3.00$, *the reachable orbits are those in the darker shaded zone, and those for* $C_J = 2.99$ *are in the lighter shaded zone. The parameters for the map are the same as in the preceding figures,* $\mu = 5.667 \times 10^{-5}$ *and* $\bar{a} = 1.35$. (b) *The variation in the location and magnitude of the maximum kick with Jacobi constant* C_J *for three values of* $\bar{a}$ *as labeled (same line labeling for both panels). Values are independent of* μ.

6. Reachable orbits, capture, and escape. Related to the previous discussion is the question of what is the reachable set of orbits as a function of n for orbits initially on the circle Γ_0? In Figure 8(a), the maximum and minimum semimajor axes reached as a function of the number of orbits are plotted for two values of the Jacobi constant. As particles migrate from an initial semimajor axis through resonance zones via lobe dynamics, the set of reachable orbits grows. The maximum (minimum) semimajor axis as a function of n is monotonically increasing (decreasing). Although our map (4.2) lacks the time-reversal symmetry, we can consider it to be close to the time-reversal symmetry of the full PCR3BP equations. In the full equations, the reachable orbit set could be extended to $n < 0$ and would be the mirror image of the $n > 0$ set. This implies that if an orbit of semimajor axis a_n can be reached in n orbits from a_0, an orbit of semimajor axis $a_{-n} = a_n$ can reach a_0 in n orbits.

For $C_J = 2.99$, the growth is more rapid than for $C_J = 3$. For lower values of Jacobi constant (higher three-body energies, per (2.5)), we expect migration in the phase space to be faster since the kick function yields larger maximum kicks; see Figure 8(b).

Variation of ω_{max} and f_{max} with $\bar{a}$ is also shown in Figure 8(b). For each $\bar{a}$, there is a critical value C_J^* corresponding to a periapsis distance of 1 and a singularity of the map. Below C_J^*, the kick function changes its character as the particle can now cross into the interior of the perturber's orbit, a regime investigated in [16].

Exits leading to capture. We can consider what the limits to the growth are for the reachable set, in terms of a lower and upper bound in a. We first consider the lower bounds, and we consider the full equations of the PCR3BP. For our case of interest, case 3 with $C_J < C_2$, a particle beginning in the exterior realm must remain there for all time in the absence of an outside perturbation (see Figure 3); it cannot collide with or enter the phase space realm around m_2. For C_J below this value, a bottleneck region opens up around the Lagrange

point L_2, permitting particles to enter into orbit around m_2. In what follows, we summarize the mechanism for this capture as discussed in detail in [14]. On each three-dimensional energy surface with $C_J < C_2$, within the L_2 bottleneck region there is a planar clockwise orbit surrounding the location of L_2, shown as a dashed curve in Figure 9(a). This orbit has two-dimensional stable and unstable manifolds, with cylindrical ($S^1 \times \mathbf{R}$) geometry. As two-dimensional objects in a three-dimensional energy surface, these cylinders partition the energy surface, acting as separatrices for the flow through the bottleneck region [7, 18]. The interior of these cylinders in the energy surface form three-dimensional tubes of trajectories, termed Conley–McGehee tubes [17]. Only particles inside the tubes will move from the exterior realm to the m_2 realm and vice versa; those outside the tubes will not.

The capture branch of Conley–McGehee tubes associated to the L_2 bottleneck is shown in Figure 9(a) as projected onto the configuration space. A tube projection appears as a strip of varying width. Trajectories within the tubes wind around them in phase space and their projection appears similar to the example trajectory within the tube in Figure 9(a). There is also an escape branch (not shown) which appears as the mirror image of the capture branch, reflected across the horizontal axis.

In order to find capture trajectories, we consider the Poincaré surface of section taken at periapsis. In terms of the Delaunay (action-angle) variables (L, G, l, ω), where $L = \sqrt{a}$ and l is the mean anomaly, the surface of section at periapsis in the exterior realm is defined as

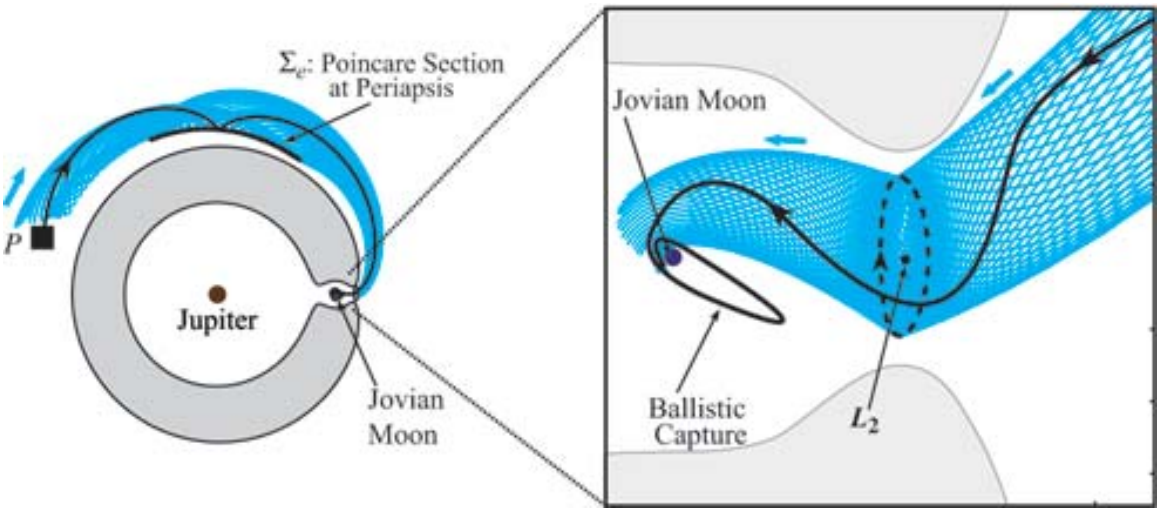
$$\Sigma_e = \{(\omega, a) \in \mathcal{A} \mid l = 0,\, a > 1\}, \tag{6.1}$$

where the condition of periapsis $l = 0$ is equivalent to setting the true anomaly ν to zero. The Hamiltonian flow induces a Poincaré return map on Σ_e, $\mathcal{F} : \Sigma_e \to \Sigma_e$ defined for almost all points on Σ_e. In Σ_e, the last cross-section of a tube before it enters the realm around m_2 appears as an exit, diffeomorphic to a disk, as shown schematically in Figure 9(b) and numerically in Figure 9(c). When trajectories of $\mathcal{F}$ reach the exit, they are transported to the realm around m_2, where we can consider them emerging within the entrance on Σ_{m_2}, a suitably defined Poincaré surface of section in the m_2 realm.

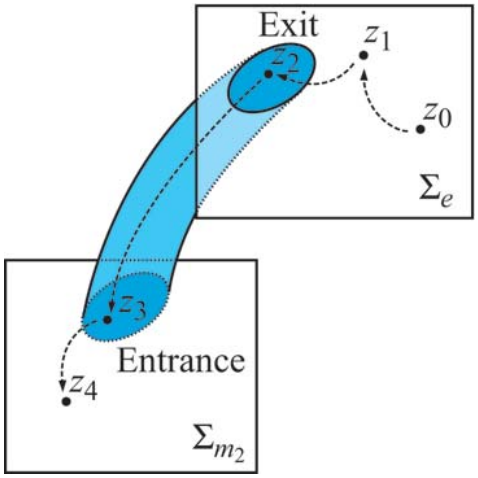
The Keplerian map F defined in (4.2) is an approximation to $\mathcal{F}$. When trajectories of F reach the exit, the Keplerian map approximation breaks down and the full equations of motion must be considered. The trajectory can no longer be approximated as near-Keplerian around the central body; it will instead be in orbit about the perturber. We can consider the location of an exit in (ω, a) space (in Σ_e) to give us an effective lower bound in the growth of a reachable set when $C_J < C_2$.

Escaping to infinity, upper bounds, and rotational invariant circles. For large values of μ, there may not be an upper bound to the reachable set as n increases. Numerically, we have found some particles which escape onto unbound parabolic and hyperbolic orbits ($K \geq 0$) in finite time ($n < \infty$) from orbits with relatively small a. The set of bound orbits in (ω, K) space which will become unbound after their next periapsis passage is given by lobes bounded above by $K = 0$ and below by $K = -\mu f(\omega)$ when $f(\omega) < 0$, as illustrated in Figure 10(a).

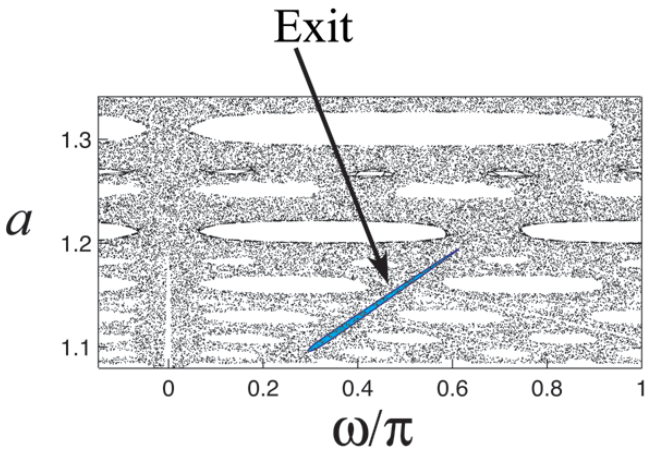
If μ is smaller than a critical value, circulating trajectories lying on invariant circles may exist, forming an upper boundary. McGehee [18] proved that for small μ in the PCR3BP, the energy surface is broken up into regions bounded by invariant tori. These invariant tori



(a)



(b)



(c)

Figure 9. (a) *A spacecraft P inside a tube of gravitational capture orbits will find itself going from an orbit about Jupiter to an orbit about a moon, as shown schematically. The spacecraft is initially inside a tube whose boundary is the stable invariant manifold of a periodic orbit about L_2. The tube, made up of individual trajectories, is shown as projected onto configuration space.* (b) *Poincaré sections in different realms—in this case in the exterior and m_2 realms, Σ_e and Σ_{m_2}, respectively—are linked by tubes in the phase space which live in surfaces of constant energy (C_J = constant). Under the Poincaré map on Σ_e, a trajectory $z_0, z_1, \ldots$ reaches an exit; the cross-section of the tube of capture orbits at the final periapsis before passage through the L_2 bottleneck. Under the Hamiltonian flow, points in the exit of Σ_e map to the entrance of Σ_{m_2}. The trajectory then evolves under the action of the Poincaré map on Σ_{m_2}.* (c) *The numerically computed location of the exit on Σ_e. The axes are the argument of periapsis in the rotating frame ω and the semimajor axis a of the instantaneous conic orbit about Jupiter, as in earlier figures. The location of the exit in configuration space is labeled in the left panel of* (a). *This surface of section was generated using the full equations of motion with $\mu = 5.667 \times 10^{-5}$ and $C_J = 3.005$.*

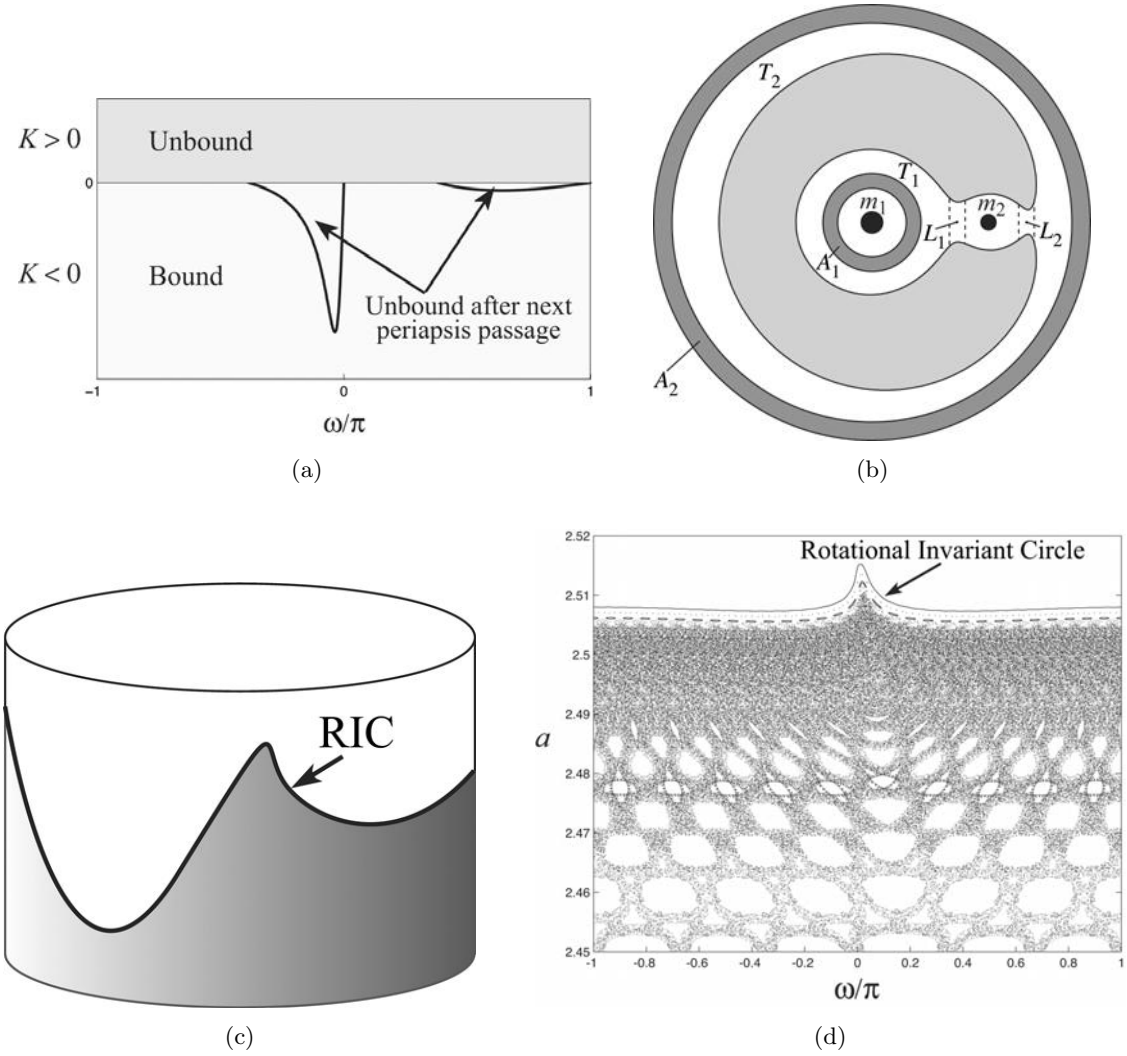


Figure 10. (a) *Bound trajectories enclosed by the lobes will become unbound after the next periapsis passage.* (b) *The projection of invariant tori (darkly shaded) on position space for case* 3. (c) *A rotational invariant circle (RIC) is an invariant loop that encircles the cylinder, i.e., a loop that cannot be contracted to a point.* (d) *A stable circulating trajectory forms an upper boundary to the chaotic sea, preventing particles from migrating to large a values. The calculation was done using the Keplerian map approximation with parameter values* $(\mu, C_J, \bar{a}) = (5 \times 10^{-6}, 3, 2.5)$.

project onto the darkly shaded annuli A_1 and A_2 shown for case 3 in Figure 10(b). These annuli separate the Hill's region into sections corresponding to the invariant regions in the energy surface. For case 3, masses m_1 and m_2 are separated from each other by an invariant torus; thus making it impossible for a particle to pass from arbitrarily close to m_1 to arbitrarily close to m_2. Similarly, the two masses are separated from infinity by an invariant torus; i.e., the exterior realm phase space is divided by a transport barrier whose projection onto configuration space is A_2. Let us call T_2 that part of the exterior realm outside a neighborhood of L_2 which extends up to the bounding surface A_2.

We find such a boundary in the Keplerian map for small μ values, understood from the point of view of absolute transport barriers in symplectic twist maps of the cylinder, following [19]. An invariant circle is a curve $\mathcal{C}$ such that $F(\mathcal{C}) = \mathcal{C}$. A rotational invariant circle (RIC) is a closed loop that encircles the cylinder (i.e., it cannot be contracted to a point; see Figure 10(c)). Birkhoff's theorem [3] implies that any RIC must be the graph of a function $a = A(\omega)$. An RIC divides the cylinder into two invariant regions. In other words, it prevents transport between the upper and lower "halves" of the cylinder. RICs are in fact the only absolute barriers to transport for symplectic twist maps of the cylinder. Therefore, in connected chaotic sets such as those shown in Figure 5 where $a_{min} < a < a_{max}$, we can say that no RIC can exist entirely in that portion of the cylinder.

If we find an RIC for our map, we have found an upper bound in the phase space beyond which particles with a greater than but close to 1 cannot pass. As a computational experiment motivated by calculations of [21], we consider the Keplerian map for $(\mu, C_J, \bar{a}) = (5 \times 10^{-6}, 3, 2.8)$. We find an orbit at the top of Figure 10(d) which marches around the cylinder, densely covering a circle. This is an RIC at the lower edge of the 1:4 resonance island and is therefore a bounding surface. By the smoothness of the map F in its parameter μ, we can expect a bounding surface for all μ below at least 5×10^{-6} for fixed C_J and $\bar{a}$.

This result is based on the structure of symplectic twist maps and not the KAM theorem [19]. However, we note that in our context the KAM theorem implies that RIC's present in the unperturbed integrable area-preserving map, i.e, (4.2) with $\mu = 0$,

$$\begin{pmatrix} \omega' \\ K' \end{pmatrix} = \begin{pmatrix} \omega - 2\pi(-2K)^{-3/2} \pmod{2\pi} \\ K \end{pmatrix}, \tag{6.2}$$

will persist under small area-preserving perturbations of the unperturbed map. For the unperturbed map, all trajectories lie on RIC's. The perturbed map for small μ is written

$$\begin{pmatrix} \omega' \\ K' \end{pmatrix} = \begin{pmatrix} \omega - 2\pi(-2K)^{-3/2} - \mu g(\omega, K) \pmod{2\pi} \\ K + \mu f(\omega) \end{pmatrix}, \tag{6.3}$$

where $g(\omega, K) = 6\pi f(\omega)(-2K)^{-5/2}$. In order for there to be invariant circles, we need the average of $f(\omega)$ over ω to be zero,

$$\int_{-\pi}^{\pi} f(\omega)\, d\omega = 0, \tag{6.4}$$

which our kick function satisfies, being odd in ω. The KAM theorem for our problem takes the form of [20], which says that for sufficiently small perturbations, most RICs will persist.

7. Discussion and conclusions. Using Picard's method of successive approximations, we derive a family of two-dimensional symplectic twist maps to approximate a particle's motion in the planar circular restricted three-body problem (PCR3BP) with Jacobi constant near 3. The maps model a particle on a near-Keplerian orbit about a central body of unit mass, where the spacecraft is perturbed by a smaller body of mass μ. The interaction of the particle with the perturber is modeled as an impulsive kick at periapsis passage, encapsulated in a kick function f. The maps are identified as an approximation of a Poincaré return map of

the full equations of motion where the surface of section is taken at periapsis, mapping each periapsis passage to the next in terms of ω, the azimuthal separation of the particle and small perturbing body, and K, the Keplerian orbital energy of the particle about the central body.

The map captures well the dynamics of the full equations of motion; namely, the phase space is densely covered by chains of stable resonant islands, in between which is a connected chaotic zone. The chaotic zone, far from being structureless, contains lanes of fast migration between orbits of different semimajor axes. The advantage of having an analytical two-dimensional map over full numerical integration is that we can apply all the machinery of the theory of transport in symplectic twist maps [24, 19, 9].

An interesting consequence of the approximation used to construct the map is that the time-reversal symmetry of the original Hamiltonian system is broken, albeit slightly. The origins of this symmetry breaking and development of a modified approximation process to avoid it are currently under investigation.

Other authors have considered similar maps, so-called Keplerian maps, to study the long-time evolution of nearly parabolic comets and comet-like objects [6, 36, 21]. By including the dependence of the map on the Keplerian energy K, we have achieved one of the goals of [16], which allows us to consider the map for orbits of moderate eccentricities and semimajor axes.

Our application is different from these papers. We apply our Keplerian map to the identification of transfer trajectories applicable to spacecraft transfers in a planet-moon system. The use of subtle gravitational effects described by the map may be feasible for future missions to explore the outer planet-moon systems where the timescale of orbits is measured in days instead of years and low energy trajectories may be considered for intermoon transfers.

Physically, particles in the regime we study undergo multiple gravity assists of a different kind than the hyperbolic flybys of, say, the Voyager missions. The gravity assists we study are for particles on orbits with semimajor axes greater than the perturber's and whose periapsis passages occur close to, but beyond, the sphere of influence of the perturbing body (as conservatively estimated by the Hill sphere). The effect of gravity assists is largest for particles whose passages occur slightly behind (resp., in front of) the perturbing body, resulting in a larger (resp., smaller) semimajor axis. This makes the apoapsis distance grow (resp., shrink) while keeping the periapsis distance relatively unchanged.

Dramatic orbital changes result from repeated gravity assists which are timed such that changes accumulate steadily in one direction (e.g., steadily shrinking apoapsis distance). This process can be understood in terms of phase space transport between resonance zones, i.e., resonant gravity assist. It is a three-body problem phenomenon not amenable to a patched-conic approach. This work therefore fills a gap in the understanding of multiple gravity assist mission design, which has been successful for Jacobi constants (Tisserand parameters) much less than 3, where the subtle effects described here play little role [30].

This paper extends earlier work which considered the dynamical connection between resonances in the exterior realm and interior realm and their relationship to escape and capture from a planetary or satellite neighborhood [14]. With straightforward modifications, the method used here can be applied to orbits entirely in the interior realm, where the Poincaré map is taken at apoapsis (where the perturbation due to the small mass is greatest) instead of periapsis.

Future work will consider extension of the Keplerian map to include (i) out of plane

motion, i.e., a four-dimensional symplectic map; (ii) multiple perturbers; (iii) eccentric orbits for the perturbers; and (iv) control and uncertainty [29]. This will increase the tools available to space mission designers and may shed light on the mechanism by which some minor bodies and impact ejecta get handed off between planets and moons of the solar system [11, 32, 13].

Given the success of the current application to celestial mechanics, we intend to investigate the general applicability of Picard's method of successive iterates to approximations of a Poincaré return map for perturbed Hamiltonian systems and other dynamical systems.

Acknowledgments. The authors thank Stefano Campagnola, Piyush Grover, Oliver Junge, Kartik Kumar, and Francois Lekien for interesting and insightful discussions, and anonymous referees for their helpful suggestions.

REFERENCES

[1] R. R. Bate, D. D. Mueller, and J. E. White, *Fundamentals of Astrodynamics*, Dover, New York, 1971.
[2] E. A. Belbruno and J. K. Miller, *Sun-perturbed Earth-to-Moon transfers with ballistic capture*, J. Guidance Control Dynam., 16 (1993), pp. 770–775.
[3] G. D. Birkhoff, *Surface transformations and their dynamical applications*, Acta Math., 43 (1920), pp. 1–119.
[4] M. Braun, *Differential Equations and Their Applications: An Introduction to Applied Mathematics*, 4th ed., Texts Appl. Math. 11, Springer-Verlag, New York, 1993.
[5] D. Brouwer and G. M. Clemence, *Methods of Celestial Mechanics*, Academic Press, New York, 1961.
[6] B. V. Chirikov and V. V. Vecheslavov, *Chaotic dynamics of comet Halley*, Astronom. and Astrophys., 221 (1989), pp. 146–154.
[7] C. C. Conley, *Low energy transit orbits in the restricted three-body problem*, SIAM J. Appl. Math., 16 (1968), pp. 732–746.
[8] C. C. Conley, *Twist mappings, linking, analyticity, and periodic solutions which pass close to an unstable periodic solution*, in Topological Dynamics, J. Auslander, ed., Benjamin, New York, 1968, pp. 129–153.
[9] M. Dellnitz, O. Junge, W. S. Koon, F. Lekien, M. W. Lo, J. E. Marsden, K. Padberg, R. Preis, S. D. Ross, and B. Thiere, *Transport in dynamical astronomy and multibody problems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (2005), pp. 699–727.
[10] R. W. Easton, J. D. Meiss, and S. Carver, *Exit times and transport for symplectic twist maps*, Chaos, 3 (1993), pp. 153–165.
[11] B. J. Gladman, J. A. Burns, M. Duncan, P. Lee, and H. F. Levison, *The exchange of impact ejecta between terrestrial planets*, Science, 271 (1996), pp. 1387–1392.
[12] G. Gómez, W. S. Koon, M. W. Lo, J. E. Marsden, J. Masdemont, and S. D. Ross, *Connecting orbits and invariant manifolds in the spatial three-body problem*, Nonlinearity, 17 (2004), pp. 1571–1606.
[13] J. Horner, N. W. Evans, M. E. Bailey, and D. J. Asher, *The populations of comet-like bodies in the Solar system*, Monthly Notices Roy. Astronom. Soc., 343 (2003), pp. 1057–1066.
[14] W. S. Koon, M. W. Lo, J. E. Marsden, and S. D. Ross, *Heteroclinic connections between periodic orbits and resonance transitions in celestial mechanics*, Chaos, 10 (2000), pp. 427–469.
[15] W. S. Koon, J. E. Marsden, S. D. Ross, and M. W. Lo, *Constructing a low energy transfer between Jovian moons*, in Celestial Mechanics (Evanston, IL, 1999), Contemp. Math. 292, AMS, Providence, RI, 2002, pp. 129–145.
[16] L. Malyshkin and S. Tremaine, *The Keplerian map for the planar restricted three-body problem as a model of comet evolution*, Icarus, 141 (1999), pp. 341–353.

[17] J. E. Marsden and S. D. Ross, *New methods in celestial mechanics and mission design*, Bull. Amer. Math. Soc. (N.S.), 43 (2006), pp. 43–73.
[18] R. McGehee, *Some Homoclinic Orbits for the Restricted Three-Body Problem*, Ph.D. thesis, University of Wisconsin, Madison, WI, 1969.
[19] J. D. Meiss, *Symplectic maps, variational principles, and transport*, Rev. Modern Phys., 64 (1992), pp. 795–848.
[20] J. Moser, *Stable and Random Motions in Dynamical Systems with Special Emphasis on Celestial Mechanics*, Princeton University Press, Princeton, NJ, 1973.
[21] M. Pan and R. Sari, *A generalization of the Lagrangian points: Studies of resonance for highly eccentric orbits*, Astronom. J., 128 (2004), pp. 1418–1429.
[22] M. E. Paskowitz and D. J. Scheeres, *Robust capture and transfer trajectories for planetary satellite orbiters*, J. Guidance Control Dynam., 29 (2006), pp. 342–353.
[23] T. Y. Petrosky and R. Broucke, *Area-preserving mappings and deterministic chaos for nearly parabolic motions*, Celestial Mech., 42 (1988), pp. 53–79.
[24] V. Rom-Kedar and S. Wiggins, *Transport in two-dimensional maps*, Arch. Rational Mech. Anal., 109 (1990), pp. 239–298.
[25] S. D. Ross, W. S. Koon, M. W. Lo, and J. E. Marsden, *Design of a multi-moon orbiter*, in 13th AAS/AIAA Space Flight Mechanics Meeting, Ponce, Puerto Rico, 2003, paper AAS 03-143.
[26] S. D. Ross, W. S. Koon, M. W. Lo, and J. E. Marsden, *Application of dynamical systems theory to a very low energy transfer*, in 14th AAS/AIAA Space Flight Mechanics Meeting, Maui, HI, 2004, paper AAS 04-289.
[27] D. J. Scheeres, *The effect of C_{22} on orbit energy and angular momentum*, Celestial Mech. Dynam. Astronom., 73 (1999), pp. 339–348.
[28] D. J. Scheeres, *Changes in rotational angular momentum due to gravitational interactions between two finite bodies*, Celestial Mech. Dynam. Astronom., 81 (2001), pp. 39–44.
[29] D. J. Scheeres, F.-Y. Hsiao, R. S. Park, B. F. Villac, and J. M. Maruskin, *Fundamental limits on spacecraft orbit uncertainty and distribution propagation*, J. Astronaut. Sci., to appear.
[30] N. J. Strange and J. M. Longuski, *Graphical method for gravity-assist trajectory design*, J. Spacecraft and Rockets, 39 (2002), pp. 9–16.
[31] V. Szebehely, *Theory of Orbits: The Restricted Problem of Three Bodies*, Academic Press, New York, 1967.
[32] M. Tiscareno and R. Malhotra, *The dynamics of known Centaurs*, Astronom. J., 126 (2003), pp. 3122–3131.
[33] B. F. Villac, *Dynamics in the Hill Problem with Applications to Spacecraft Maneuvers*, Ph.D. thesis, University of Michigan, Ann Arbor, MI, 2003.
[34] B. F. Villac and D. J. Scheeres, *On the concept of periapsis in Hill's problem*, Celestial Mech. Dynam. Astronom., 90 (2004), pp. 165–178.
[35] G. J. Whiffen, *An investigation of a Jupiter Galilean moon orbiter trajectory*, in AAS/AIAA Astrodynamics Specialist Conference, Big Sky, MN, 2003, paper AAS 03-544.
[36] J.-L. Zhou, Y.-S. Sun, J.-Q. Zheng, and M. J. Valtonen, *The transfer of comets from near-parabolic to short-period orbits: Map approach*, Astronom. and Astrophys., 364 (2000), pp. 887–893.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 3, pp. 597–644

How Do Synaptic Coupling and Spatial Temporal Delay Influence Traveling Waves in Nonlinear Nonlocal Neuronal Networks?*

Linghai Zhang†

Abstract. We are concerned with existence, uniqueness, and stability of the traveling wave of a nonlocal model equation which incorporates spatial temporal delay due to the finite propagation velocity of action potentials along axons. In particular, we investigate how wave shape, speed, and stability vary as the synaptic coupling and the model parameters change. The synaptic coupling may be of pure excitation, lateral inhibition, or lateral excitation. We introduce two concepts: the speed index function and the stability index function. One interesting point is that we can define the stability index function through the speed index function. By using this relationship, the stability of the traveling wave can be analyzed easily. These concepts (the speed index function and the stability index function) may play very important roles in rigorous mathematical analysis of traveling waves of nonlinear singularly perturbed systems of integral differential equations. The analysis and results on the speed, the speed index function, and the stability index function can be applied to dynamical systems and computational neuroscience.

Key words. nonlinear nonlocal neuronal network, synaptic coupling, spatial temporal delay, traveling wave front, speed index function, stability index function, existence and stability

AMS subject classifications. 35B25, 35R10, 92B20, 92C20

DOI. 10.1137/06066789X

1. Introduction.

The main focus of this paper is on mathematical analysis of an infinite-dimensional dynamical system, particularly the existence, uniqueness, and stability of a traveling wave front and analysis of its speed.

Since the fundamental work of the pioneers (e.g., Hodgkin and Huxley [26] and Wilson and Cowan [38]), neuronal network models have attracted many experts to investigate the rich dynamics underlying various spatial temporal pattern formations, such as traveling wave fronts, pulses, spirals, and target waves, as well as Hopf bifurcations and dynamic Turing Hopf instabilities. The formation and propagation of excitation/inhibition patterns in synaptically coupled neuronal networks play a very important role in understanding basic information processing in the nervous system; see [20] and [21]. Experimental results and numerical simulations of the brain demonstrate the occurrence of propagating patterns of activity in cortex neurons and in thalamus neurons; see Pinto and Ermentrout [29], [30], and Terman, Ermentrout, and Yew [36]. The propagating activities can appear spontaneously or can be caused by external stimulation.

Nonlinear traveling waves are of fundamental importance in neurobiology and in applied mathematics. Motivated by several important papers, we are concerned with a biophysically

*Received by the editors August 21, 2006; accepted for publication (in revised form) by B. Ermentrout March 13, 2007; published electronically August 1, 2007.
http://www.siam.org/journals/siads/6-3/66789.html
†Department of Mathematics, Lehigh University, 14 East Packer Avenue, Bethlehem, PA 18015 (liz5@lehigh.edu).

motivated nonlinear nonlocal firing rate model equation, which involves the convolution product of a kernel function and the Heaviside step function. The model also involves spatial temporal delay resulting from the finite conduction velocity of action potentials along axons. There are three parameters in the model, and all of them are closely related to nonlocal interactions between neurons in a network.

We are going to investigate how various neurobiological mechanisms (in particular, synaptic coupling and spatial temporal delay) influence traveling wave front, speed, and wave stability. We will be concerned with three kinds of synaptic couplings: pure excitation (modeled by a nonnegative kernel function), lateral inhibition (modeled by a Mexican hat kernel function), and lateral excitation (modeled by an upside down Mexican hat kernel function). We are not going to consider the influence of pure inhibition (modeled by a negative kernel function), as there exists no traveling wave in this case. We are concerned with asymptotic behaviors of the speed as the model parameters approach zero or infinity.

We introduce two concepts, the speed index function and stability index function (the latter is often referred to as the Evans function in PDEs and IDEs; see Sandstede [34]), and we plan to use them to investigate the relationship between the existence and the stability of the front. We will establish the existence, uniqueness, and stability of a traveling wave front of the model. We also will derive a formula on speed and then use examples to compare speeds corresponding to different kinds of kernels. These results can be applied to computational neuroscience and dynamical systems.

The speed index function is very interesting and important for the following reasons. There exists a unique solution to an equation involving the speed index function and the intrinsic parameters, and this unique solution is precisely the wave speed of the front; the details will be given later. Through this we will be able to investigate how the speed depends on various parameters as well as the synaptic coupling. Many estimates and asymptotic behaviors of the speed as the parameters approach zero or infinity can be investigated very effectively. By using properties of the speed index function, we are able to prove a simple but elegant identity, which connects the speed of the front of the model where there is a delay to the speed of the front where there is no delay. It is worth mentioning that Golomb and Ermentrout [20] have obtained a similar identity before. There is a nice relationship between the speed index function and the stability index function; see also Coombes [10]. By using this relationship, the stability index function can be analyzed easily. As is well known, the zeros of the stability index function coincide with the eigenvalues of an associated linear operator, and the eigenvalues determine wave stability/instability; see Coombes and Owen [8], Jones [27], Kapitula, Kutz, and Sandstede [28], Pinto, Jackson, and Wayne [31], Sandstede [34], and Zhang [40], [41], [42].

1.1. The model equation and its biological background. An important ingredient of neuronal activity is the input-output behavior of synapses which convert incoming pulses to postsynaptic potentials. In the coarse-grained population model, at position $\mathbf{x}$ and time t in the field, ensembles of excitatory and inhibitory chemical synapses respond to incoming pulse activity, yielding an effective postsynaptic potential $u(\mathbf{x}, t)$. In classical models of a homogeneous field, the dynamics of u can be described by nonlocal equations. Consider the following nonlinear scalar integral differential equation with spatial temporal delay:

$$u_t + u = \alpha \int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y}) H\left(u\left(\mathbf{y}, t - \frac{1}{c}|\mathbf{x}-\mathbf{y}|\right) - \theta\right) d\mathbf{y} + \mathcal{I}(\mathbf{x}, t), \tag{1}$$

where $0 < c \leq \infty$, $\alpha > 0$, and $\theta > 0$ are constants. The parameter α denotes the synaptic rate constant in a neuronal network, θ represents the threshold for excitation of the neuronal network, c represents the finite propagation speed of action potentials along axons, and $\frac{1}{c}|\mathbf{x}-\mathbf{y}|$ denotes the spatial temporal delay, where the distance is defined in the usual way: $|\mathbf{x}-\mathbf{y}| = \sqrt{|x_1-y_1|^2 + |x_2-y_2|^2 + \cdots + |x_n-y_n|^2}$, with $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T \in \mathbb{R}^n$.

In (1), $u(\mathbf{x}, t)$ stands for the average membrane potential (mean voltage activity level) of a neuron population at position $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$ and time t, and $H(u-\theta)$ denotes the output firing rate of a neuron. K represents synaptic coupling between neurons in the tissue (which is considered to be a continuous medium because the spatial density of the neurons is very large: it is approximately 10^4 neurons/mm^3). In addition to regular nonnegative symmetric kernels, K may be a Mexican hat kernel function or upside down Mexican hat kernel function, on-center or off-center, symmetric or asymmetric, and with or without compact support. Overall, the synaptic coupling in this work is very general. The symbol H denotes the Heaviside gain function: $H(u-\theta) = 0$ for all $u < \theta$, $H(\theta) = \frac{1}{2}$, and $H(u-\theta) = 1$ for all $u > \theta$. This gain function will make the model equation (1) analytically tractable. The convolution product represents nonlocal interactions between synapses through axon connections. In (1), $\mathcal{I}$ stands for an external applied current, such as $\mathcal{I}(\mathbf{x}, t) = \kappa \exp\left[-\lambda(\mathbf{v}\cdot\mathbf{x} + \mu t)^2\right]$ or $\mathcal{I}(\mathbf{x}, t) = 1 + \beta \tanh(\gamma \mathbf{v}\cdot\mathbf{x} + \delta t)$, where β, γ, δ, λ, μ, and κ are real constants and $\mathbf{v} = (v_1, v_2, \ldots, v_n)^T \in \mathbb{R}^n$ is a unit constant vector. This model was derived with the assumption that action potential propagation along axons with finite propagation velocity c is not damped and no additional constant delay is present. See Atay and Hutt [4], Coombes, Lord, and Owen [7], Coombes and Owen [8], Hutt and Atay [24], Ermentrout [12], and Pinto and Ermentrout [29], [30], for the same or very similar models. In particular, the following integral equation is closely connected to (1):

$$u(\mathbf{x}, t) = \int_{-\infty}^{t} \eta(t-s) \left[\alpha \int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y}) H\left(u\left(\mathbf{y}, s - \frac{1}{c}|\mathbf{x}-\mathbf{y}|\right) - \theta\right) d\mathbf{y} + \mathcal{I}(\mathbf{x}, s)\right] ds. \tag{2}$$

Here the synaptic input current $u(\mathbf{x}, t)$ is equal to the convolution product over space and time of the kernel function and the presynaptic firing rate function $H(u-\theta)$. It is straightforward to show that this equation is equivalent to (1) if the synaptic processing kernel η is chosen to be $\eta(t) = e^{-t}$ and the solution satisfies the boundary condition $\lim_{t\to-\infty}\left[e^t u(\mathbf{x}, t)\right] = 0$ for all real vectors $\mathbf{x}$.

For related wave problems modeled by reaction diffusion equations, see Aronson and Weinberger [2], [3] and Jones [27].

The observed waves of excitation/inhibition in synaptically coupled neuronal networks propagate more slowly than the action potentials along axons: the propagation speed of the traveling wave front is typically of the order 0.06 meters per second (m/s), while the action potential propagates along axons with a characteristic velocity of 0.5 m/s. Biologically, it makes sense to investigate the wave with speed μ so that $0 < \mu < c$, as Coombes and Owen [8] and Pinto and Ermentrout [29], [30] did in their papers. On the other hand, both biologically

and mathematically, there exists no wave solution with speed $\mu \leq 0$ or $\mu > c$ under the assumption $0 < 2\theta < \alpha$.

In many previous studies, the effect of the spatial temporal delay in neuronal networks has been neglected from mathematical analysis by letting $c \to \infty$ (i.e., the propagation of action potentials along axons is sufficiently fast). Note that for this special choice $c = \infty$, that is, there is no time delay; the field at point $\mathbf{y}$ has an instantaneous influence on the field at point $\mathbf{x}$. Then (1) reduces to

$$u_t + u = \alpha \int_{\mathbb{R}^n} K(\mathbf{x}-\mathbf{y}) H\big(u(\mathbf{y},t)-\theta\big) d\mathbf{y} + \mathcal{I}(\mathbf{x},t). \tag{3}$$

In [24] Hutt and Atay introduced a more general integral differential equation incorporating transmission velocity distribution:

$$u_t + u = \alpha \int_a^b \eta(c) \left[\int_{\mathbb{R}} K(x-y)\phi\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy \right] dc + \mathcal{I}(x,t), \tag{4}$$

where $\phi = H$ or ϕ has a sigmoidal shape, $\eta(c)$ is a statistical distribution of action potential velocities with compact support, $\eta \geq 0$ on (a,b), and $\int_a^b \eta(c)dc = 1$. Additionally, a and b are positive numbers denoting, respectively, the lower and upper bounds of biologically possible transmission velocities. They studied stability and bifurcations of equilibrium solutions and give a brief discussion on traveling wave solutions. Using a perturbation approach, they determined the effect of distributed velocities on bifurcations of equilibrium. In [25] Hutt considered the more interesting model equation

$$\begin{aligned} u_t + u = {} & \alpha \int_{\mathbb{R}} K(x-y) H\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy \\ & + \beta \int_{\mathbb{R}} J(x-y) H(u(y,t-\tau)-\theta) dy, \end{aligned}$$

where c and τ are the propagation velocity of action potential and feedback delay, respectively. Hutt examined briefly the dependence of the speed of the front on various parameters.

1.2. More previous results and open problems. There have been many very interesting research results on traveling wave solutions of these model equations. Based on different biophysical interpretations, by using concrete examples, through analytic approaches and numerical simulations, important properties of waves, such as their propagation speeds and their dependence on the parameters and degree of homogeneity of the network, have been established for (1)–(3) and the following neuronal network equations:

$$u_t + u + w = \alpha \int_{\mathbb{R}} K(x-y)\phi\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy, \tag{5}$$

$$w_t = \varepsilon u - \gamma w, \tag{6}$$

where $\phi(u) = H(u)$ or $\phi(u) = \frac{1}{1+\exp(-\beta u)}$ for some positive constant β, and the model

$$u_t + f(u) + w = \alpha \int_{\mathbb{R}} K(x-y) H(u(y,t)-\theta) dy + \mathcal{I}(x,t), \tag{7}$$

$$w_t = \varepsilon g(u) - \gamma w. \tag{8}$$

The systems (5)–(6) and (7)–(8) include the so-called effects of spike frequency adaptation. Here is a partial collection of previous results relevant to this work. Due to space limitation, it is impossible to give a full list. Amari [1] investigated the existence and stability of standing waves of (3) with lateral inhibition. Atay and Hutt [4] studied stability and Hopf bifurcations of equilibrium solutions of (1) for a general class of connectivity kernels. They give sufficient conditions for asymptotic stability. They also studied stationary and nonstationary Turing instability; the propagation delays play a significant role in nonstationary bifurcations of equilibrium. Bressloff [5], Bressloff and Folias [6], and Folias and Bressloff [18], [19] studied weakly interacting pulses in [5], front bifurcations in [6], breathing pulses in [18] as well as stimulus-locked traveling waves and breathers in [19] of system (7)–(8). Coombes, Lord, and Owen [7], Coombes and Owen [8], [9], and Coombes [10] investigated existence and stability of traveling wave fronts, traveling wave pulses, and time-independent bumps of scalar equations (2) and systems of integral equations (5)–(6), where $\mathcal{I} = 0$. In particular, Coombes and Owen [8] constructed Evans functions for the waves of (2) and (5)–(6) with Heaviside gain function and established the stability of traveling wave solutions. They used certain examples of positive, Mexican hat, and inverted Mexican hat kernels to model synaptic couplings. Moreover, some of their model equations [7] incorporate two types of gain functions: either the Heaviside step function or the sigmoid smooth function. In some of their papers the synaptic processing kernel $\eta(t) = \gamma \exp(-\gamma t)$ or $\eta(t) = \gamma^2 t \exp(-\gamma t)$, where $\gamma > 0$ is a constant. Kapitula, Kutz, and Sandstede [28] constructed Evans functions for traveling pulses of (7)–(8) as well as a nonlocal Schrödinger-type equation (the master mode-locking equation, that is, a model for the solid state cavity laser). Sandstede [34] proved the equivalence of linear stability and nonlinear stability of traveling pulses for (7)–(8). We hope that his ideas and results can be extended to this model. This may not be obvious, as the eigenvalue problem is nonlinear in terms of the eigenvalue parameter. Venkov, Coombes, and Matthews [37] used Laplace transform and Fourier–Laplace transform as well as Taylor expansions to establish pattern formation beyond a dynamic Turing instability (Turing–Hopf bifurcation). Assuming the existence and uniqueness of a monotonically increasing traveling wave front of (1) without inhomogeneity, i.e. $U(0) = \theta$, $U(-\infty) = 0$, $U(\infty) = \alpha$, $U' \geq 0$ on $\mathbb{R}$, and $U'(\pm\infty) = 0$, Pinto and Ermentrout [29], [30] found a nice relationship between the speed and the model parameters of (1). The spike frequency adaptation model was also investigated in great detail in the work of Pinto and Ermentrout [29], [30] and Pinto, Jackson, and Wayne [31]. Zhang [40] established the existence and stability of a front of (3) with $\mathcal{I} = 0$ and a back of (3) with $\mathcal{I} = 2\theta - \alpha$ for the special case $c = \infty$ with the case $c < \infty$ left open. Hopf bifurcations, Turing–Hopf bifurcations, and Turing dynamic instability of a standing pulse leading to a pair of traveling pulses are also investigated, either briefly or in great detail, in the work of [4], [5], [6], [8], [9], [10], and [37].

As we can easily see, the main body of the existence, uniqueness, and stability proofs of traveling waves for all kernels in the three classes (pure excitation, lateral inhibition, and lateral excitation) are still open. For example, there has been very little rigorous mathematical analysis on (1) with asymmetric, Mexican hat, or upside down Mexican hat kernel functions. The rigorous mathematical analysis and application of stability index functions for all three classes of kernel functions have been open for a long time. The propagation velocities of the action potentials along the axon depend on a number of factors, such as its myelination. The

myelin is a fatty material, composed mainly of lipids and lipoproteins, that encloses certain axons and nerve fibers and affects their electrical conductance. Natural diversity in the degree of myelination of the axons leads to a diversity in their propagation velocities. Thus it is of great biological interest to investigate how the wave speed changes as the velocity of the action potentials varies. Overall, the dependence of the speed on the kernel function and other parameters is a very important topic. Only very special cases (see Coombes, Lord, and Owen [7], Coombes and Owen [8], Pinto and Ermentrout [29], [30], and Zhang [40]) have been investigated. Our goal is to provide positive solutions to these open problems. The theoretical results of this paper (Theorem 4) partly validate the experimental results (modulation of neuronal threshold can speed up, slow down, and even block traveling waves in neocortical slices) conducted by Richardson, Schiff, and Gluckman [32]. We hope to find more real applications to their work.

1.3. Main goal. We will establish the existence and stability of traveling waves of (1) by applying fundamental ideas in differential equations and functional analysis (see [14], [15], [16], [17], [33], [39]) and by using the method of linearization. In particular, we will introduce two concepts: the speed index function and stability index function. We will derive formulas for the speed in terms of the given data, namely, the kernel function, as well as the parameters α, θ, and c. We will consider the asymptotic behaviors of the wave and the speed as the kernel function and the parameters tend to certain critical cases, such as $c \to 0$, $c \to \infty$, $\frac{\theta}{\alpha} \to 0$, and $\frac{\theta}{\alpha} \to \frac{1}{2}$. We will also demonstrate the valuable relation between the speed index function and stability index function (i.e., complex analytic function). This is extremely important for studying the stability of the traveling wave. Note that there have been many nice results using the so-called Evans function (i.e., the stability index function) to investigate stability of traveling waves; see [8], [9], [27], [28], [31], [34].

1.4. Mathematical assumptions. Suppose that the positive parameters α, θ, and c satisfy the conditions $0 < 2\theta < \alpha$ and $0 < c \le \infty$. To include the case $c \to \infty$, it is understood that $\frac{c-\mu}{c\mu} = \frac{1}{\mu}$ when $c = \infty$. We will consider three general classes of kernel functions. Note that the kernel function can be on-center (the maximum of K is attained only at the center $x = 0$) or off-center (the maximum of K is attained only at points other than the center $x = 0$), symmetric or asymmetric, with or without compact support. For each of the kernel functions mentioned below, suppose that K is continuous at $x = 0$, almost everywhere smooth, satisfying the conditions

$$\int_{\mathbb{R}} K(x)dx = 1, \qquad |K(x)| \le C\exp(-\rho|x|) \quad \text{on } \mathbb{R}, \tag{9}$$

for two positive constants C and ρ.

(A) The first class consists of nonnegative kernel functions (pure excitation), such as $K_1(x) = \frac{1}{2}(1 + \sin x)\exp(-|x|)$ and $K_2(x) = \sqrt{\frac{1}{\pi}}\exp(-x^2)$.

(B) The second class consists of Mexican hat kernel functions (lateral inhibition), that is, $K \ge 0$ on $(-M, N)$ and $K \le 0$ on $(-\infty, -M) \cup (N, \infty)$ for some positive constants M and N such that

$$\int_{-\infty}^{0} |x|K(x)dx \ge 0. \tag{10}$$

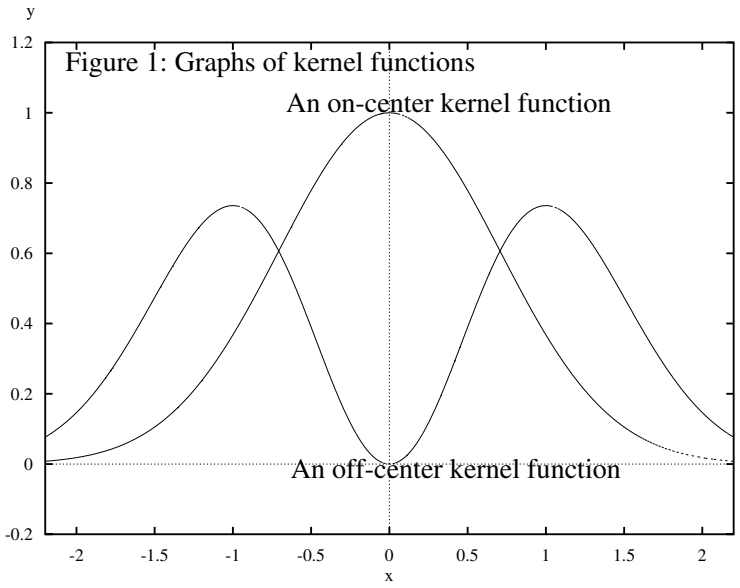


Figure 1. *Comparison of synaptic couplings (between an on-center kernel and an off-center kernel).*

This condition is sufficient to guarantee the existence and uniqueness of the speed.

(C) The third class consists of upside down Mexican hat kernel functions (lateral excitation), that is, $K \le 0$ on $(-M, N)$ and $K \ge 0$ on $(-\infty, -M) \cup (N, \infty)$, for two positive constants M and N, such that

$$\frac{\alpha}{2} + \alpha \int_0^N K(x)dx \ge \theta. \tag{11}$$

For examples, $K_5(x) = (1+\tanh x)[|x|\exp(-|x|)-2|x|\exp(-2|x|)]$ and $K_6(x) = |x|\exp(-\frac{1}{2}x^2)-2|x|\exp(-2x^2)$ belong to this class; see Figures 1 and 2. See Curtu and Ermentrout [11], Guo and Chow [22], [23], and Pinto and Ermentrout [29], [30] for the biological interpretations of these functions. For simplicity, we assume that

$$\int_{-\infty}^0 K(x)dx = \int_0^\infty K(x)dx = \frac{1}{2}. \tag{12}$$

These assumptions will make the notation and presentation very simple. Nevertheless, we should mention that the same results of this paper remain valid even if $0 < \int_{-\infty}^0 K(x)dx \ne \frac{1}{2}$ and $0 < \int_0^\infty K(x)dx \ne \frac{1}{2}$. Overall, these above conditions imply that excitation dominates inhibition in a neuronal network. Each of the integrals $\int_{\mathbb{R}} |K(x)|dx$ in the last two classes may be arbitrarily large, although $\int_{\mathbb{R}} K(x)dx = 1$.

1.5. Difficulty. When $c = \infty$ and $K \ge 0$ on $\mathbb{R}$, it is very easy to establish the existence, uniqueness, and stability of the traveling wave front together with a unique positive speed; see Pinto and Ermentrout [29], [30] and Zhang [40]. When $0 < c < \infty$ and $K \ge 0$ is not guaranteed, in particular for Mexican hat kernel functions and for upside down Mexican hat

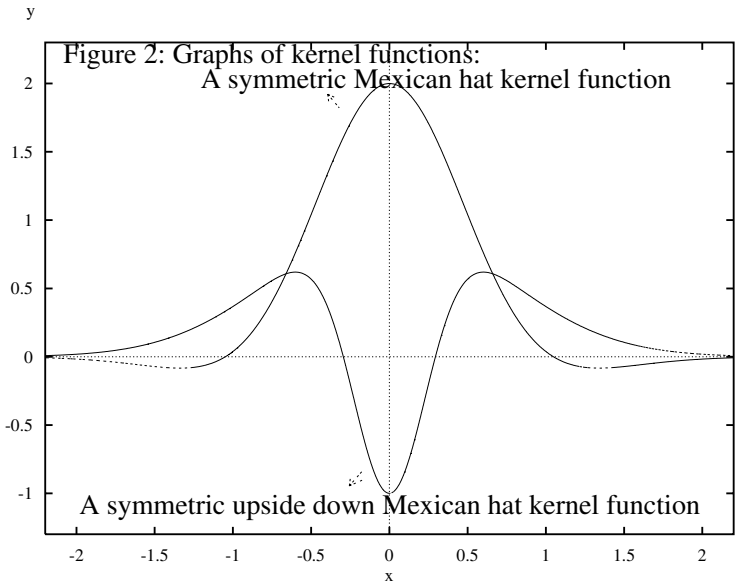


Figure 2. *Comparison of synaptic couplings (between a Mexican hat kernel and an upside down Mexican hat kernel).*

kernel functions, it is completely not straightforward to prove these results; even the existence and uniqueness of the speed is very difficult. We need to impose certain conditions such as those in (9)–(11). When we investigate the stability of the wave, we must solve an eigenvalue problem. It turns out that the eigenvalue problem corresponding to (1) is nonlinear in terms of the eigenvalue parameter λ. We have to overcome certain difficulties to investigate the eigenvalues. We will extend the mathematical analysis and application of the speed index function and the stability index function from nonnegative symmetric kernels to very general kernel functions.

2. Existence analysis of the traveling wave front. We have to make rigorous mathematical analysis before providing delicate estimates and limits on wave speeds. The analysis on speed given in later sections can be potentially applied to dynamical systems and neuroscience. Setting $n = 1$ and $\mathcal{I}(x,t) \equiv \mathcal{I}$, which is a constant independent of x and t, without loss of generality, let $\mathcal{I} \equiv 0$ in this section. Therefore, we will consider the homogeneous scalar integral differential equation

$$u_t + u = \alpha \int_{\mathbb{R}} K(x-y) H\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy. \tag{13}$$

The model (13) has two constant solutions: $U_0 \equiv 0$ and $U_1 \equiv \alpha$. The constant solutions exist only when $0 < \theta < \alpha$. They play a fundamental role for our existence and uniqueness problem. Traveling waves of (13) are functions of the form $u(x,t) = U(x + \mu t)$, where $z = x + \mu t$ and μ represents a speed, satisfying certain boundary conditions at plus/minus infinity, and the

equation

$$\mu U' + U = \alpha \int_{\mathbb{R}} K(z-y) H\left(U\left(y - \frac{\mu}{c}|z-y|\right) - \theta\right) dy.$$

A fundamental property of the traveling wave is the so-called translation invariance: if $U = U(z)$ is a solution of (13), then $\widetilde{U} = U(z+h)$ is also a solution of that equation, for any real constant $h \in \mathbb{R}$. Another important property is the exponential convergence at plus/minus infinity: there exist two positive constants C and ρ such that

$$\begin{aligned} |U(z)| &\le C\exp(-\rho|z|) && \text{for all } z < 0, \\ |U(z) - \alpha| &\le C\exp(-\rho|z|) && \text{for all } z > 0, \\ |U'(z)| &\le C\exp(-\rho|z|) && \text{for all } z \in \mathbb{R}. \end{aligned}$$

The main goal of this section is to use the sufficient conditions provided earlier to establish the existence and uniqueness (up to translation invariance) of the traveling wave front of the model equation (13) with general kernel functions described in classes (A), (B), and (C). We will also establish the relationship between the speed and the model parameters (in particular, the kernel function and the constants α, θ, and c).

The main goal of this section is to state the existence and uniqueness of the front together with the speed. Some asymptotic behaviors of the front and the speed of the wave as the parameters approach certain finite numbers or infinity are also presented.

2.1. Existence and uniqueness of the traveling wave front.

Theorem 1. *For any kernel function K in the three classes* (A), (B), *and* (C), *and for any positive parameters α, θ, and c including the case $c = \infty$, if $0 < 2\theta < \alpha$, then there exists a unique traveling wave front $U = U(c, \mu_0, \cdot)$ together with a unique speed $\mu_0 = \mu_0(c, \alpha, \theta)$ such that $0 < \mu_0 < c$.* (I) *The front satisfies the traveling wave equation*

$$\mu_0 U' + U = \alpha \int_{\mathbb{R}} K(z-y) H\left(U\left(y - \frac{\mu_0}{c}|z-y|\right) - \theta\right) dy, \tag{14}$$

the phase conditions $U(c, \mu_0, 0) = \theta$ and $U_z(c, \mu_0, 0) > 0$, and $U < \theta$ on $(-\infty, 0)$ and $U > \theta$ on $(0, \infty)$, and the following boundary conditions:

$$\lim_{z\to-\infty} U(c, \mu_0, z) = 0, \qquad \lim_{z\to\infty} U(c, \mu_0, z) = \alpha, \qquad \lim_{z\to\pm\infty} U_z(c, \mu_0, z) = 0. \tag{15}$$

(II) *Furthermore,*

$$\lim_{c\to 0} \mu_0(c, \alpha, \theta) = 0, \qquad \lim_{c\to\infty} \mu_0(c, \alpha, \theta) = \nu_0(\alpha, \theta), \tag{16}$$

$$\lim_{c\to 0} U(c, \mu_0, z) = \left[\frac{\alpha}{2} + \alpha \int_0^{z/2} K(x)dx\right] H(z), \tag{17}$$

$$\lim_{c\to\infty} U(c, \mu_0, z) = \alpha \int_{-\infty}^{z} K(x)dx - \alpha \int_{-\infty}^{z} \exp\left(\frac{x-z}{\nu_0}\right) K(x)dx, \tag{18}$$

where $\mu_0(c,\alpha,\theta)$ *and* $\nu_0(\alpha,\theta)$ *satisfy*

$$\alpha\int_{-\infty}^{0}\exp\left[\frac{c-\mu_0(c,\alpha,\theta)}{c\mu_0(c,\alpha,\theta)}\xi\right]K(\xi)d\xi=\frac{\alpha}{2}-\theta \tag{19}$$

and

$$\alpha\int_{-\infty}^{0}\exp\left[\frac{1}{\nu_0(\alpha,\theta)}\xi\right]K(\xi)d\xi=\frac{\alpha}{2}-\theta, \tag{20}$$

respectively. (III) *Additionally, if we define the function* $\mathcal{U}$ *by* $\mathcal{U}(z)=0$ *on* $(-\infty,0)$, $\mathcal{U}(0)=\frac{\alpha}{2}$, *and*

$$\mathcal{U}(z)=\frac{\alpha}{2}+\alpha\int_0^{z/2}K(x)dx-\frac{\alpha}{2}\int_0^z\exp\left(\frac{x-z}{c}\right)K\left(\frac{x}{2}\right)dx$$

on $(0,\infty)$, *then*

$$\lim_{\frac{\theta}{\alpha}\to 0}\mu_0(c,\alpha,\theta)=c,\qquad \lim_{\frac{\theta}{\alpha}\to\frac{1}{2}}\mu_0(c,\alpha,\theta)=0, \tag{21}$$

$$\lim_{\frac{\theta}{\alpha}\to 0}U(c,\mu_0,z)=\mathcal{U}(z),\qquad \lim_{\frac{\theta}{\alpha}\to\frac{1}{2}}U(c,\mu_0,z)=\alpha\int_{-\infty}^{z}K(x)dx. \tag{22}$$

See Appendix 1 (section 6.1) for the complete proof of Theorem 1.

Remark 1. Under the above conditions, there also exists a unique monotone front with the negative speed $-\mu_0$ such that

$$\lim_{z\to-\infty}U(c,-\mu_0,z)=\alpha,\qquad \lim_{z\to\infty}U(c,-\mu_0,z)=0,\qquad \lim_{z\to\pm\infty}U_z(c,-\mu_0,z)=0.$$

To keep the analysis and results clear, we will focus on the wave front with the positive speed μ_0.

2.2. The speed index function—Existence and uniqueness of the speed. We now define the so-called speed index function φ on the open domain $(0,c)$ by

$$\varphi(\mu)=\alpha\int_{-\infty}^{0}\exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi. \tag{23}$$

This function will play a pivotal role in this paper. Clearly we have

$$\varphi'(\mu)=\frac{\alpha}{\mu^2}\int_{-\infty}^{0}|\xi|\exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi,$$

$$\lim_{c\to\infty}\varphi(\mu)=\alpha\int_{-\infty}^{0}\exp\left(\frac{\xi}{\mu}\right)K(\xi)d\xi,$$

$$\lim_{c\to\infty}\varphi'(\mu)=\frac{\alpha}{\mu^2}\int_{-\infty}^{0}|\xi|\exp\left(\frac{\xi}{\mu}\right)K(\xi)d\xi.$$

Therefore the speed index function φ is defined on $(0,\infty)$ if $c=\infty$. Overall, it is a smooth function in its domain. To see the behavior of $\varphi'(\mu)$ when μ is very small, let us make the change of variable $\eta=\frac{\xi}{\mu}$. Then

$$\begin{aligned}\varphi'(\mu) &= \alpha\int_{-\infty}^{0}|\eta|\exp\left(\frac{c-\mu}{c}\eta\right)K(\mu\eta)d\eta\\ &\to \alpha\int_{-\infty}^{0}|\eta|\exp(\eta)K(0)d\eta=\alpha K(0)\end{aligned}$$

as $\mu\to 0$, provided that the kernel function K is at least piecewise continuous on $\mathbb{R}$. Therefore, $\varphi'(\mu)>0$ on $[0,\delta]$ for some constant $\delta>0$ if $K(0)>0$, and $\varphi'(\mu)<0$ on $[0,\delta]$ for some constant $\delta>0$ if $K(0)<0$. Obviously

$$\lim_{\mu\to 0^+}\varphi'(\mu)=0$$

if $K(0)=0$. However, we can easily establish that if $K(0)=0$, $K\le 0$ on $(-\infty,-M)$, and $K\ge 0$ on $(-M,0)$, then

$$\begin{aligned}\varphi'(\mu) &= \alpha\int_{-\infty}^{0}|\eta|\exp\left(\frac{c-\mu}{c}\eta\right)K(\mu\eta)d\eta\\ &> \alpha\exp\left(-\frac{c-\mu}{c\mu}M\right)\int_{-\infty}^{0}|\eta|K(\mu\eta)d\eta>0\end{aligned}$$

on $(0,c)$; if $K(0)=0$, $K\ge 0$ on $(-\infty,-M)$, and $K\le 0$ on $(-M,0)$, then $\varphi'(\mu)<0$ on $(0,\delta)$.

Similarly, we can study the behavior of $\varphi(\mu)$ as $\mu\to c$. Clearly

$$\lim_{\mu\to c}\frac{\alpha}{\mu^2}\int_{-\infty}^{0}|\xi|\exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi=\frac{\alpha}{c^2}\int_{-\infty}^{0}|\xi|K(\xi)d\xi>0$$

if K is a nonnegative or upside down Mexican hat kernel function. This limit is nonnegative if K is a Mexican hat kernel function. First, let us consider the case

$$\int_{-\infty}^{0}|\xi|K(\xi)d\xi>0.$$

We have $\varphi'(\mu)>0$ on $[c-\delta,c]$. Second, if $\int_{-\infty}^{0}|\xi|K(\xi)d\xi=0$, then we can also show that there exists a constant $0<\delta\ll 1$ such that $\varphi'(\mu)>0$ on $(c-\delta,c)$. Moreover, even if $c=\infty$, there exists a constant $0<\delta\ll 1$ such that $\varphi'(\mu)>0$ on $\left(\frac{1}{\delta},\infty\right)$.

Lemma 1. *There exists a unique positive number $\mu_0\in(0,c)$ such that*

$$\varphi(\mu_0)=\frac{\alpha}{2}-\theta. \tag{24}$$

Furthermore,

$$\lim_{c\to 0}\mu_0(c,\alpha,\theta)=0,\qquad \lim_{c\to\infty}\mu_0(c,\alpha,\theta)=\nu_0(\alpha,\theta), \tag{25}$$

$$\lim_{\frac{\theta}{\alpha}\to 0}\mu_0(c,\alpha,\theta)=c,\qquad \lim_{\frac{\theta}{\alpha}\to\frac{1}{2}}\mu_0(c,\alpha,\theta)=0. \tag{26}$$

See section 6.2 for the complete proof of Lemma 1.

Remark 2. For all kernel functions in classes (A) and (B), the speed index function

$$\varphi(\mu) = \alpha \int_{-\infty}^{0} \exp\left(\frac{c-\mu}{c\mu}\xi\right) K(\xi)d\xi$$

is strictly increasing on $(0, c)$.

Remark 3. For all kernel functions in class (C), $\varphi'(\mu_0) > 0$ if $\varphi(\mu_0) = \frac{\alpha}{2} - \theta > 0$. For any given upside down Mexican hat kernel function K, there exists a unique number $\mu_* = \mu_*(K) > 0$ such that

$$\varphi(\mu_*) = \alpha \int_{-\infty}^{0} \exp\left(\frac{c-\mu_*}{c\mu_*}\xi\right) K(\xi)d\xi = 0.$$

It turns out that μ_* is the lower bound for the speed, that is, $0 < \mu_* < \mu_0 < c$.

3. Speed index function and speed analysis. The main focus of this section is on speed analysis. First, we present some examples. Then we study the behaviors (estimates as well as limits) of the speed.

We have defined the speed index function φ earlier by

$$\varphi(\mu) = \alpha \int_{-\infty}^{0} \exp\left(\frac{c-\mu}{c\mu}\xi\right) K(\xi)d\xi$$

on $(0, c)$. The speed μ_0 of the front is the unique solution of the equation

$$\varphi(\mu) = \frac{\alpha}{2} - \theta.$$

The speed index function will be associated with the stability of the front of (13).

3.1. Examples of speed index function and speed. Nothing helps us understand how the wave speed depends on the intrinsic parameters better than lots of good examples. We also compare the speeds in the next four examples.

Example 1. Let $K(x) = \frac{\rho}{2}\exp(-\rho|x|)$, where $\rho > 0$ is a constant. This is a positive, on-center, symmetric kernel function in class (A). The speed index function and the speed of the front satisfy

$$\varphi(\mu) = \frac{c\alpha\rho\mu}{2\big(c\rho\mu + c - \mu\big)} \to \frac{\alpha\rho\mu}{2(1+\rho\mu)} \qquad \text{as } c \to \infty,$$
$$\mu_0 = \frac{(\alpha - 2\theta)c}{\alpha - 2\theta + 2c\rho\theta}.$$

Obviously, $0 < \mu_0 < c$ and $0 < \mu_0 < \frac{\alpha-2\theta}{2\rho\theta}$. There hold the following limits (here we are abusing the notation a bit: $\mu_0 = \mu_0(c, \alpha, \theta, \rho)$):

$$\lim_{\rho\to\infty} \mu_0(c,\alpha,\theta,\rho) = 0, \qquad \lim_{\rho\to 0} \mu_0(c,\alpha,\theta,\rho) = c,$$
$$\lim_{c\to\infty} \mu_0(c,\alpha,\theta,\rho) = \frac{\alpha-2\theta}{2\rho\theta}, \qquad \lim_{c\to 0} \mu_0(c,\alpha,\theta,\rho) = 0.$$

Example 2. Let $K(x) = \frac{1}{2}\rho^2|x|\exp(-\rho|x|)$, where $\rho > 0$ is a constant. This is a nonnegative, off-center, symmetric kernel function also in class (A). The speed index function and the speed of the front satisfy

$$\varphi(\mu) = \frac{c^2\alpha\rho^2\mu^2}{2\big(c\rho\mu + c - \mu\big)^2} \to \frac{\alpha\rho^2\mu^2}{2(1+\rho\mu)^2} \qquad \text{as } c \to \infty,$$
$$\mu_0 = \frac{c\sqrt{\alpha - 2\theta}}{\sqrt{\alpha - 2\theta} + c\rho\big(\sqrt{\alpha} - \sqrt{\alpha - 2\theta}\big)}.$$

This speed is larger than the speed in Example 1 if the parameters c, α, θ, and ρ are held the same in both examples. There hold the estimates

$$0 < \mu_0 < c, \qquad 0 < \mu_0 < \frac{\sqrt{\alpha - 2\theta}}{\rho\big(\sqrt{\alpha} - \sqrt{\alpha - 2\theta}\big)},$$

and the limits

$$\lim_{\rho\to\infty} \mu_0(c,\alpha,\theta,\rho) = 0, \qquad \lim_{\rho\to 0} \mu_0(c,\alpha,\theta,\rho) = c,$$
$$\lim_{c\to\infty} \mu_0(c,\alpha,\theta,\rho) = \frac{\sqrt{\alpha - 2\theta}}{\rho\big(\sqrt{\alpha} - \sqrt{\alpha - 2\theta}\big)}, \qquad \lim_{c\to 0} \mu_0(c,\alpha,\theta,\rho) = 0.$$

Example 3. Let $K(x) = A\exp(-a|x|) - B\exp(-b|x|)$, where $A > B > 0$ and $a > b > 0$ are constants, such that

$$\frac{2A}{a} - \frac{2B}{b} = 1, \qquad \frac{B}{b^2} \le \frac{A}{a^2}.$$

For example, the choice $A = 4$, $a = 4$, $B = 1$, $b = 2$ is fine. This is a symmetric Mexican hat kernel function in class (B). The speed index function and the speed of the front satisfy

$$\varphi(\mu) = \frac{Ac\alpha\mu}{ac\mu + c - \mu} - \frac{Bc\alpha\mu}{bc\mu + c - \mu} \to \frac{A\alpha\mu}{1 + a\mu} - \frac{B\alpha\mu}{1 + b\mu} \qquad \text{as } c \to \infty,$$
$$\frac{\alpha - 2\theta}{2\alpha} = \left[\frac{A}{ac\mu_0 + c - \mu_0} - \frac{B}{bc\mu_0 + c - \mu_0}\right] c\mu_0,$$
$$\frac{1}{\mu_0} = \frac{1}{c} + \frac{\alpha}{\alpha - 2\theta}(A - B) - \frac{a+b}{2} + \frac{1}{2}\sqrt{(a-b)^2 + \left[\frac{2\alpha}{\alpha - 2\theta}(A - B)\right]^2 + \frac{4\alpha}{\alpha - 2\theta}(A + B)(b - a)}.$$

This speed is smaller than the speed in Example 1 if the parameters c, α, θ are kept the same in these examples and if $A - B > \frac{\rho}{2}$.

Example 4. Let $K(x) = A|x|\exp(-a|x|) - B|x|\exp(-b|x|)$, where $B > A > 0$ and $b > a > 0$ are constants, such that

$$\frac{2A}{a^2} - \frac{2B}{b^2} = 1.$$

For example, $A = 1$, $a = 1$, $B = 2$, $b = 2$ is fine. This is a symmetric upside down Mexican hat kernel function in class (C). The speed index function and the speed of the front satisfy

$$\varphi(\mu) = \frac{Ac^2\alpha\mu^2}{\big(ac\mu + c - \mu\big)^2} - \frac{Bc^2\alpha\mu^2}{(bc\mu + c - \mu)^2} \to \frac{A\alpha\mu^2}{(1+a\mu)^2} - \frac{B\alpha\mu^2}{(1+b\mu)^2} \qquad \text{as } c \to \infty,$$

$$\frac{\alpha - 2\theta}{2\alpha} = \left[\frac{A}{(ac\mu_0 + c - \mu_0)^2} - \frac{B}{(bc\mu_0 + c - \mu_0)^2}\right] c^2{\mu_0}^2.$$

This speed is larger than the speed in Example 2 if the parameters c, α, θ are kept the same in these examples and if $\rho > a$. In these speed comparisons, we used a simple result: the equation

$$\sum_{k=1}^{n} c_k \exp(\rho_k x) = 0,$$

where $c_k \neq 0$ and $\rho_k \neq \rho_l$ for $k \neq l$, has at most n solutions.

Overall

$$\mu_0(4) > \mu_0(2) > \mu_0(1) > \mu_0(3),$$

where $\mu_0(i)$ represents the speed of the front in example i. To illustrate these estimates, we have to introduce an important concept. Given a kernel function K, define the moment

$$m = \int_{\mathbb{R}} |x| K(x) dx.$$

Intuitively, a nonnegative off-center kernel function has a larger moment than a nonnegative on-center kernel function. An upside down Mexican hat kernel function has a larger moment than a nonnegative kernel function, and a nonnegative kernel function has a larger moment than a Mexican hat kernel function. The speed of the front is not really proportional to the moment. But roughly speaking, the larger the moment is, the bigger the speed is.

3.2. Estimates and limits of the speed. Note that the speed index function

$$\begin{aligned}
&\alpha \int_{-\infty}^{0} \exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right) K(\xi) d\xi \\
&= \frac{\alpha}{2} - \alpha\frac{c-\mu_0}{c\mu_0} \int_{-\infty}^{0} \exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right) \left[\int_{-\infty}^{\xi} K(x) dx\right] d\xi \\
&= \frac{\alpha}{2} - \alpha\frac{c-\mu_0}{c\mu_0} \int_{-\infty}^{0} |\xi| K(\xi) d\xi \\
&+ \alpha\left(\frac{c-\mu_0}{c\mu_0}\right)^2 \int_{-\infty}^{0} \exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right) \left[\int_{-\infty}^{\xi} (\xi - x) K(x) dx\right] d\xi.
\end{aligned}$$

If K is nonnegative on $\mathbb{R}$, then we have the optimal estimates

$$0 < \frac{c\mu_0}{c-\mu_0} \ln\frac{\alpha}{\alpha - 2\theta} < 2\int_{-\infty}^{0} |x| K(x) dx.$$

Let K be a fixed kernel function such as $K(x) = \frac{1}{2}\exp(-|x|)$ or $K(x) = \sqrt{\frac{1}{\pi}}\exp(-x^2)$, or any other function in the three classes. Define a new kernel by rescaling: $\mathcal{K}(\rho, x) = \rho K(\rho x)$, where $\rho > 0$ is a parameter. Then

$$\int_{\mathbb{R}} \mathcal{K}(\rho, x)dx = 1,$$
$$\lim_{\rho\to 0} \mathcal{K}(\rho, x) = 0 \quad \text{uniformly on } \mathbb{R}, \text{ and}$$
$$\lim_{\rho\to\infty} \mathcal{K}(\rho, x) = \begin{cases} \delta(x) & \text{if } K(0) > 0, \\ -\delta(x) & \text{if } K(0) < 0, \\ 0 & \text{if } K(0) = 0, \end{cases}$$

where δ is the Dirac delta impulse function.

Theorem 2. *Suppose that $\rho > 0$ is a constant and that $\mathcal{K}(\rho, x) = \rho K(\rho x)$. Let $\mu_0(c, \alpha, \theta, \rho)$ and $U(c, \mu_0(c, \alpha, \theta, \rho), \cdot)$ denote the speed and the front, respectively, of (13), where K is replaced with $\mathcal{K}$. The speeds $\mu_0 = \mu_0(c, \alpha, \theta, \rho)$ of the fronts satisfy*

$$\mu_0(c, \alpha, \theta, \rho) = \frac{c\mu_0(c, \alpha, \theta, 1)}{\mu_0(c, \alpha, \theta, 1) + \rho[c - \mu_0(c, \alpha, \theta, 1)]}, \tag{27}$$

$$\lim_{\rho\to 0} \mu_0(c, \alpha, \theta, \rho) = c, \qquad \lim_{\rho\to 0} \frac{c - \mu_0(c, \alpha, \theta, \rho)}{\rho} = \frac{c[c - \mu_0(c, \alpha, \theta, 1)]}{\mu_0(c, \alpha, \theta, 1)}, \tag{28}$$

$$\lim_{\rho\to\infty} \mu_0(c, \alpha, \theta, \rho) = 0, \qquad \lim_{\rho\to\infty} \left[\rho\mu_0(c, \alpha, \theta, \rho)\right] = \frac{c\mu_0(c, \alpha, \theta, 1)}{c - \mu_0(c, \alpha, \theta, 1)}, \tag{29}$$

and

$$\lim_{\rho\to 0} U(c, \mu_0(c, \alpha, \theta, \rho), z) = \frac{\alpha}{2} \quad \text{uniformly on } \mathbb{R},$$
$$\lim_{\rho\to\infty} U(c, \mu_0(c, \alpha, \theta, \rho), z) = \begin{cases} \alpha H(z) & \text{for } z \neq 0, \\ \theta & \text{for } z = 0. \end{cases} \tag{30}$$

Remark 4. Let $c = \infty$. Then

$$\nu_0(\alpha, \theta, \rho) = \frac{1}{\rho}\nu_0(\alpha, \theta, 1).$$

Proof. The identity (27) follows from the uniqueness of the wave speed. Note that

$$\begin{aligned} &U(c, \mu_0(c, \alpha, \theta, \rho), z) \\ &= \alpha \int_{-\infty}^{cz/[c+s(z)\mu_0(c,\alpha,\theta,\rho)]} \rho K(\rho x)dx \\ &\quad - \alpha \int_{-\infty}^{z} \exp\left(\frac{x - z}{\mu_0(c, \alpha, \theta, \rho)}\right) \frac{c\rho}{c + s(x)\mu_0(c, \alpha, \theta, \rho)} K\left(\frac{c\rho}{c + s(x)\mu_0(c, \alpha, \theta, \rho)}x\right) dx \\ &= \alpha \int_{-\infty}^{c\rho z/[c+s(z)\mu_0(c,\alpha,\theta,\rho)]} K(x)dx \\ &\quad - \alpha \int_{-\infty}^{\rho z} \exp\left(\frac{x - \rho z}{\rho\mu_0(c, \alpha, \theta, \rho)}\right) \frac{c}{c + s(x)\mu_0(c, \alpha, \theta, \rho)} K\left(\frac{c}{c + s(x)\mu_0(c, \alpha, \theta, \rho)}x\right) dx, \end{aligned}$$

where $s(x/\rho) = s(x)$ since $\rho > 0$. The details of the proof of (28)–(30) are omitted because they are straightforward. ■

3.3. How the speed depends on various parameters. In general, it is not very clear how the speed depends on the model parameters. To study this topic, we have four choices: (i) keeping α, θ, and K fixed and letting c vary; (ii) keeping c, θ, and K fixed and letting α vary; (iii) keeping c, α, and K fixed and letting θ vary; (iv) keeping c, α, and θ fixed and letting K vary. The first three choices will be studied below. These results tell us how the speed changes as the parameters vary. In particular, they allow us to see the asymptotic behaviors as the parameters approach certain numbers or infinity. The last choice has been briefly investigated by using the Examples 1–4. We will have to calculate the partial derivatives and investigate the asymptotic behaviors of the speed μ_0. In particular, we have to compute

$$\frac{\partial \mu_0}{\partial c}, \qquad \frac{\partial \mu_0}{\partial \alpha}, \qquad \frac{\partial \mu_0}{\partial \theta}.$$

Let $\mu_0(c, \alpha, \theta)$ and $\nu_0(\alpha, \theta)$ denote the speeds of the traveling waves corresponding to the cases $0 < c < \infty$ and $c = \infty$, respectively. First, we will prove an identity associating $\mu_0(c, \alpha, \theta)$ to c and $\nu_0(\alpha, \theta)$.

Theorem 3. *Suppose that $0 < 2\theta < \alpha$. Then the speeds of the traveling wave fronts of (13) satisfy*

$$\frac{1}{\mu_0(c, \alpha, \theta)} = \frac{1}{c} + \frac{1}{\nu_0(\alpha, \theta)}, \tag{31}$$

$$0 < \mu_0(c, \alpha, \theta) < c, \qquad 0 < \mu_0(c, \alpha, \theta) < \nu_0(\alpha, \theta), \tag{32}$$

$$\lim_{c\to\infty} \mu_0(c, \alpha, \theta) = \nu_0(\alpha, \theta), \qquad \lim_{c\to 0} \mu_0(c, \alpha, \theta) = 0, \qquad \lim_{c\to 0} \frac{\mu_0(c, \alpha, \theta)}{c} = 1. \tag{33}$$

These results are elegant because the formula does not depend on α, θ, and K directly. A similar identity has been found by Golomb and Ermentrout in [20] for the propagation velocity of pulses in a one-dimensional model of integrate-and-fire neurons that are coupled by excitatory synapses with delay.

Proof. The speeds $\mu_0(c, \alpha, \theta)$ and $\nu_0(\alpha, \theta)$ satisfy the equations

$$\alpha \int_{-\infty}^{0} \exp\left[\frac{c - \mu_0(c, \alpha, \theta)}{c\mu_0(c, \alpha, \theta)}\xi\right] K(\xi)d\xi = \frac{\alpha}{2} - \theta$$

and

$$\alpha \int_{-\infty}^{0} \exp\left[\frac{\xi}{\nu_0(\alpha, \theta)}\right] K(\xi)d\xi = \frac{\alpha}{2} - \theta,$$

respectively. By uniqueness, we obtain the relationship

$$\frac{c - \mu_0(c, \alpha, \theta)}{c\mu_0(c, \alpha, \theta)} = \frac{1}{\nu_0(\alpha, \theta)},$$

that is,

$$\frac{1}{\mu_0(c, \alpha, \theta)} = \frac{1}{c} + \frac{1}{\nu_0(\alpha, \theta)}.$$

The estimates and limits follow immediately from this above formula. The proof is now complete. ■

The goal of next theorem is to investigate various behaviors of the speed when the parameters c, α, θ change. For example, we would like to see whether μ_0 is an increasing function of α, and whether it converges to c as $\alpha \to \infty$. If yes, we want to see how fast the convergence is. This requires us to study the limit $[c - \mu_0(c, \alpha, \theta)]/\alpha$ as $\alpha \to \infty$. The behaviors of the speed influence the behaviors of the stability index function as the parameters change. This may play a role when we study the stability of the wave.

Theorem 4. *Let the parameters satisfy $0 < 2\theta < \alpha$. Then the speed of the traveling wave front is monotonically increasing with respect to c and with respect to α, and it is monotonically decreasing with respect to θ. More precisely,* (I) *when $c = \infty$, we have*

$$\frac{\partial \nu_0}{\partial \alpha}(\alpha, \theta) > 0, \qquad \frac{\partial \nu_0}{\partial \theta}(\alpha, \theta) < 0, \tag{34}$$

$$\lim_{\alpha \to \infty} \nu_0(\alpha, \theta) = \infty, \qquad \lim_{\theta \to 0} \nu_0(\alpha, \theta) = \infty, \tag{35}$$

$$\lim_{\alpha \to \infty} \frac{\nu_0(\alpha, \theta)}{\alpha} = \frac{1}{\theta} \int_{-\infty}^{0} |\xi| K(\xi) d\xi, \tag{36}$$

$$\lim_{\theta \to 0} \big[\theta \nu_0(\alpha, \theta)\big] = \alpha \int_{-\infty}^{0} |\xi| K(\xi) d\xi, \tag{37}$$

$$\lim_{\frac{\theta}{\alpha} \to 0} \nu_0(\alpha, \theta) = \infty, \qquad \lim_{\frac{\theta}{\alpha} \to \frac{1}{2}} \nu_0(\alpha, \theta) = 0, \tag{38}$$

$$\lim_{\frac{\theta}{\alpha} \to 0} \frac{\theta \nu_0(\alpha, \theta)}{\alpha} = \int_{-\infty}^{0} |\xi| K(\xi) d\xi, \tag{39}$$

$$\lim_{\frac{\theta}{\alpha} \to \frac{1}{2}} \frac{2\alpha}{\alpha - 2\theta} \nu_0(\alpha, \theta) = \frac{1}{|K(0)|}, \tag{40}$$

where in the last limit, we assume that $K(0) \neq 0$; and (II) *when $0 < c < \infty$, we have*

$$\frac{\partial \mu_0}{\partial c}(c, \alpha, \theta) > 0, \qquad \frac{\partial \mu_0}{\partial \alpha}(c, \alpha, \theta) > 0, \qquad \frac{\partial \mu_0}{\partial \theta}(c, \alpha, \theta) < 0, \tag{41}$$

$$\lim_{\alpha \to \infty} \mu_0(c, \alpha, \theta) = c, \qquad \lim_{\theta \to 0} \mu_0(c, \alpha, \theta) = c, \qquad \lim_{\frac{\theta}{\alpha} \to \frac{1}{2}} \mu_0(c, \alpha, \theta) = 0, \tag{42}$$

$$\lim_{c \to \infty} \big\{c[\nu_0(\alpha, \theta) - \mu_0(c, \alpha, \theta)]\big\} = \big[\nu_0(\alpha, \theta)\big]^2, \tag{43}$$

$$\lim_{\alpha \to \infty} \big\{\alpha[c - \mu_0(c, \alpha, \theta)]\big\} = \frac{c^2 \theta}{\int_{-\infty}^{0} |\xi| K(\xi) d\xi}, \tag{44}$$

$$\lim_{\theta \to 0} \frac{c - \mu_0(c, \alpha, \theta)}{\theta} = \frac{c^2}{\alpha} \int_{-\infty}^{0} |\xi| K(\xi) d\xi, \tag{45}$$

$$\lim_{\frac{\theta}{\alpha} \to 0} \left\{\frac{\alpha}{\theta}[c - \mu_0(c, \alpha, \theta)]\right\} = c^2 \int_{-\infty}^{0} |\xi| K(\xi) d\xi, \tag{46}$$

$$\lim_{\frac{\theta}{\alpha} \to \frac{1}{2}} \mu_0(c, \alpha, \theta) \frac{2\alpha}{\alpha - 2\theta} = \frac{1}{|K(0)|} \text{ if } K(0) \neq 0. \tag{47}$$

Proof. First of all, differentiating the equation $\frac{1}{\mu_0(c,\alpha,\theta)} = \frac{1}{c} + \frac{1}{\nu_0(\alpha,\theta)}$ with respect to c, and noting that $\nu_0(\alpha,\theta)$ is independent of c, we have

$$-\frac{1}{\big[\mu_0(c,\alpha,\theta)\big]^2}\frac{\partial \mu_0}{\partial c}(c,\alpha,\theta) = -\frac{1}{c^2}.$$

Now we find that

$$\frac{\partial \mu_0}{\partial c}(c,\alpha,\theta) = \frac{\big[\mu_0(c,\alpha,\theta)\big]^2}{c^2} > 0.$$

Secondly, differentiating $\frac{1}{\mu_0(c,\alpha,\theta)} = \frac{1}{c} + \frac{1}{\nu_0(\alpha,\theta)}$ with respect to α, we obtain

$$-\frac{1}{\big[\mu_0(c,\alpha,\theta)\big]^2}\frac{\partial \mu_0}{\partial \alpha}(c,\alpha,\theta) = -\frac{1}{\big[\nu_0(\alpha,\theta)\big]^2}\frac{\partial \nu_0}{\partial \alpha}(\alpha,\theta).$$

That is,

$$\frac{\partial \mu_0}{\partial \alpha}(c,\alpha,\theta) = \frac{\big[\mu_0(c,\alpha,\theta)\big]^2}{\big[\nu_0(\alpha,\theta)\big]^2}\frac{\partial \nu_0}{\partial \alpha}(\alpha,\theta). \tag{48}$$

Now upon differentiating

$$\frac{\alpha - 2\theta}{2\alpha} = \int_{-\infty}^{0} \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi)d\xi$$

with respect to α, we get

$$\frac{\theta}{\alpha^2} = \frac{1}{\big[\nu_0(\alpha,\theta)\big]^2}\frac{\partial \nu_0}{\partial \alpha}(\alpha,\theta)\int_{-\infty}^{0} |\xi| \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi)d\xi.$$

Therefore

$$\frac{\partial \nu_0}{\partial \alpha}(\alpha,\theta) = \frac{\theta}{\alpha^2}\big[\nu_0(\alpha,\theta)\big]^2 \Big/ \int_{-\infty}^{0} |\xi| \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi)d\xi = \frac{\theta}{\alpha}\frac{1}{\varphi'(\nu_0)} > 0, \tag{49}$$

and by (48),

$$\frac{\partial \mu_0}{\partial \alpha}(c,\alpha,\theta) = \frac{\theta}{\alpha^2}\big[\mu_0(c,\alpha,\theta)\big]^2 \Big/ \int_{-\infty}^{0} |\xi| \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi)d\xi > 0. \tag{50}$$

Similarly, we can obtain the following partial derivatives with respect to θ:

$$\frac{\partial \mu_0}{\partial \theta}(c,\alpha,\theta) = \frac{\big[\mu_0(c,\alpha,\theta)\big]^2}{\big[\nu_0(\alpha,\theta)\big]^2}\frac{\partial \nu_0}{\partial \theta}(\alpha,\theta), \tag{51}$$

$$\frac{\partial \nu_0}{\partial \theta}(\alpha,\theta) = -\frac{\big[\nu_0(\alpha,\theta)\big]^2}{\alpha} \Big/ \int_{-\infty}^{0} |\xi| \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi)d\xi < 0, \tag{52}$$

$$\frac{\partial \mu_0}{\partial \theta}(c,\alpha,\theta) = -\frac{\big[\mu_0(c,\alpha,\theta)\big]^2}{\alpha} \Big/ \int_{-\infty}^{0} |\xi| \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi)d\xi < 0. \tag{53}$$

Recall that

$$\int_{-\infty}^{0} \exp\left[\frac{c-\mu_0(c,\alpha,\theta)}{c\mu_0(c,\alpha,\theta)}\xi\right] K(\xi)d\xi = \frac{\alpha-2\theta}{2\alpha} = \beta,$$

$$\frac{\partial\mu_0}{\partial\beta}(c,\alpha,\theta) = \big[\mu_0(c,\alpha,\theta)\big]^2 / \int_{-\infty}^{0} |\xi| \exp\left[\frac{c-\mu_0(c,\alpha,\theta)}{c\mu_0(c,\alpha,\theta)}\xi\right] K(\xi)d\xi = \frac{\alpha}{\varphi'(\mu_0)} > 0,$$

$$\frac{\partial\mu_0}{\partial(\theta/\alpha)}(c,\alpha,\theta) = -\big[\mu_0(c,\alpha,\theta)\big]^2 / \int_{-\infty}^{0} |\xi| \exp\left[\frac{c-\mu_0(c,\alpha,\theta)}{c\mu_0(c,\alpha,\theta)}\xi\right] K(\xi)d\xi = -\frac{\alpha}{\varphi'(\mu_0)} < 0,$$

and similarly

$$\int_{-\infty}^{0} \exp\left[\frac{1}{\nu_0(\alpha,\theta)}\xi\right] K(\xi)d\xi = \frac{\alpha-2\theta}{2\alpha},$$

$$\frac{\partial\nu_0}{\partial\beta}(\alpha,\theta) = \big[\nu_0(\alpha,\theta)\big]^2 / \int_{-\infty}^{0} |\xi| \exp\left[\frac{1}{\nu_0(\alpha,\theta)}\xi\right] K(\xi)d\xi$$

$$= 1/\int_{-\infty}^{0} |x| \exp(x) K(\nu_0(\alpha,\theta)x)dx,$$

where

$$\int_{-\infty}^{0} |\xi| \exp\left[\frac{c-\mu_0(c,\alpha,\theta)}{c\mu_0(c,\alpha,\theta)}\xi\right] K(\xi)d\xi > 0 \quad \text{and} \quad \int_{-\infty}^{0} |\xi| \exp\left[\frac{1}{\nu_0(\alpha,\theta)}\xi\right] K(\xi)d\xi > 0.$$

Now let us look at the limits (35)–(36). As a function of α, $\nu_0(\alpha,\theta)$ is increasing and positive, and hence

$$0 < l := \lim_{\alpha\to\infty} \nu_0(\alpha,\theta) \le \infty.$$

We conclude that $l = \infty$. Suppose that this is not true; i.e., the limit is finite, $0 < l < \infty$. In the speed equation

$$\int_{-\infty}^{0} \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi)d\xi = \frac{1}{2} - \frac{\theta}{\alpha},$$

letting $\alpha \to \infty$, we find that

$$\frac{1}{2} = \int_{-\infty}^{0} K(\xi)d\xi > \int_{-\infty}^{0} \exp\left(\frac{\xi}{l}\right) K(\xi)d\xi = \frac{1}{2}.$$

This is a contradiction. Therefore

$$l = \lim_{\alpha\to\infty} \nu_0(\alpha,\theta) = \infty. \tag{54}$$

Roughly speaking, μ_0 is proportional to α. By L'Hospital's rule, we have

$$L := \lim_{\alpha\to\infty} \frac{\nu_0(\alpha,\theta)}{\alpha} = \lim_{\alpha\to\infty} \frac{\partial\nu_0}{\partial\alpha}(\alpha,\theta)$$

$$= \left[\lim_{\alpha\to\infty} \frac{\nu_0(\alpha,\theta)}{\alpha}\right]^2 \lim_{\alpha\to\infty} \left\{\theta / \int_{-\infty}^{0} |\xi| \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi)d\xi\right\} \qquad \text{[by (49)]}$$

$$= L^2 \left[\theta / \int_{-\infty}^{0} |\xi| K(\xi)d\xi\right] \qquad \text{[by the limit (54)]}.$$

Thus

$$L = \lim_{\alpha\to\infty} \frac{\nu_0(\alpha,\theta)}{\alpha} = \frac{1}{\theta}\int_{-\infty}^{0} |\xi| K(\xi) d\xi. \tag{55}$$

Following this idea, the limits (37)–(39) are straightforward to establish. Let us prove (40). Recall that $\beta = \frac{\alpha-2\theta}{2\alpha}$. Now we have

$$\begin{aligned}
&\lim_{\beta\to 0} \frac{2\alpha}{\alpha-2\theta}\nu_0(\alpha,\theta) = \lim_{\beta\to 0} \frac{\nu_0(\alpha,\theta)}{\beta} \\
&= \lim_{\beta\to 0} \frac{\partial \nu_0(\alpha,\theta)}{\partial\beta} = \lim_{\beta\to 0} \big[\nu_0(\alpha,\theta)\big]^2 \Big/ \int_{-\infty}^{0} |\xi| \exp\left[\frac{\xi}{\nu_0(\alpha,\theta)}\right] K(\xi) d\xi \\
&= \lim_{\beta\to 0} 1 \Big/ \int_{-\infty}^{0} |x| \exp(x) K(\nu_0(\alpha,\theta)x) dx = \frac{1}{|K(0)|}.
\end{aligned}$$

Now let us investigate the limits involving the delay: $0 < c < \infty$. Obviously

$$\lim_{c\to\infty} c\big[\nu_0(\alpha,\theta) - \mu_0(c,\alpha,\theta)\big] = \lim_{c\to\infty} \big[\mu_0(c,\alpha,\theta)\nu_0(\alpha,\theta)\big] = \big[\nu_0(\alpha,\theta)\big]^2.$$

Thus (43) is true. We also have

$$\begin{aligned}
&\lim_{\alpha\to\infty} \mu_0(c,\alpha,\theta) = \lim_{\alpha\to\infty} \frac{c\nu_0(\alpha,\theta)}{c+\nu_0(\alpha,\theta)} = c, \\
&\lim_{\alpha\to\infty} \alpha\big[c - \mu_0(c,\alpha,\theta)\big] = \lim_{\alpha\to\infty} \frac{c\alpha\mu_0(c,\alpha,\theta)}{\nu_0(\alpha,\theta)} = c^2\theta \Big/ \int_{-\infty}^{0} |\xi| K(\xi) d\xi.
\end{aligned}$$

Now (44) is proved. Moreover, we have

$$\begin{aligned}
&\lim_{\theta\to 0} \mu_0(c,\alpha,\theta) = \lim_{\theta\to 0} \frac{c\nu_0(\alpha,\theta)}{c+\nu_0(\alpha,\theta)} = c, \\
&\lim_{\theta\to 0} \frac{c-\mu_0(c,\alpha,\theta)}{\theta} = \lim_{\theta\to 0} \frac{c\mu_0(c,\alpha,\theta)}{\theta\nu_0(\alpha,\theta)} \\
&\qquad = \frac{c^2}{\alpha} \Big/ \int_{-\infty}^{0} |\xi| K(\xi) d\xi.
\end{aligned}$$

Hence (45) is valid. Finally

$$\begin{aligned}
&\lim_{\frac{\theta}{\alpha}\to 0} \mu_0(c,\alpha,\theta) = \lim_{\frac{\theta}{\alpha}\to 0} \frac{c\nu_0(\alpha,\theta)}{c+\nu_0(\alpha,\theta)} = c, \\
&\lim_{\frac{\theta}{\alpha}\to 0} \frac{\alpha}{\theta}\big[c - \mu_0(c,\alpha,\theta)\big] = \lim_{\frac{\theta}{\alpha}\to 0} \frac{c\alpha\mu_0(c,\alpha,\theta)}{\theta\nu_0(\alpha,\theta)} = c^2 \Big/ \int_{-\infty}^{0} |\xi| K(\xi) d\xi.
\end{aligned}$$

The proof is complete. ■

Throughout this section, we find that

$$\text{the moment } m := \int_{-\infty}^{0} |x| K(x) dx$$

is a very important quantity, because many limits are related to it. Roughly speaking, the moment measures how fast the front propagates.

4. Stability index function and stability analysis. The main goal of this section is to build a relationship between the speed index function and the stability index function and use this relationship to establish the stability of the front.

Define the operator

$$\mathcal{L}\psi = -\mu_0\psi_z - \psi,$$

where $\psi \in C^1(\mathbb{R}) \cap W^{1,\infty}(\mathbb{R})$. Later, we will explain why we define the operator in this way. The following definition is different from those of traditional eigenvalue problems.

DEFINITION. *The complex number λ_0 is called an eigenvalue of the linear differential operator $\mathcal{L}$ if there exists a bounded, uniformly continuous function $\psi = \psi(\lambda_0, z)$ such that*

$$\mu_0\psi_z + (\lambda+1)\psi$$

$$= \frac{\alpha}{U_z(c,\mu_0,0)}\frac{c}{c+s(z)\mu_0}K\left(\frac{c}{c+s(z)\mu_0}z\right)\exp\left[-\frac{\lambda|z|}{c+s(z)\mu_0}\right]\psi(\lambda,0). \tag{56}$$

The function ψ is called an eigenfunction of the operator $\mathcal{L}$ corresponding to the eigenvalue λ_0.

This eigenvalue problem is nonlinear in the eigenvalue parameter λ. The nonlinearity arises due to the presence of the spatial temporal delay. Obviously, if there is no delay, then the eigenvalue problem becomes linear, $\nu_0\psi_z + (\lambda+1)\psi = \frac{\alpha}{U_z(\nu_0,0)}K(z)\psi(\lambda,0)$, which has been studied before; see Zhang [40]. Without loss of generality, suppose that $\psi(\lambda,0) \neq 0$ for each λ. Note that the eigenvalue problem cannot be represented as $\mathcal{L}\psi = \lambda\psi$ so we may not be able to apply the general theory established by Evans [14], [15], [16], [17], or Sattinger [35]. In particular, we are not sure whether the linearized stability criterion is still valid. It is an open problem whether the spectral stability, linear stability, and nonlinear stability of the front are equivalent to each other, since this is a nonlinear eigenvalue problem. For linear eigenvalue problems, Sandstede [34] recently proved the equivalence.

To study the stability of the traveling wave of the integral differential equation (13), we are going to compute the essential spectrum and the eigenvalues of the associated linear differential operator $\mathcal{L}$ relative to the Banach space $C^0(\mathbb{R}) \cap L^\infty(\mathbb{R})$. The essential spectrum $\sigma_{\text{essential}}(\mathcal{L})$ is relatively easy to calculate: $\sigma_{\text{essential}}(\mathcal{L}) = \{\lambda \in \mathbb{C}\colon \text{Re}\lambda = -1\}$. We will use the speed index function to construct the stability index function, which is defined in the right half plane $\Omega = \{\lambda \in \mathbb{C}\colon \text{Re}\lambda > -1\}$, to locate all of the eigenvalues of $\mathcal{L}$: $\sigma_{\text{normal}}(\mathcal{L})$.

4.1. The eigenvalue problem. Let $z = x+\mu_0 t$ be the moving coordinate, and set $P(z,t) = u(x,t)$; then $u_t = P_t + \mu_0 P_z$. Now (13) becomes

$$P_t + \mu_0 P_z + P = \alpha\int_{\mathbb{R}} K(z-y)H\left(P\left(y - \frac{\mu_0}{c}|z-y|, t - \frac{1}{c}|z-y|\right) - \theta\right)dy. \tag{57}$$

Recall that $s(x)$ is the sign function, and keep in mind that $0 < \mu_0 < c$. As before, we will make the change of variable $\eta = y - \frac{\mu_0}{c}|z-y|$. Then the nonlocal term becomes

$$\int_{\mathbb{R}} K(z-y)H\left(P\left(y - \frac{\mu_0}{c}|z-y|, t - \frac{1}{c}|z-y|\right) - \theta\right)dy$$
$$= \int_{\mathbb{R}} \frac{c}{c+s(z-\eta)\mu_0}K\left(\frac{c}{c+s(z-\eta)\mu_0}(z-\eta)\right)H\left(P\left(\eta, t - \frac{1}{c+s(z-\eta)\mu_0}|z-\eta|\right) - \theta\right)d\eta.$$

The traveling wave front is a stationary solution of (57), and we can linearize it to get

$$\begin{aligned} &Q_t + \mu_0 Q_z + Q \\ &= \frac{\alpha}{U_z(c,\mu_0,0)} \cdot \frac{c}{c+s(z)\mu_0} \cdot K\left(\frac{c}{c+s(z)\mu_0} z\right) \cdot Q\left(0, t - \frac{|z|}{c+s(z)\mu_0}\right). \end{aligned} \tag{58}$$

Suppose that $Q(z,t) = \exp(\lambda t)\psi(z)$ is a solution of (58), which grows or decays in time depending on whether the real part of λ is positive or negative, respectively, therefore determining the stability of the traveling wave front. Upon substituting such solutions, we obtain the following eigenvalue problem:

$$\mu_0 \psi_z + (\lambda+1)\psi = \frac{\alpha}{U_z(c,\mu_0,0)} \cdot \frac{c}{c+s(z)\mu_0} \cdot K\left(\frac{c}{c+s(z)\mu_0} z\right) \cdot \exp\left[-\frac{\lambda|z|}{c+s(z)\mu_0}\right] \psi(\lambda,0).$$

The aforementioned definition of the differential operator is motivated by this eigenvalue problem.[1] As mentioned before, this problem is very interesting and is quite different from previous eigenvalue problems arising from wave stability analysis, because the eigenvalue parameter λ appears in a nonlinear way.

4.2. The solutions of the eigenvalue problem. One solution of the simple differential equation $\mu_0\psi' + (\lambda+1)\psi = 0$ is $\psi(\lambda,z) = \exp(-\frac{\lambda+1}{\mu_0} z)$ for all complex numbers λ. Define the piecewise constant function

$$a(z) = \frac{\alpha}{\mu_0 U_z(c,\mu_0,0)} \cdot \frac{c}{c+s(z)\mu_0}.$$

Recall that

$$U_z(c,\mu_0,0) = \frac{\alpha - 2\theta}{2\mu_0}.$$

Hence

$$a(z) = \frac{2c\alpha}{(\alpha-2\theta)[c+s(z)\mu_0]}.$$

The general solutions of the eigenvalue problem are given by

$$\begin{aligned} \psi(\lambda,z) &= \mathcal{C}(\lambda) \exp\left(-\frac{\lambda+1}{\mu_0} z\right) \\ &+ \psi(\lambda,0) \int_{-\infty}^{z} a(x) K\left(\frac{c}{c+s(x)\mu_0} x\right) \exp\left[\frac{\lambda+1}{\mu_0}(x-z)\right] \exp\left[-\frac{\lambda|x|}{c+s(x)\mu_0}\right] dx, \end{aligned}$$

where $\mathcal{C}(\lambda)$ is a complex constant to be specified later. Due to the condition $|K(x)| \le C\exp(-\rho|x|)$ on $\mathbb{R}$, it can be demonstrated straightforwardly that, for each fixed λ with

[1] We could not define the operator in this way:

$$\mathcal{L}\psi = -\mu_0\psi_z - \psi + \frac{\alpha}{U_z(c,\mu_0,0)} \cdot \frac{c}{c+s(z)\mu_0} \cdot K\left(\frac{c}{c+s(z)\mu_0} z\right) \cdot \exp\left[-\frac{\lambda|z|}{c+s(z)\mu_0}\right] \psi(\lambda,0).$$

Note that the eigenvalue parameter λ is involved in this definition.

$\mathrm{Re}\lambda > -1$, the solution ψ is bounded on $\mathbb{R}$ if and only if $\mathcal{C}(\lambda) = 0$. Setting $z = 0$ in this solution, we have

$$\mathcal{C}(\lambda) = \left\{1 - \int_{-\infty}^{0} a(x) K\left(\frac{c}{c + s(x)\mu_0} x\right) \exp\left(\frac{\lambda+1}{\mu_0} x\right) \exp\left[-\frac{\lambda|x|}{c + s(x)\mu_0}\right] dx\right\} \psi(\lambda, 0). \tag{59}$$

If $\psi(\lambda, 0) = 0$, then the eigenvalue problem has only one solution—the trivial solution $\psi(\lambda, \cdot) \equiv 0$ on $\mathbb{R}$. It turns out that for nonzero solutions there must hold $\psi(\lambda, 0) \neq 0$. It is easy to show that for each fixed λ the number $\mathcal{C}(\lambda) = 0$ if and only if

$$1 - \int_{-\infty}^{0} a(x) K\left(\frac{c}{c + s(x)\mu_0} x\right) \exp\left(\frac{\lambda+1}{\mu_0} x\right) \exp\left[-\frac{\lambda|x|}{c + s(x)\mu_0}\right] dx = 0.$$

4.3. The speed index function vs the stability index function. Note that

$$\begin{aligned}
&\int_{-\infty}^{0} a(x) K\left(\frac{c}{c + s(x)\mu_0} x\right) \exp\left(\frac{\lambda+1}{\mu_0} x\right) \exp\left[-\frac{\lambda|x|}{c + s(x)\mu_0}\right] dx \\
&= \frac{2\alpha}{\alpha - 2\theta} \int_{-\infty}^{0} \exp\left[(\lambda+1)\frac{c - \mu_0}{c\mu_0}\xi\right] \exp\left(\frac{\lambda}{c}\xi\right) K(\xi) d\xi \\
&= \frac{2\alpha}{\alpha - 2\theta} \int_{-\infty}^{0} \exp\left[\left(\frac{\lambda+1}{\mu_0} - \frac{1}{c}\right)\xi\right] K(\xi) d\xi \\
&= \frac{1}{\varphi(\mu_0)} \varphi\left(\frac{\mu_0}{\lambda+1}\right).
\end{aligned}$$

We construct the stability index function for the front, with the open domain $\Omega = \{\lambda \in \mathbb{C}: \mathrm{Re}\lambda > -1\}$, by

$$\mathcal{E}(\lambda) = 1 - \frac{1}{\varphi(\mu_0)} \varphi\left(\frac{\mu_0}{\lambda+1}\right), \tag{60}$$

where we recall that the speed index function $\varphi(\mu)$ is defined by

$$\varphi(\mu) = \alpha \int_{-\infty}^{0} \exp\left(\frac{c - \mu}{c\mu}\xi\right) K(\xi) d\xi.$$

It would be nice to see the behaviors of the stability index function as the synaptic coupling and the parameters vary. We may gain important insights on the eigenvalues of the operator (thus stability of the front). Obviously, we have the following limits: (I)

$$\begin{aligned}
\lim_{c\to\infty} \mathcal{E}(\lambda) &= 1 - \frac{\int_{-\infty}^{0} \exp\left(\frac{\lambda+1}{\nu_0}\xi\right) K(\xi) d\xi}{\int_{-\infty}^{0} \exp\left(\frac{\xi}{\nu_0}\right) K(\xi) d\xi} \\
&= 1 - \frac{1}{\varphi(\nu_0)} \varphi\left(\frac{\nu_0}{\lambda+1}\right),
\end{aligned}$$

where $\nu_0 = \nu_0(\alpha, \theta)$ is the speed of the front of (13) when there is no delay. Recall that for $0 < c < \infty$ we have

$$\lim_{\rho\to 0} \mu_0(c,\alpha,\theta,\rho) = c, \qquad \lim_{\rho\to 0} \frac{c-\mu_0(c,\alpha,\theta,\rho)}{\rho} = \frac{c[c-\mu_0(c,\alpha,\theta,1)]}{\mu_0(c,\alpha,\theta,1)},$$
$$\lim_{\rho\to 0} \frac{1}{\rho}\frac{c-\mu_0(c,\alpha,\theta,\rho)}{c\mu_0(c,\alpha,\theta,\rho)} = \frac{c-\mu_0(c,\alpha,\theta,1)}{c\mu_0(c,\alpha,\theta,1)},$$
$$\lim_{\rho\to \infty} \mu_0(c,\alpha,\theta,\rho) = 0, \qquad \lim_{\rho\to \infty} \rho\mu_0(c,\alpha,\theta,\rho) = \frac{c\mu_0(c,\alpha,\theta,1)}{c-\mu_0(\alpha,\theta,1)},$$
$$\lim_{\rho\to \infty} \frac{1}{\rho}\frac{c-\mu_0(c,\alpha,\theta,\rho)}{c\mu_0(c,\alpha,\theta,\rho)} = \frac{c-\mu_0(c,\alpha,\theta,1)}{c\mu_0(\alpha,\theta,1)}.$$

Note that if $\rho > 0$ and $K(x)$ is replaced with $\rho K(\rho x)$, then the stability index function is also a function of ρ. We have

$$\begin{aligned}
\lim_{\rho\to 0} \mathcal{E}(\lambda,\rho) &= 1 - \lim_{\rho\to 0} \frac{\int_{-\infty}^0 \exp\left[\left(\frac{\lambda+1}{\mu_0(c,\alpha,\theta,\rho)} - \frac{1}{c}\right)\xi\right] \rho K(\rho\xi)d\xi}{\int_{-\infty}^0 \exp\left[\left(\frac{1}{\mu_0(c,\alpha,\theta,\rho)} - \frac{1}{c}\right)\xi\right] \rho K(\rho\xi)d\xi} \\
&= 1 - \lim_{\rho\to 0} \frac{\int_{-\infty}^0 \exp\left[\frac{1}{\rho}\left(\frac{\lambda+1}{\mu_0(c,\alpha,\theta,\rho)} - \frac{1}{c}\right)x\right] K(x)dx}{\int_{-\infty}^0 \exp\left[\frac{1}{\rho}\left(\frac{1}{\mu_0(c,\alpha,\theta,\rho)} - \frac{1}{c}\right)x\right] K(x)dx} \\
&= 1 \qquad \text{for all } \lambda \neq 0 \text{ with } \mathrm{Re}\lambda \geq 0, \text{ if } 0 < c < \infty, \\
&= 1 - \frac{\int_{-\infty}^0 \exp\left[\frac{\lambda+1}{\nu_0(\alpha,\theta,1)}x\right] K(x)dx}{\int_{-\infty}^0 \exp\left[\frac{1}{\nu_0(\alpha,\theta,1)}x\right] K(x)dx} \qquad \text{if } c = \infty,
\end{aligned}$$

for all nonzero complex numbers with $\mathrm{Re}\lambda \geq 0$ except for $\lambda = 0$; similarly

$$\begin{aligned}
\lim_{\rho\to \infty} \mathcal{E}(\lambda,\rho) &= 1 - \lim_{\rho\to \infty} \frac{\int_{-\infty}^0 \exp\left[\left(\frac{\lambda+1}{\mu_0(c,\alpha,\theta,\rho)} - \frac{1}{c}\right)\xi\right] \rho K(\rho\xi)d\xi}{\int_{-\infty}^0 \exp\left[\left(\frac{1}{\mu_0(c,\alpha,\theta,\rho)} - \frac{1}{c}\right)\xi\right] \rho K(\rho\xi)d\xi} \\
&= 1 - \frac{\int_{-\infty}^0 \exp\left[(\lambda+1)\frac{c-\mu_0(c,\alpha,\theta,1)}{c\mu_0(c,\alpha,\theta,1)}x\right] K(x)dx}{\int_{-\infty}^0 \exp\left[\frac{c-\mu_0(c,\alpha,\theta,1)}{c\mu_0(c,\alpha,\theta,1)}x\right] K(x)dx}
\end{aligned}$$

for all $\lambda \in \Omega$. Recall that

$$\lim_{\frac{\theta}{\alpha}\to 0} \mu_0(c,\alpha,\theta) = c, \qquad \lim_{\frac{\theta}{\alpha}\to \frac{1}{2}} \mu_0(c,\alpha,\theta) = 0.$$

Therefore

$$\lim_{\frac{\theta}{\alpha}\to 0} \mathcal{E}(\lambda) = \begin{cases} 1 - 2\int_{-\infty}^0 \exp\left(\frac{\lambda}{c}\xi\right) K(\xi)d\xi & \text{if} \quad c < \infty, \\ 0 & \text{if} \quad c = \infty, \end{cases}$$

and

$$\lim_{\frac{\theta}{\alpha}\to\frac{1}{2}} \mathcal{E}(\lambda) = 1 - \frac{1}{\lambda+1} = \frac{\lambda}{\lambda+1}$$

if $K(0) \neq 0$.

Certainly the limits of the stability index function help us understand much better the dynamics of the waves.

For the traveling wave front, we also have

$$\mathcal{E}'(\lambda) = \frac{\mu_0}{\varphi(\mu_0)} \frac{1}{(\lambda+1)^2} \varphi'\left(\frac{\mu_0}{\lambda+1}\right). \tag{61}$$

Clearly, $\mathcal{E}'(\lambda) > 0$ for all real numbers $\lambda > -1$, provided that the kernel function is nonnegative (pure excitation) or Mexican hat (lateral inhibition). In particular for the traveling wave front, we have

$$\mathcal{E}'(0) = \mu_0 \frac{\varphi'(\mu_0)}{\varphi(\mu_0)} > 0. \tag{62}$$

The stability index function constructed in this paper is different from the stability index function in our former work; see Zhang [40], [41].

4.4. Examples of the stability index function and eigenvalues. Again concrete examples always help us understand the stability index function better than the abstract theory.

Example 5. Let $K(x) = \frac{\rho}{2}\exp(-\rho|x|)$, where $\rho > 0$ is a constant. Then the stability index function for the front is

$$\mathcal{E}(\lambda) = \frac{\lambda}{\lambda + 1 + \rho\mu_0 - \mu_0/c} \to \frac{2\theta\lambda}{\alpha + 2\theta\lambda},$$

as $c \to \infty$. The only eigenvalue of $\mathcal{L}$ in the right half plane is $\lambda = 0$.

Example 6. Let $K(x) = \frac{1}{2}\rho^2|x|\exp(-\rho|x|)$, where $\rho > 0$ is a constant. The stability index function for the front

$$\begin{aligned}\mathcal{E}(\lambda) &= 1 - \frac{(1+\rho\mu_0-\mu_0/c)^2}{(\lambda+1+\rho\mu_0-\mu_0/c)^2} = \frac{\lambda + 2(1+\rho\mu_0-\mu_0/c)}{(\lambda+1+\rho\mu_0-\mu_0/c)^2}\lambda \\ &\to \frac{\lambda+2+2\rho\nu_0}{(\lambda+1+\rho\nu_0)^2}\lambda\end{aligned}$$

as $c \to \infty$. There is only one solution to $\mathcal{E}(\lambda) = 0$, that is, $\lambda = 0$.

Example 7. Let $K(x) = A\exp(-a|x|) - B\exp(-b|x|)$, where $A > B > 0$ and $a > b > 0$ are constants, such that

$$\frac{2A}{a} - \frac{2B}{b} = 1, \qquad \frac{B}{b^2} \leq \frac{A}{a^2}.$$

The stability index function for the front is

$$\mathcal{E}(\lambda) = 1 - \frac{\left[\frac{A}{\lambda+1+a\mu_0-\mu_0/c} - \frac{B}{\lambda+1+b\mu_0-\mu_0/c}\right]}{\left[\frac{A}{1+a\mu_0-\mu_0/c} - \frac{B}{1+b\mu_0-\mu_0/c}\right]}$$
$$\to 1 - \frac{\left[\frac{A}{\lambda+1+a\nu_0} - \frac{B}{\lambda+1+b\nu_0}\right]}{\left[\frac{A}{1+a\nu_0} - \frac{B}{1+b\nu_0}\right]}$$

as $c \to \infty$. There exists only one eigenvalue to $\mathcal{L}$: $\lambda = 0$.

Example 8. Let $K(x) = A|x|\exp(-a|x|) - B|x|\exp(-b|x|)$, where $B > A > 0$ and $b > a > 0$ are constants, such that

$$\frac{2A}{a^2} - \frac{2B}{b^2} = 1.$$

The stability index function for the front

$$\mathcal{E}(\lambda) = 1 - \frac{\left[\frac{A}{(\lambda+1+a\mu_0-\mu_0/c)^2} - \frac{B}{(\lambda+1+b\mu_0-\mu_0/c)^2}\right]}{\left[\frac{A}{(1+a\mu_0-\mu_0/c)^2} - \frac{B}{(1+b\mu_0-\mu_0/c)^2}\right]}$$
$$\to 1 - \frac{\left[\frac{A}{(\lambda+1+a\nu_0)^2} - \frac{B}{(\lambda+1+b\nu_0)^2}\right]}{\left[\frac{A}{(1+a\nu_0)^2} - \frac{B}{(1+b\nu_0)^2}\right]}$$

as $c \to \infty$. There is only one solution to $\mathcal{E}(\lambda) = 0$: $\lambda = 0$.

By very careful examination of these stability index functions, we find that the only eigenvalue of the operator $\mathcal{L}$ in the right half plane $\{\lambda \in \mathbb{C}\colon \mathrm{Re}\lambda \geq 0\}$ is the neutral eigenvalue $\lambda = 0$, which is simple. Therefore each of these traveling wave fronts *may* be stable. See Figures 3 and 4.

4.5. Properties of the stability index function and eigenvalues.

Theorem 5. *The following results are true relative to the Banach space* $C^0(\mathbb{R}) \cap L^\infty(\mathbb{R})$:

(I) *The stability index function* $\mathcal{E}(\lambda)$ *is complex analytic in* λ, *and it is real-valued if* λ *is real.*

(II) $\mathcal{E}(0) = 0$. *The complex number* λ_0 *is an eigenvalue of the linear differential operator* $\mathcal{L}$ *if and only if* $\mathcal{E}(\lambda_0) = 0$. *(This is the one-to-one correspondence.)*

(III) *The algebraic multiplicity of any eigenvalue* λ_0 *of* $\mathcal{L}$ *is equal to the order of* λ_0 *as a zero of the stability index function, which is equal to one.*

(IV) *The stability index function enjoys the limit*

$$\lim_{|\lambda|\to\infty} \mathcal{E}(\lambda) = 1$$

in the right half plane $\{\lambda \in \mathbb{C} : \mathrm{Re}\lambda \geq 0\}$.

(V) *There exists no nonzero solution to* $\mathcal{E}(\lambda) = 0$ *inside* $\{\lambda \in \mathbb{C} : \mathrm{Re}\lambda \geq 0\}$.

Proof. We only sketch the proof of (II) and (V). See Zhang [40] and [41] for the proofs of other similar results. The complex number λ_0 is an eigenvalue of the linear differential

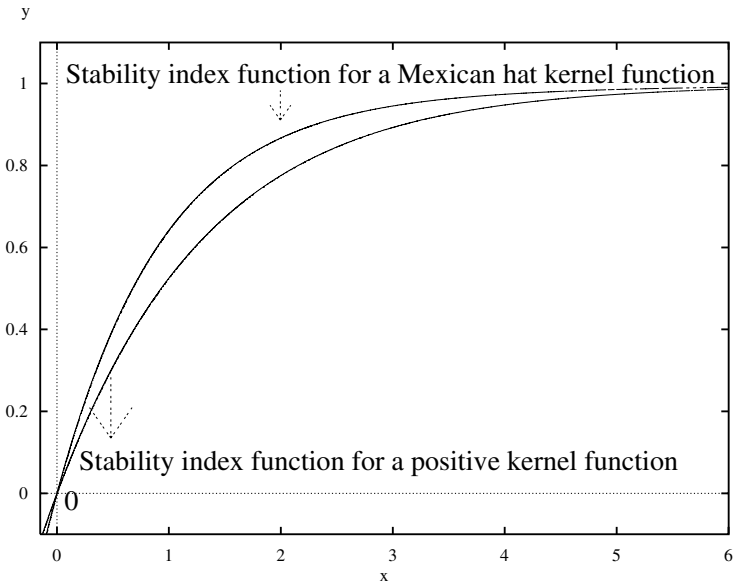


Figure 3. *Comparison of stability index functions (between a positive kernel and a Mexican hat kernel).*

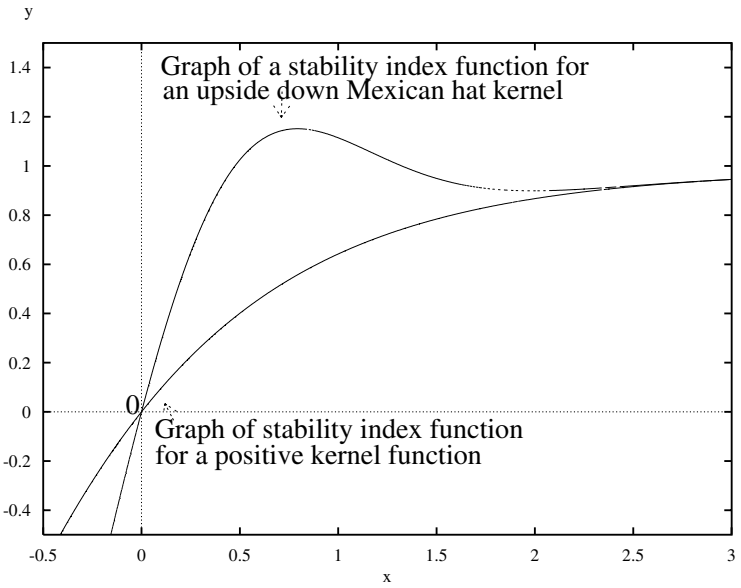


Figure 4. *Comparison of stability index functions (between a positive kernel and an upside down Mexican hat kernel).*

operator $\mathcal{L}$ if and only if there exists a bounded and uniformly continuous function ψ defined on the entire real line $\mathbb{R}$ such that

$$\mu_0\psi_z + (\lambda_0+1)\psi = \frac{\alpha}{U_z(c,\mu_0,0)}\frac{c}{c+s(z)\mu_0}\exp\left[-\frac{\lambda|z|}{c+s(z)\mu_0}\right]K\left(\frac{c}{c+s(z)\mu_0}z\right)\psi(\lambda_0,0).$$

By the assumptions on the kernel function and also by using L'Hospital's rule, we can claim that the solution of the eigenvalue problem is bounded on the real line if and only if $\mathcal{C}(\lambda_0) = 0$. By the above analysis and (59), we obtain the relationship $\mathcal{C}(\lambda_0) = \mathcal{E}(\lambda_0)\psi(\lambda_0, 0)$. Hence λ_0 is an eigenvalue of $\mathcal{L}$ if and only if $\mathcal{E}(\lambda_0) = 0$, since $\psi(\lambda_0, 0) \neq 0$. Due to translation invariance of the traveling wave front, $\lambda = 0$ is an eigenvalue of $\mathcal{L}$, so $\mathcal{E}(0) = 0$. The proof of (II) is finished.

Suppose that the nonzero complex number λ satisfies $\mathrm{Re}\lambda \geq 0$. Then for any nonnegative kernel function with $\int_{\mathbb{R}} K(x)dx = 1$, there holds the estimate

$$\left| \int_{-\infty}^{0} \exp\left(\frac{\lambda + 1}{\mu_0} x\right) K(x)dx \right| < \int_{-\infty}^{0} \exp\left(\frac{x}{\mu_0}\right) K(x)dx. \tag{63}$$

The estimate is also true for Mexican hat kernel functions and for upside down Mexican hat kernel functions as long as

$$\int_{-\infty}^{0} |x| K(x)dx \geq 0. \tag{64}$$

The proof of (V) is finished immediately. ■

4.6. Stability of the traveling wave front.

Assumption—The linearized stability criterion. *If* $\max\{\mathrm{Re}\lambda : \lambda \in \sigma(\mathcal{L}),\ \lambda \neq 0\} \leq -c_0$ *and* $\lambda = 0$ *is algebraically simple, where* $\sigma(\mathcal{L})$ *denotes the spectrum of* $\mathcal{L}$ *and* $c_0 > 0$ *is a constant, then the traveling wave of the integral differential equation* (13) *is exponentially stable.*

Obviously this assumption is true for $c = \infty$; see Sandstede [34]. Most possibly, it is also true for $0 < c < \infty$ even if we cannot represent the eigenvalue problem as $\mathcal{L}\psi = \lambda\psi$.

Intuitively, other than the traveling wave itself, there exists no other biological mechanism to generate a second eigenvalue of $\mathcal{L}$.

Theorem 6. *Let* $0 < 2\theta < \alpha$ *and* $0 < c \leq \infty$. *Then the traveling wave front of* (13) *is exponentially stable.*

Proof. Let the synaptic coupling be of a pure excitation or lateral inhibition-type kernel. By virtue of similar analysis to the existence and uniqueness of the speed μ_0, and by using (61), we can show that $\mathcal{E}'(\lambda) > 0$ for all real numbers $\lambda > -1$. Let the synaptic coupling be of lateral excitation. Recall that there exists a lower bound $\mu_* > 0$ for the speed such that $\varphi(\mu_*) = 0$ and $\mu_* < \mu_0 < c$. Then there exists a unique positive number $\lambda_* = \frac{\mu_0}{\mu_*} - 1$ such that $\mathcal{E}'(\lambda) > 0$ for $-1 < \lambda < \lambda_*$, $\mathcal{E}'(\lambda_*) = 0$, $\mathcal{E}'(\lambda) < 0$ for $\lambda > \lambda_*$. Note that

$$\lim_{|\lambda| \to \infty} \mathcal{E}(\lambda) = 1. \tag{65}$$

Therefore $\mathcal{E}(\lambda) > 0$ for all real numbers $0 < \lambda < \infty$. By using Theorem 5, there exists no nonzero eigenvalue in the right half plane $\{\lambda \in \mathbb{C}\colon \mathrm{Re}\lambda \geq 0\}$, and the neutral eigenvalue $\lambda = 0$ is simple. In particular, there holds the important relationship

$$\mathcal{E}'(0) = \mu_0 \frac{\varphi'(\mu_0)}{\varphi(\mu_0)} > 0. \tag{66}$$

By the linearized stability criterion, we claim that the front is exponentially stable. ■

5. Concluding remarks on neuronal waves. We proved the existence and stability of the traveling wave front of (13) for all kernel functions in classes (A), (B), (C) and for all $c \in (0, \infty]$. The analysis of the model provides valuable information regarding how the speed of the traveling wave, which is relatively straightforward to measure experimentally, depends on various features of the underlying network. Keeping the parameters fixed, we are able to compare the speeds of the fronts corresponding to different synaptic couplings. Our analysis and results on the speed, the speed index function, and the stability index function have not been obtained before. The solid rigorous mathematical analysis is needed because we can build up important properties for the speed index function and the stability index function. It is based on these properties that we may find interesting applications to applied mathematics and computational neuroscience. The analysis for the existence of the front together with the speed is quite different from one class of kernel functions to another. Our results generalize those of Coombes and Owen, Pinto and Ermentrout; see [8], [9], [29], [30]. The main goal of this section is to present remarks on model equations and their wave solutions closely related to (1). For more general nonlinear models, we may not be able to find the index functions. Nevertheless, by using similar ideas, we can give nice estimates about the speed. For example, for the front of the integral differential equation

$$u_t + u(u-1)(Au-1) = \alpha \int_{\mathbb{R}} K(x-y)H(u(y,t)-\theta)dy, \tag{67}$$

where $A > 1$ is a constant, we have the estimates

$$\frac{\alpha}{2} - \theta \le \varphi(\mu_0), \qquad \varphi(\mu_0) \le \frac{\alpha}{2} - \theta(\theta-1)(A\theta-1), \tag{68}$$

where $\varphi(\mu)$ is the speed index function defined in (23). The speed index function and stability index function may play very important roles in rigorous mathematical analysis of traveling waves of nonlinear singularly perturbed systems of integral differential equations, such as the spike frequency adaptation model

$$u_t + u + w = \alpha \int_{\mathbb{R}} K(x-y)H\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy, \tag{69}$$

$$w_t = \varepsilon(u - \gamma w). \tag{70}$$

Remark 5. For lateral inhibition (modeled by a Mexican hat kernel function), the condition $\int_{-\infty}^{0} |x|K(x)dx \ge 0$ is sufficient to guarantee the existence and uniqueness of the speed μ_0 of the front, but we are not sure if it is also necessary. Suppose that $\int_{\mathbb{R}} K(x)dx = 1$ and $\int_{-\infty}^{0} |x|K(x)dx < 0$ for a Mexican hat kernel function. Then interesting phenomena (e.g., bifurcations of steady states into stable and unstable waves) may occur. See Coombes and Owen [8] and Coombes [10].

5.1. The case $\theta < \alpha < 2\theta$. Mathematically, we can also study the existence and stability of the traveling wave for the case $\theta < \alpha < 2\theta$. The analysis would be very similar to that of the case $0 < 2\theta < \alpha$. However, we would like to offer an alternate approach.

Theorem 7. *Suppose that the parameters satisfy $0 < \theta < \alpha$. Let $u = u(x,t)$ be a solution of the integral equation*

$$u(x,t) = \alpha \int_{-\infty}^{t} \eta(t-s) \left[\int_{\mathbb{R}} K(x-y)H\left(u\left(y, s - \frac{1}{c}|x-y|\right) - \theta\right) dy + \mathcal{I}(x,s) \right] ds \tag{71}$$

or

$$u_t + u = \alpha \int_{\mathbb{R}} K(x-y) H\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy + \mathcal{I}(x,t), \tag{72}$$

respectively, where

$$\int_0^\infty \eta(t)dt = 1 \qquad \textit{and} \qquad \int_{\mathbb{R}} K(x)dx = 1. \tag{73}$$

See [7], [8], *and* [29], [30]. *Define* $\widetilde{u} = \widetilde{u}(x,t)$ *by*

$$\widetilde{u}(x,t) - \theta = \frac{\theta}{\alpha - \theta}[\theta - u(x,t)]. \tag{74}$$

Then $\widetilde{u}$ *is a solution of*

(75)

$$\widetilde{u}(x,t) = \frac{\alpha\theta}{\alpha-\theta} \int_{-\infty}^t \eta(t-s) \left[\int_{\mathbb{R}} K(x-y) H\left(\widetilde{u}\left(y, s - \frac{1}{c}|x-y|\right) - \theta\right) dy - \frac{\theta}{\alpha-\theta}\mathcal{I}(x,s)\right] ds$$

or

$$\widetilde{u}_t + \widetilde{u} = \frac{\alpha\theta}{\alpha-\theta} \int_{\mathbb{R}} K(x-y) H\left(\widetilde{u}\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy - \frac{\theta}{\alpha-\theta}\mathcal{I}(x,t), \tag{76}$$

respectively.

Proof. For any real number x, there holds $H(x) + H(-x) = 1$. Thus

$$H\left(u\left(y, s - \frac{1}{c}|x-y|\right) - \theta\right) + H\left(\widetilde{u}\left(y, s - \frac{1}{c}|x-y|\right) - \theta\right) = 1.$$

Note that

$$\int_0^\infty \eta(t)dt = 1 \qquad \text{and} \qquad \int_{\mathbb{R}} K(x)dx = 1.$$

The rest of the proof is omitted because it is straightforward. ∎

Remark 6. The two solutions u and $\widetilde{u}$ are complementary in the sense that excited intervals of u correspond to nonexcited intervals of $\widetilde{u}$ and vice versa. Furthermore, let $\widetilde{\alpha} = \frac{\alpha\theta}{\alpha-\theta}$. If $0 < 2\theta < \alpha$ and U is the wave front with speed μ_0 found in Theorem 1, then $\theta < \widetilde{\alpha} < 2\theta$ and $\widetilde{U}$ is the wave front of (13) with the same speed μ_0. If $2\theta = \alpha$, then the complementary solution $\widetilde{u}$ coincides with the solution u, because $2\theta = \alpha = \widetilde{\alpha}$.

5.2. A similar model with a parameter. Consider the scalar integral differential equation with a parameter W:

$$u_t + u + W = \alpha \int_{\mathbb{R}} K(x-y) H\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy. \tag{77}$$

Theorem 8. *Suppose that $W = \alpha - 2\theta > 0$. Then there exists a unique stable traveling wave solution to (77) with the same speed μ_0 as the front of (13).*

Proof. Suppose that $0 < 2\theta < \alpha$. If there exists a traveling wave with the speed μ_0, then we must have

$$\mu_0 U' + U + W = \alpha \int_{\mathbb{R}} K(z-y) H\left(U\left(y - \frac{\mu_0}{c}|z-y|\right) - \theta\right) dy,$$

where $z = x + \mu_0 t$. We are seeking a monotone (decreasing) traveling wave solution so that $U(c, \mu_0, 0) = \theta$, $U_z(c, \mu_0, 0) < 0$, $U > \theta$ on $(-\infty, 0)$, and $U < \theta$ on $(0, \infty)$. Additionally,

$$\lim_{z\to-\infty} U(c, \mu_0, z) = \alpha, \qquad \lim_{z\to\infty} U(c, \mu_0, z) = 0, \qquad \lim_{z\to\pm\infty} U_z(c, \mu_0, z) = 0.$$

As before, by making the change of variables $\eta = y - \frac{\mu_0}{c}|z-y|$, the traveling wave equation becomes

$$\mu_0 U' + U + W = \alpha \int_{cz/(c+s(z)\mu_0)}^{\infty} K(x)dx.$$

Now using the integrating factor idea to solve this above equation, we find the solution

$$\begin{aligned} U(c, \mu_0, z) &= \alpha \int_{cz/(c+s(z)\mu_0)}^{\infty} K(x)dx \\ &+ \alpha \int_{-\infty}^{z} \exp\left(\frac{x-z}{\mu_0}\right) \frac{c}{c+s(x)\mu_0} K\left(\frac{cx}{c+s(x)\mu_0}\right) dx - W. \end{aligned}$$

Setting $z = 0$ yields

$$\begin{aligned} U(c, \mu_0, 0) &= \alpha \int_0^{\infty} K(x)dx + \alpha \int_{-\infty}^{0} \exp\left(\frac{x}{\mu_0}\right) \frac{c}{c+s(x)\mu_0} K\left(\frac{cx}{c+s(x)\mu_0}\right) dx - W \\ &= \frac{\alpha}{2} + \alpha \int_{-\infty}^{0} \exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right) K(\xi)d\xi - W. \end{aligned} \tag{78}$$

By comparing this with $\alpha \int_{-\infty}^{0} \exp(\frac{c-\mu_0}{c\mu_0}\xi) K(\xi)d\xi = \frac{\alpha}{2} - \theta$, we find that only when $W = \alpha - 2\theta$ does there hold $U(c, \mu_0, 0) = \theta$. The stability analysis is very similar to that of the front and is omitted. This result on the traveling wave back regarding the speed and crossing is the same as that in Zhang [40]. ∎

5.3. Singularly perturbed system of integral differential equations. Consider the singularly perturbed system of integral differential equations

$$u_t + f(u) + w = \alpha \int_{\mathbb{R}} K(x-y) H\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy, \tag{79}$$

$$w_t = \varepsilon(g(u) - \gamma w), \tag{80}$$

where f and g are smooth functions of u and w. Typical examples are $(f(u), g(u)) = (u, u)$ and $(f(u), g(u)) = (u(u-1)(Au-1), \tanh u)$. For the linear choice, a perturbation argument is not absolutely needed to find traveling pulses.

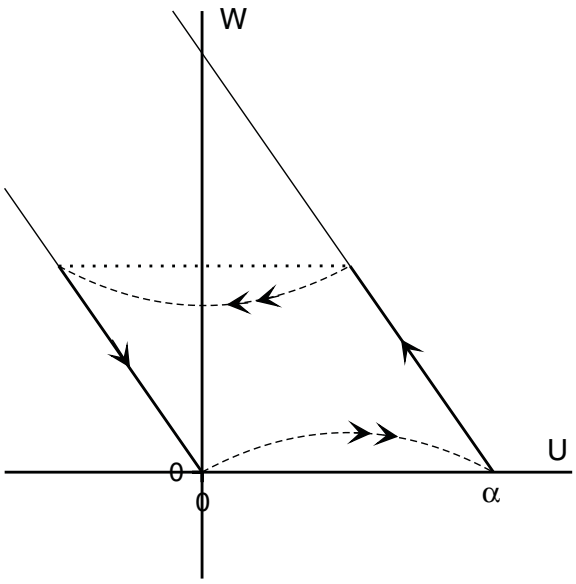


Figure 5. *A singular homoclinic orbit of system* (69)–(70).

By the construction of a singular homoclinic orbit S_1 (see Figure 5), under the assumption that $0 < 2\theta < \alpha$, $0 < \alpha\gamma < (1+\gamma)\theta$, and $0 < \varepsilon \ll 1$, we can prove the existence, uniqueness, and stability of a fast homoclinic orbit, whose speed is very close to that of the front. Due to the existence and uniqueness of a steady-state solution, we can also prove the existence and uniqueness of a slow homoclinic orbit, which is unstable.

By the construction of a singular heteroclinic orbit S_2 (see Figure 6), under the assumption that $0 < 2\theta < \alpha$, $0 < (1+\gamma)\theta < \alpha\gamma$, and $0 < \varepsilon \ll 1$, we can demonstrate the existence, uniqueness, and stability of a unique heteroclinic orbit, whose speed is also very close to that of the front.

The speed index function and the stability index function play a very important role in the mathematical analysis of these results. The details of the proof of these results are too long and are omitted. See Coombes and Owen [8] and Coombes [10].

5.4. Sigmoid firing rate functions. Additional to the Heaviside step gain function, a typically reasonable gain function corresponding to experimental data is given by

$$\mathcal{H}(\beta,\theta,u) = \frac{\exp\left[\beta(u-\theta)\right]}{1+\exp\left[\beta(u-\theta)\right]}, \tag{81}$$

where θ represents a threshold and β stands for a gain or steepness parameter. However, $\mathcal{H}$ is not derived from a biophysical model. The firing rate function $\mathcal{H}$ is monotonically increasing and nonlinear, saturating to a constant for sufficiently large u. Note that $\mathcal{H}$ is a solution of the nonlinear differential equation $\mathcal{H}' = \beta\mathcal{H}(1-\mathcal{H})$ subject to the initial condition $\mathcal{H}(\theta) = \frac{1}{2}$. The existence and uniqueness of a traveling wave front together with a positive speed of (2), where H is replaced with $\mathcal{H}$ and $\mathcal{I} = 0$, have been proved by using the homotopy ideas offered by Ermentrout and McLeod [13]. The speed index function and the stability index function

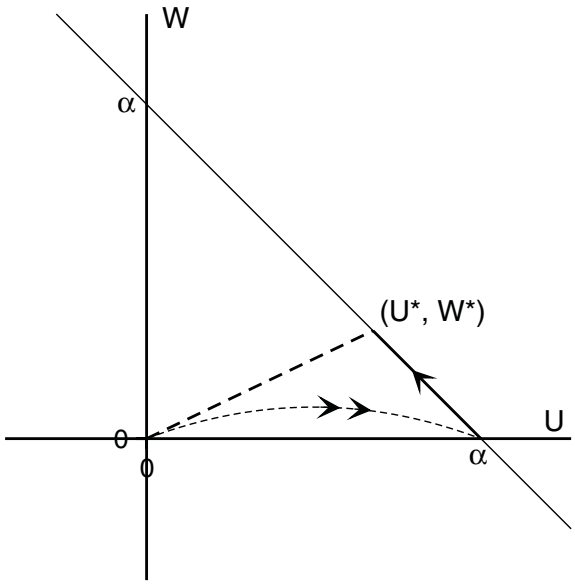


Figure 6. *A singular heteroclinic orbit of system* (69)–(70).

have not been found. However, by using the speed index function defined in (23), we may provide nice estimates on the speed. Note also that

$$\lim_{\beta\to\infty} \mathcal{H}(\beta, u) = H(u-\theta), \tag{82}$$

where H is the well-known Heaviside step function, whose derivative is the famous Dirac delta impulse function. This may be a useful hint on how to construct the speed index function with large $\beta > 0$.

5.5. Stimulus-locked traveling wave front. Consider the inhomogeneous scalar equation

$$u_t + u = \alpha \int_{\mathbb{R}} K(x-y) H\left(u\left(x, t - \frac{1}{c}|x-y|\right) - \theta\right) dy + \mathcal{I}(x,t), \tag{83}$$

where $\mathcal{I}(x,t) = \mathcal{I}(x+\zeta t)$ is a given function of the new variable $z = x+\zeta t$, for some constant ζ. By using the ideas presented in this work, it is not too difficult to show that there exists a solution of the form $u(x,t) = U(x+\zeta t)$, provided that $K \geq 0$ and $\mathcal{I}' \geq 0$ on $\mathbb{R}$, and $\theta < \frac{\alpha}{2} + \mathcal{I}(\mathcal{Z})$, for some real number $\mathcal{Z}$. This result might also be correct under weaker conditions on K and $\mathcal{I}$. We are not going to discuss the details. This part is motivated by the interesting work of Folias and Bressloff [19, section 2].

5.6. Steady state solutions. Consider the following scalar inhomogeneous integral differential equation:

$$u_t + u = \alpha \int_{\mathbb{R}} K(x-y) H\left(u\left(y, t - \frac{1}{c}|x-y|\right) - \theta\right) dy + \mathcal{I}(x). \tag{84}$$

Steady-state solutions are independent of time. Thus, they should satisfy

$$U(x) = \alpha \int_{\mathbb{R}} K(x-y)H(U(y)-\theta)dy + \mathcal{I}(x).$$

If $\alpha = 2\theta$ and $\mathcal{I} = 0$, then there are two trivial steady states,

$$U_+(z) = \alpha \int_{-\infty}^{z} K(x)dx \quad \text{and} \quad U_-(z) = \alpha \int_{z}^{\infty} K(x)dx,$$

such that

$$U_+(0) = \theta, \qquad \lim_{z\to-\infty} U_+(z) = 0, \qquad \lim_{z\to\infty} U_+(z) = \alpha, \qquad \lim_{z\to\pm\infty} U_+{}'(z) = 0,$$

for the first steady state, and

$$U_-(0) = \theta, \qquad \lim_{z\to-\infty} U_-(z) = \alpha, \qquad \lim_{z\to\infty} U_-(z) = 0, \qquad \lim_{z\to\pm\infty} U_-{}'(z) = 0,$$

for the second steady state.

Now let us consider the case $\mathcal{I} \neq 0$. It is straightforward to show that (I) the solution

$$U(x) = \alpha \int_{-\infty}^{x-\mathcal{Z}} K(\xi)d\xi + \mathcal{I}(x)$$

if (i) $\frac{\alpha}{2} + \mathcal{I}(\mathcal{Z}) = \theta$ for some real number $\mathcal{Z}$, (ii) the kernel function $K \geq 0$, and (iii) the derivative of the inhomogeneity $\mathcal{I}' \geq 0$ on $\mathbb{R}$; and (II) the solution

$$U(x) = \alpha \int_{x-\mathcal{Z}}^{\infty} K(\xi)d\xi + \mathcal{I}(x)$$

if (i) $\frac{\alpha}{2} + \mathcal{I}(\mathcal{Z}) = \theta$ for some real $\mathcal{Z}$, (ii) $K \geq 0$, and (iii) $\mathcal{I}' \leq 0$ on $\mathbb{R}$, are stable, monotone steady states of (84). For solution (I), which is increasing on the entire real line, the nonexcited interval is $(-\infty, \mathcal{Z})$ and the excited interval is $(\mathcal{Z}, \infty)$; for solution (II), which is decreasing on the entire real line, the excited interval is $(-\infty, \mathcal{Z})$ and the nonexcited interval is $(\mathcal{Z}, \infty)$. Additionally, we can prove that (III) the solution

$$U(x) = \alpha \int_{x-\Lambda}^{x-\Gamma} K(\xi)d\xi + \mathcal{I}(x)$$

if (i) $K \geq 0$ on $\mathbb{R}$ and $K' \leq 0$ on $\mathbb{R}^+$, (ii)

$$\alpha \int_{0}^{\Lambda-\Gamma} K(x)dx + \mathcal{I}(\Gamma) = \alpha \int_{0}^{\Lambda-\Gamma} K(x)dx + \mathcal{I}(\Lambda) = \theta,$$

and (iii) $\mathcal{I}'(x) \geq 0$ on $(-\infty, \frac{1}{2}(\Gamma+\Lambda))$ and $\mathcal{I}'(x) \leq 0$ on $(\frac{1}{2}(\Gamma+\Lambda), \infty)$, for two real constants Γ and Λ, with $-\infty < \Gamma < \Lambda < \infty$; and (IV) the solution

$$U(x) = \alpha - \alpha \int_{x-\Lambda}^{x-\Gamma} K(\xi)d\xi + \mathcal{I}(x)$$

if (i) $K \geq 0$ on $\mathbb{R}$ and $K' \leq 0$ on $\mathbb{R}^+$, (ii)

$$\alpha - \alpha \int_0^{\Lambda-\Gamma} K(x)dx + \mathcal{I}(\Gamma) = \alpha - \alpha \int_0^{\Lambda-\Gamma} K(x)dx + \mathcal{I}(\Lambda) = \theta,$$

and (iii) $\mathcal{I}'(x) \leq 0$ on $(-\infty, \frac{1}{2}(\Gamma+\Lambda))$ and $\mathcal{I}'(x) \geq 0$ on $(\frac{1}{2}(\Gamma+\Lambda), \infty)$, for two constants Γ and Λ, with $-\infty < \Gamma < \Lambda < \infty$, are unstable steady states with a global maximum or a global minimum of (84). For solution (III), the nonexcited interval is $(-\infty, \Gamma) \cup (\Lambda, \infty)$ and the excited interval is (Γ, Λ); for solution (IV), the excited interval is $(-\infty, \Gamma) \cup (\Lambda, \infty)$ and the nonexcited interval is (Γ, Λ). The solutions in (I) and (II) are complementary to each other, and the solutions in (III) and (IV) are also complementary to each other. These steady states are bounded on $\mathbb{R}$ if and only if the inhomogeneous functions $\mathcal{I}$ are bounded on $\mathbb{R}$. Examples for each of the above four cases are

$$\begin{aligned}
&\text{(I)} \ \mathcal{I}(x) = \lambda\big(1 + \alpha \sinh x + \beta \tanh x + \gamma e^x\big),\\
&\text{(II)} \ \mathcal{I}(x) = \mu - \lambda\big(1 + \alpha \sinh x + \beta \tanh x + \gamma e^x\big),\\
&\text{(III)} \ \mathcal{I}(x) = \lambda \tanh\big(\alpha + \beta(x-\Gamma)(\Lambda - x)\big),\\
&\text{(IV)} \ \mathcal{I}(x) = \lambda \tanh\big(\alpha + \beta(x-\Gamma)(x - \Lambda)\big),
\end{aligned}$$

respectively, where $\alpha \geq 0$, $\beta \geq 0$, $\gamma \geq 0$, $\lambda \geq 0$, $\mu \geq 0$, and $-\infty < \Gamma < \Lambda < \infty$ are real constants.

In this paper we treated the synaptic rate α, the threshold θ, and the speed of action potentials to be constants. In the real neuronal network, however, they may be functions of position x, thus leading to inhomogeneous equations. It turns out that the inhomogeneity may block, reflect, or transmit waves, depending on the intrinsic properties of the neuronal medium. It has been confirmed experimentally in [32] that modulation of the intrinsic properties of the network may speed up, slow down, block, or reflect traveling waves. This will be our next research goal.

6. Appendices.

The main goal of this section is to prove the existence and uniqueness of the front and the speed to make this paper complete and self-contained. The second goal is to provide a summary of influences on wave, speed, and stability, based on the proofs.

6.1. Appendix 1—The proof of Theorem 1.

In this subsection we will first solve an ODE subject to certain boundary conditions at negative infinity to obtain the representation of the front. Then, we present some preliminary analysis on the wave solution and the speed.

Proof. Due to the presence of the Heaviside step function in the IDE (13), it is very possible to derive the exact solution in terms of the kernel function as well as the parameters α, θ, and c. To do this, let us simplify the nonlocal term. Because of the translation invariance of the traveling wave, we may assume without loss of generality that $U(0) = \theta$ and $U'(0) > 0$, and $U < \theta$ on $(-\infty, 0)$, and $U > \theta$ on $(0, \infty)$.

Let us first solve the existence of the front for the special case $\mu = c$. The traveling wave

equation is

$$\begin{aligned} cU' + U &= \alpha \int_{\mathbb{R}} K(z-y) H\big(U(y-|z-y|)-\theta\big) dy \\ &= \frac{\alpha}{2} \int_{-\infty}^{z} K\left(\frac{z-y}{2}\right) H(U(y)-\theta) dy + \frac{\alpha}{2} H(U(z)-\theta). \end{aligned}$$

In other words,

$$\begin{aligned} cU' + U &= 0 && \text{on } (-\infty, 0), \\ cU' + U &= \alpha \int_0^{z/2} K(x) dx + \frac{\alpha}{2} && \text{on } (0, \infty). \end{aligned}$$

Solving this equation, we find the following solution:

$$\begin{aligned} U(z) &= 0 && \text{for } z < 0, \\ U(z) &= \frac{\alpha}{2} + \alpha \int_0^{z/2} K(x) dx - \frac{\alpha}{2} \int_0^z \exp\left(\frac{x-z}{c}\right) K\left(\frac{x}{2}\right) dx && \text{for } z > 0, \\ U'(z) &= \frac{\alpha}{2c} \int_0^z \exp\left(\frac{x-z}{c}\right) K\left(\frac{x}{2}\right) dx && \text{for } z > 0, \\ \lim_{z\to\infty} U(z) &= \alpha. \end{aligned}$$

This solution is discontinuous at $z = 0$ but is continuous everywhere else.

Now let us consider the case $0 < \mu < c$. Let $s(x) = H(x) - H(-x)$ be the usual sign function, that is, $s(x) = -1$ for all $x < 0$, $s(0) = 0$, and $s(x) = 1$ for all $x > 0$. To simplify the nonlocal integral, we make the following change of variables:

$$\eta = y - \frac{\mu}{c}|z - y|.$$

Then for each fixed number z, η is a strictly increasing function of y, and $s(z-y) = s(z-\eta)$. Clearly, $z > \eta$ if $z > y$, and $z < \eta$ if $z < y$. Moreover, we have the important relationships

$$z - y = \frac{c}{c + s(z-\eta)\mu}(z-\eta)$$

and

$$dy = \frac{c}{c + s(z-\eta)\mu} d\eta - \frac{c\mu(z-\eta)}{[c + s(z-\eta)\mu]^2} s'(z-\eta) d\eta.$$

Now the traveling wave equation reduces to the simple ODE

$$\begin{aligned} \mu U' + U &= \alpha \int_{\mathbb{R}} \frac{c}{c + s(z-\eta)\mu} K\left(\frac{c}{c + s(z-\eta)\mu}(z-\eta)\right) H(U(\eta)-\theta) d\eta \\ &= \alpha \int_0^{\infty} \frac{c}{c + s(z-\eta)\mu} K\left(\frac{c}{c + s(z-\eta)\mu}(z-\eta)\right) d\eta \\ &= \alpha \int_{-\infty}^{cz/(c+s(z)\mu)} K(x) dx, \end{aligned}$$

where

$$x = \frac{c}{c + s(z-\eta)\mu}(z-\eta), \qquad dx = -\frac{c}{c + s(z-\eta)\mu}d\eta + \frac{c\mu(z-\eta)}{[c + s(z-\eta)\mu]^2}s'(z-\eta)d\eta.$$

If we apply fundamental ideas in differential equations, then we obtain the wave solution representation

$$\begin{aligned} U(c,\mu,z) &= \alpha \int_{-\infty}^{cz/(c+s(z)\mu)} K(x)dx \\ &- \alpha \int_{-\infty}^{z} \exp\left(\frac{x-z}{\mu}\right)\left\{\frac{c}{c+s(x)\mu}K\left(\frac{c}{c+s(x)\mu}x\right)\right\}dx. \end{aligned} \tag{85}$$

Moreover, we have

$$U_z(c,\mu,z) = \frac{\alpha}{\mu}\int_{-\infty}^{z} \exp\left(\frac{x-z}{\mu}\right)\left\{\frac{c}{c+s(x)\mu}K\left(\frac{c}{c+s(x)\mu}x\right)\right\}dx. \tag{86}$$

Clearly there hold the following limits:

$$\lim_{z\to-\infty} U(c,\mu,z) = 0, \qquad \lim_{z\to\infty} U(c,\mu,z) = \alpha, \qquad \lim_{z\to\pm\infty} U_z(c,\mu,z) = 0. \tag{87}$$

Below we will investigate whether the solution really crosses the threshold as expected. To do that, we have to study on which interval the solution is increasing or decreasing, and where it attains a maximum or a minimum.

We have prescribed the phase condition by $U(c,\mu,0) = \theta$ at the beginning. That is,

$$\frac{\alpha}{2} - \alpha\int_{-\infty}^{0} \exp\left(\frac{x}{\mu}\right)\frac{c}{c-\mu}K\left(\frac{c}{c-\mu}x\right)dx = \theta, \tag{88}$$

for some number μ satisfying $0 < \mu < c$. If we can find the unique speed μ_0 and demonstrate that the solution $U(c,\mu_0,z)$ really crosses the threshold θ at $z = 0$, and it really lies below and above the threshold on $(-\infty, 0)$ and $(0,\infty)$, respectively, then we can finish the proof immediately. The existence and uniqueness of the positive speed μ_0 will be proved in Lemma 1. The crossing can be guaranteed by the assumptions on K.

First, if $U(c,\mu_0,0) = \theta$, then $U_z(c,\mu_0,0) = \frac{1}{\mu_0}(\frac{\alpha}{2} - \theta) > 0$. Clearly, if $K \geq 0$, then $U_z \geq 0$ (i.e., if the synaptic coupling is of pure excitation, then the front U is increasing). In particular, if K is positive almost everywhere, then U is strictly increasing. Therefore, kernel functions corresponding to pure excitation give rise to monotone solutions, thus crossing the threshold as desired. Let us look at solutions of (13) with Mexican hat or upside down Mexican hat kernel functions. For all $z \leq 0$ we have

$$\begin{aligned} U_z(c,\mu,z) &= \frac{\alpha}{\mu}\int_{-\infty}^{z} \exp\left(\frac{x-z}{\mu}\right)\frac{c}{c-\mu}K\left(\frac{c}{c-\mu}x\right)dx \\ &= \frac{\alpha}{\mu}\exp\left(-\frac{z}{\mu}\right)\int_{-\infty}^{cz/(c-\mu)} \exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi, \end{aligned}$$

and for all $z > 0$ we have

$$\begin{aligned}
&U_z(c,\mu,z)\\
&= \frac{\alpha}{\mu}\int_{-\infty}^{0} \exp\left(\frac{x-z}{\mu}\right)\frac{c}{c-\mu}K\left(\frac{c}{c-\mu}x\right)dx + \frac{\alpha}{\mu}\int_{0}^{z} \exp\left(\frac{x-z}{\mu}\right)\frac{c}{c+\mu}K\left(\frac{c}{c+\mu}x\right)dx\\
&= \frac{\alpha}{\mu}\exp\left(-\frac{z}{\mu}\right)\int_{-\infty}^{0} \exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi + \frac{\alpha}{\mu}\exp\left(-\frac{z}{\mu}\right)\int_{0}^{cz/(c+\mu)} \exp\left(\frac{c+\mu}{c\mu}\xi\right)K(\xi)d\xi\\
&= \frac{\alpha}{\mu}\exp\left(-\frac{z}{\mu}\right)\left\{\int_{-\infty}^{0} \exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi + \int_{0}^{cz/(c+\mu)} \exp\left(\frac{c+\mu}{c\mu}\xi\right)K(\xi)d\xi\right\}.
\end{aligned}$$

Moreover,

$$U_z(c,\mu_0,z) = \frac{\alpha}{\mu_0}\exp\left(-\frac{z}{\mu_0}\right)\int_{-\infty}^{cz/(c-\mu_0)} \exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right)K(\xi)d\xi, \qquad z \le 0,$$

and

$$U_z(c,\mu_0,z) = \frac{1}{\mu_0}\exp\left(-\frac{z}{\mu_0}\right)\left\{\frac{\alpha}{2} - \theta + \alpha\int_{0}^{cz/(c+\mu_0)} \exp\left(\frac{c+\mu_0}{c\mu_0}\xi\right)K(\xi)d\xi\right\}, \qquad z \ge 0.$$

Let the synaptic coupling be of lateral inhibition type. Recall that $K \ge 0$ on $(-M,N)$, and $K \le 0$ on $(-\infty,-M)\cup(N,\infty)$. Obviously $U_z \le 0$ on $(-\infty,-M)$; in particular, $U_z(c,\mu_0,-M) < 0$ and $U_z(c,\mu_0,0) > 0$. Therefore, there exists a unique negative number $z_0 \in (-M,0)$ such that $U_z(c,\mu_0,z_0) = 0$, $U_z \le 0$ on $(-\infty,z_0)$, and $U_z \ge 0$ on (z_0,∞); i.e., the front is decreasing on $(-\infty,z_0)$ and it is increasing on (z_0,N). Therefore the front U attains a local minimum at some $z_0 \in (-\infty,0)$. Let us investigate the behavior of U on the right half line (N,∞). Differentiating the traveling wave equation with respect to z, we get

$$\mu_0 U_{zz} + U_z = \frac{c\alpha}{c+s(z)\mu_0}K\left(\frac{c}{c+s(z)\mu_0}z\right) \le 0.$$

If we multiply this inequality by the integrating factor $\exp(z/\mu_0)$, we get

$$\left[\mu_0\exp\left(\frac{z}{\mu_0}\right)U_z(c,\mu_0,z)\right]_z = \exp\left(\frac{z}{\mu_0}\right)(\mu_0 U_{zz} + U_z) \le 0.$$

Now, if we integrate this inequality with respect to z over (z,∞), we obtain the estimate

$$0 < \lim_{z\to\infty}\left[\exp\left(\frac{z}{\mu_0}\right)U_z(c,\mu_0,z)\right] \le \exp\left(\frac{z}{\mu_0}\right)U_z(c,\mu_0,z).$$

Therefore

$$U_z(c,\mu_0,z) > 0$$

for all $z > N$. Overall, $U(c,\mu_0,z)$ is decreasing on $(-\infty,z_0)$, and it is increasing on (z_0,∞), and $U(c,\mu_0,z_0)$ is a global minimum.

Let the synaptic coupling be of lateral excitation. Note that $K \geq 0$ on $(-\infty, -M) \cup (N, \infty)$ and $K \leq 0$ on $(-M, N)$. Clearly $U_z \geq 0$ on $(-\infty, -M)$. The fact $U_z(c, \mu_0, 0) > 0$ implies that $U_z \geq 0$ on $(-\infty, 0)$. For each upside down Mexican hat kernel function, it is easy to show that

$$\int_0^{cz/(c+\mu_0)} \exp\left(\frac{c+\mu_0}{c\mu_0}\xi\right) K(\xi)d\xi > 0$$

for all sufficiently large $z > 0$.

Let us find the point z_N where the synaptic coupling makes the most negative contribution, starting from $z = -M$. To this end, we set the upper limit $\frac{cz_N}{c+\mu_0} = N$; thus $z_N = \frac{c+\mu_0}{c}N$. Additionally, if $U_z\left(c, \mu_0, \frac{c+\mu_0}{c}N\right) \geq 0$, then $U_z \geq 0$ on $\mathbb{R}$. In this case the front is a monotone increasing curve and crosses the threshold exactly once, at $z = 0$, as expected. If, however, $U_z\left(c, \mu_0, \frac{c+\mu_0}{c}N\right) < 0$, then based on the profile of the upside down Mexican hat kernel function, U_z must change sign from positive to negative and then to positive; hence there exists a unique pair of positive numbers z_1 and z_2, $z_1 \in (0, \frac{c+\mu_0}{c}N)$ and $z_2 \in (\frac{c+\mu_0}{c}N, \infty)$, such that $U_z(c, \mu_0, z_1) = 0$ and $U_z(c, \mu_0, z_2) = 0$, $U_z \geq 0$ on $(-\infty, z_1) \cup (z_2, \infty)$ and $U_z \leq 0$ on (z_1, z_2). The front attains a local maximum at $z = z_1$ and a local minimum at $z = z_2$. We have to investigate whether the solution crosses the threshold on the right half line. It is easy to see that

$$\begin{aligned}
U(c, \mu_0, z_1) &= \alpha \int_{-\infty}^{cz_1/(c+s(z_1)\mu_0)} K(x)dx \\
&= \alpha \int_{-\infty}^{0} K(x)dx + \alpha \int_0^{cz_1/(c+s(z_1)\mu_0)} K(x)dx \\
&= \frac{\alpha}{2} + \alpha \int_0^{cz_1/(c+\mu_0)} K(x)dx \\
&< \alpha, \qquad \text{by the assumption } K \leq 0 \text{ on } (-M, N), \\
U(c, \mu_0, z_2) &= \frac{\alpha}{2} + \alpha \int_0^{N} K(x)dx + \alpha \int_N^{cz_2/(c+s(z_2)\mu_0)} K(x)dx \\
&> \frac{\alpha}{2} + \alpha \int_0^{N} K(x)dx \geq \theta, \qquad \text{by assumption (11).}
\end{aligned}$$

Thus the local maximum $U(c, \mu_0, z_1)$ is smaller than α, and the local minimum $U(c, \mu_0, z_2)$ is larger than the threshold. Therefore U crosses the threshold θ only once.

Overall, for each of the three classes of synaptic couplings, the front crosses the threshold only once. See Figures 7 and 8. For the extreme case $\mu = c$, although the solution is not continuous, the mathematical analysis of the existence and uniqueness of the front is very similar to the case $0 < \mu < c$, and we can demonstrate that the wave crosses the threshold exactly once. The limits of the speed μ_0 as $c \to 0$, $c \to \infty$, $\frac{\theta}{\alpha} \to 0$, and $\frac{\theta}{\alpha} \to \frac{1}{2}$ have been proved in Theorems 3 and 4. The limits of the front U as $c \to 0$, $c \to \infty$, $\frac{\theta}{\alpha} \to 0$, and $\frac{\theta}{\alpha} \to \frac{1}{2}$ can be computed straightforwardly. The details are omitted. ■

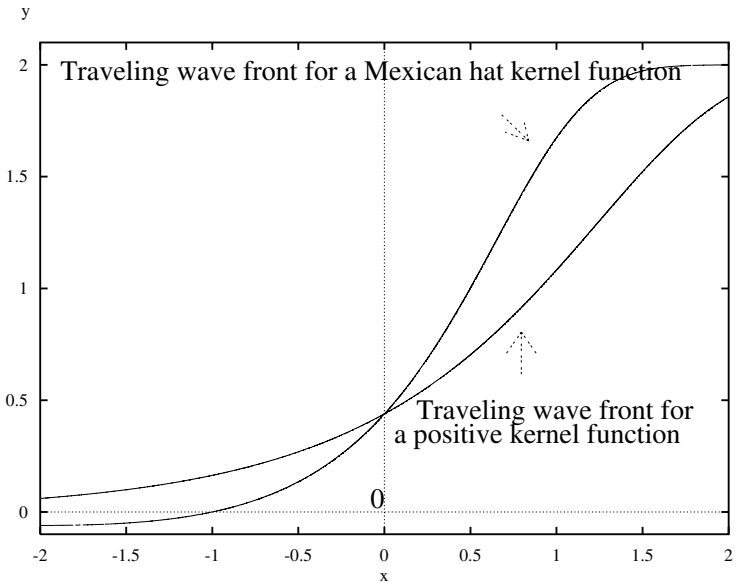


Figure 7. *Comparison of traveling wave fronts (between a positive kernel and a Mexican hat kernel).*

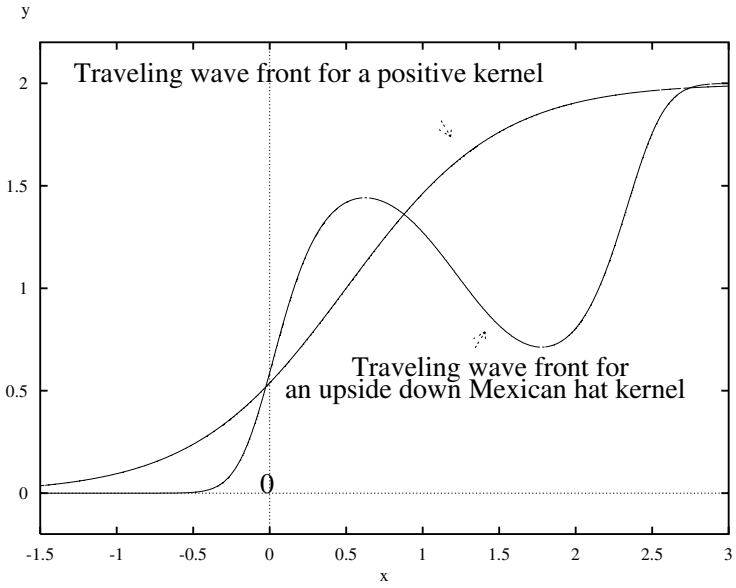


Figure 8. *Comparison of traveling wave fronts (between a positive kernel and an upside down Mexican hat kernel).*

6.2. Appendix 2—The proof of Lemma 1. We have defined the speed index function φ earlier by

$$\varphi(\mu) = \alpha \int_{-\infty}^{0} \exp\left(\frac{c-\mu}{c\mu}\xi\right) K(\xi)d\xi$$

on $(0, c)$. The speed μ_0 of the front will be proved to be the unique solution of the equation

$$\varphi(\mu) = \frac{\alpha}{2} - \theta.$$

Clearly we have

$$\lim_{\mu\to 0} \varphi(\mu) = 0 < \frac{\alpha}{2} - \theta < \frac{\alpha}{2} = \lim_{\mu\to c} \varphi(\mu).$$

The existence of a positive number μ_0 satisfying $0 < \mu_0 < c$ is obvious. It suffices to establish the uniqueness. We will show that $\varphi' > 0$ for all kernel functions in classes (A) and (B). Also we will demonstrate that $\varphi' < 0$ on $(0, \mu_\#)$, $\varphi'(\mu_\#) = 0$, and $\varphi' > 0$ on $(\mu_\#, c)$, where the positive number $\mu_\#$ depends on the kernel K, for all kernels in class (C). Overall, the graph of φ crosses the line $\varphi = \frac{\alpha}{2} - \theta$ only once if $\alpha > 2\theta > 0$.

See Figures 9 and 10.

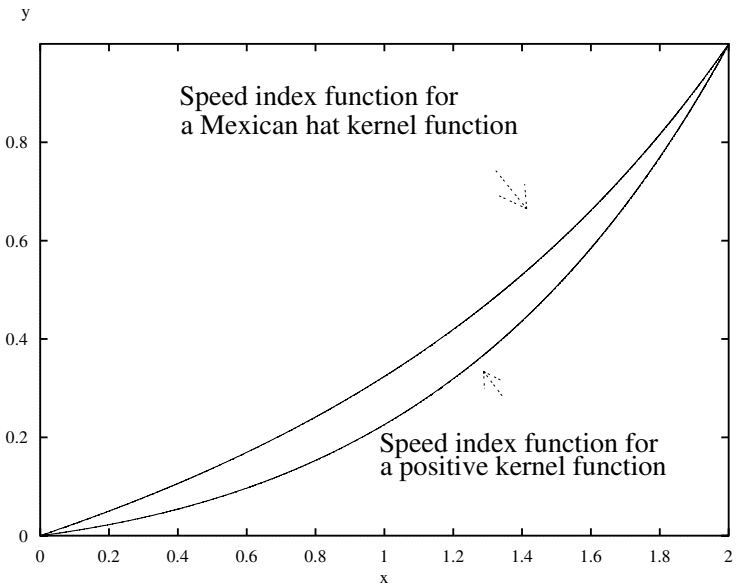


Figure 9. *Comparison of speed index functions (between a positive kernel and a Mexican hat kernel).*

(A) For all nonnegative kernels (either symmetric or asymmetric, either on-center or off-center, either with or without compact support), clearly $\int_{-\infty}^{0} K(x)dx = \frac{1}{2}$, so we must have $\varphi'(\mu) > 0$ on $(0, c)$.

(B) For each Mexican hat kernel function, recall that $K \geq 0$ on $(-M, 0)$ and $K \leq 0$ on $(-\infty, -M)$. Therefore, there hold the estimates

$$|\xi| \exp\left(\frac{c-\mu}{c\mu}\xi\right) K(\xi) \geq \exp\left(-\frac{c-\mu}{c\mu}M\right) |\xi| K(\xi) \geq 0$$

on $(-M, 0)$, and

$$0 \geq |\xi| \exp\left(\frac{c-\mu}{c\mu}\xi\right) K(\xi) \geq \exp\left(-\frac{c-\mu}{c\mu}M\right) |\xi| K(\xi)$$

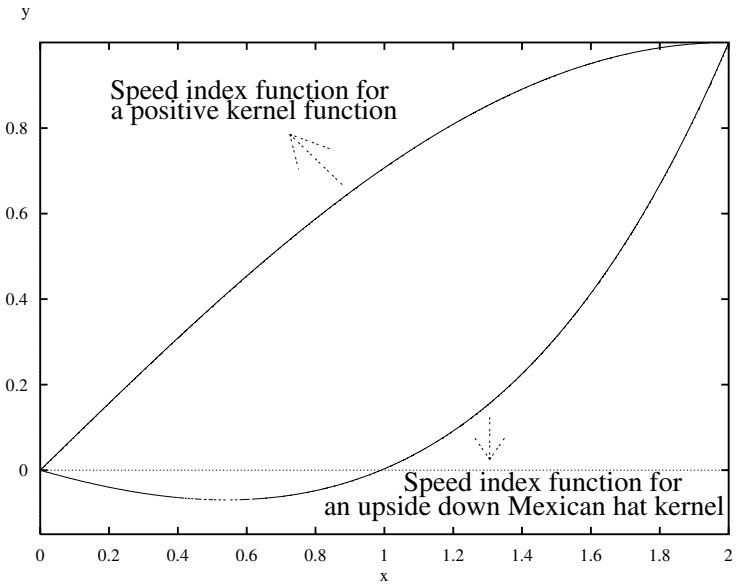


Figure 10. *Comparison of speed index functions (between a positive kernel and an upside down Mexican hat kernel).*

on $(-\infty, -M)$. On some nonempty open interval $(\tau_1, \tau_2) \subset (-\infty, 0)$, there must hold the inequality ">," because $\int_{-\infty}^{0} K(x)dx = \frac{1}{2}$. Hence for all μ with $0 < \mu < c$ we have $\varphi'(\mu) > 0$, because

$$\varphi'(\mu) > \frac{\alpha}{\mu^2} \exp\left(-\frac{c-\mu}{c\mu} M\right) \int_{-\infty}^{0} |\xi| K(\xi) d\xi \geq 0.$$

(C) This is the most challenging and interesting case. Define a sequence of nonlinear smooth functions $\{\varphi_n(\mu)\}$ on $(0, c)$, where

$$\varphi_n(\mu) = \alpha \int_{-\infty}^{0} |\xi|^n \exp\left(\frac{c-\mu}{c\mu}\xi\right) K(\xi) d\xi,$$

for all integers $n \geq 1$. Then

$$\varphi_n{}'(\mu) = \frac{\alpha}{\mu^2} \int_{-\infty}^{0} |\xi|^{n+1} \exp\left(\frac{c-\mu}{c\mu}\xi\right) K(\xi) d\xi = \frac{\varphi_{n+1}(\mu)}{\mu^2},$$

$$\lim_{\mu \to 0} \varphi_n(\mu) = 0, \qquad \lim_{\mu \to c} \varphi_n(\mu) = \alpha \int_{-\infty}^{0} |\xi|^n K(\xi) d\xi > 0.$$

For each upside down Mexican hat kernel function, recall that $K \leq 0$ on $(-M, 0)$, $K \geq 0$ on $(-\infty, -M)$, and that $\int_{-\infty}^{0} K(x)dx = \frac{1}{2}$. Therefore, we can pick up two positive numbers δ and Δ, with $\Delta > \delta$ (note that the δ is not the delta function), such that

$$\int_{-M}^{0} K(\xi) d\xi + \int_{-M(1+\Delta)}^{-M(1+\delta)} K(\xi) d\xi > 0.$$

Fix μ such that $0 < \mu < c$. Now we have (note that below $\xi = Mx$)

$$\begin{aligned}
\frac{1}{\alpha}\varphi_{n+1}(\mu) &= \int_{-\infty}^{0} |\xi|^{n+1} \exp\left(\frac{c-\mu}{c\mu}\xi\right) K(\xi)d\xi \\
&= M^{n+2}\int_{-\infty}^{0} |x|^{n+1} \exp\left(\frac{c-\mu}{c\mu}Mx\right) K(Mx)dx \\
&= M^{n+2}\int_{-1}^{0} |x|^{n+1} \exp\left(\frac{c-\mu}{c\mu}Mx\right) K(Mx)dx \\
&\quad + M^{n+2}\int_{-\infty}^{-1} |x|^{n+1} \exp\left(\frac{c-\mu}{c\mu}Mx\right) K(Mx)dx \\
&\geq M^{n+2}\int_{-1}^{0} |x|^{n+1} \exp\left(\frac{c-\mu}{c\mu}Mx\right) K(Mx)dx \\
&\quad + M^{n+2}\int_{-1-\Delta}^{-1-\delta} |x|^{n+1} \exp\left(\frac{c-\mu}{c\mu}Mx\right) K(Mx)dx \\
&\geq M^{n+2}\int_{-1}^{0} K(Mx)dx + M^{n+2}\int_{-1-\Delta}^{-1-\delta} K(Mx)dx \\
&= M^{n+1}\int_{-M}^{0} K(\xi)d\xi + M^{n+1}\int_{-M(1+\Delta)}^{-M(1+\delta)} K(\xi)d\xi > 0,
\end{aligned}$$

where we have applied the following simple estimates:

$$\begin{aligned}
&|x|^{n+1} \exp\left(\frac{c-\mu}{c\mu}Mx\right) \leq 1 \quad \text{and} \quad K(Mx) \leq 0 \text{ on } (-1,0) \text{ for all } n \geq 0, \\
&|x|^{n+1} \exp\left(\frac{c-\mu}{c\mu}Mx\right) \geq 1 \quad \text{and} \quad K(Mx) \geq 0 \text{ on } (-1-\Delta, -1-\delta),
\end{aligned}$$

provided that the integer n is sufficiently large, say $n \geq N$, where

$$N = 1 + \left[\left[\frac{c-\mu}{c\mu}\frac{M(1+\Delta)}{\ln(1+\delta)}\right]\right].$$

Of course, here $[[x]]$ represents the greatest integer function of x. Therefore we obtain the result

$$\varphi_n{}'(\mu) = \frac{1}{\mu^2}\varphi_{n+1}(\mu) > \frac{M^{n+1}\alpha}{\mu^2}\left[\int_{-M}^{0} K(\xi)d\xi + \int_{-M(1+\Delta)}^{-M(1+\delta)} K(\xi)d\xi\right] > 0.$$

Indeed, this result is also true on the open interval (ν, c). For the same reason as above, we

have

$$\begin{aligned}\int_{-\infty}^{0}|\xi|^n K(\xi)d\xi &= M^{n+1}\int_{-\infty}^{0}|x|^n K(Mx)dx\\ &= M^{n+1}\int_{-1}^{0}|x|^n K(Mx)dx + M^{n+1}\int_{-\infty}^{-1}|x|^n K(Mx)dx\\ &\geq M^{n+1}\int_{-1}^{0}|x|^n K(Mx)dx + M^{n+1}\int_{-1-\Delta}^{-1-\delta}|x|^n K(Mx)dx\\ &\geq M^{n+1}\int_{-1}^{0}K(Mx)dx + M^{n+1}\int_{-1-\Delta}^{-1-\delta}K(Mx)dx\\ &= M^{n}\int_{-M}^{0}K(\xi)d\xi + M^{n}\int_{-M(1+\Delta)}^{-M(1+\delta)}K(\xi)d\xi > 0\end{aligned}$$

for all integers $n \geq 1$. From the above analysis we may also conclude that if $\varphi_n(\mu) = 0$ at some $0 < \mu < c$, then for the same μ, we have $\varphi_{n+k}(\mu) > 0$, because

$$\begin{aligned}\varphi_{n+k}(\mu) &= \alpha\int_{-\infty}^{0}|\xi|^{n+k}\exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi\\ &= \alpha M^{n+k+1}\int_{-\infty}^{0}|x|^{n+k}\exp\left(\frac{c-\mu}{c\mu}Mx\right)K(Mx)dx\\ &= \alpha M^{n+k+1}\int_{-1}^{0}|x|^{n+k}\exp\left(\frac{c-\mu}{c\mu}Mx\right)K(Mx)dx\\ &\quad + \alpha M^{n+k+1}\int_{-\infty}^{-1}|x|^{n+k}\exp\left(\frac{c-\mu}{c\mu}Mx\right)K(Mx)dx\\ &> \alpha M^{n+k+1}\int_{-1}^{0}|x|^{n}\exp\left(\frac{c-\mu}{c\mu}Mx\right)K(Mx)dx\\ &\quad + \alpha M^{n+k+1}\int_{-\infty}^{-1}|x|^{n}\exp\left(\frac{c-\mu}{c\mu}Mx\right)K(Mx)dx\\ &= \alpha M^{n+k+1}\int_{-\infty}^{0}|x|^{n}\exp\left(\frac{c-\mu}{c\mu}Mx\right)K(Mx)dx\\ &= \alpha M^{k}\int_{-\infty}^{0}|\xi|^{n}\exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi\\ &= 0,\end{aligned}$$

at the same μ, for all integers $k \geq 1$.

On the other hand, by fixing the integer $n = N$ and making the change of variable $\xi = \mu x$ with $0 < \mu < c$, we have

$$\int_{-\infty}^{0}|\xi|^{n+1}\exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi = \mu^{n+2}\int_{-\infty}^{0}|x|^{n+1}\exp\left(\frac{c-\mu}{c}x\right)K(\mu x)dx,$$

where

$$\lim_{\mu\to 0}\int_{-\infty}^{0}|x|^{n+1}\exp\left(\frac{c-\mu}{c}x\right)K(\mu x)dx$$
$$=K(0)\int_{-\infty}^{0}|x|^{n+1}\exp(x)dx=(n+1)!K(0)<0$$

if $K(0)<0$. Therefore, there exists a small number $\mu=\mu_{n+1}>0$ such that

$$\varphi_{n+1}(\mu)=\alpha\int_{-\infty}^{0}|\xi|^{n+1}\exp\left(\frac{c-\mu}{c\mu}\xi\right)K(\xi)d\xi$$
$$=\alpha\mu^{n+2}\int_{-\infty}^{0}|x|^{n+1}\exp\left(\frac{c-\mu}{c}x\right)K(\mu x)dx<0,$$

for all μ with $0<\mu<\mu_{n+1}$. This result is also true if $K(0)=0$, $K\geq 0$ on $(-\infty,-M)$, and $K\leq 0$ on $(-M,0)$, for some positive number $M>0$. Set $A_*=\inf\{\mu\leq c:\varphi_{n+1}(\mu)>0\}$ and $A^*=\sup\{\mu\geq 0:\varphi_{n+1}(\mu)<0\}$. If $\varphi_{n+1}(\mu_\#)=0$ for some $\mu_\#\in(0,c)$, then using analysis similar to the above, we see $\varphi_{n+2}(\mu_\#)={\mu_\#}^2{\varphi_{n+1}}'(\mu_\#)>0$. Thus the graph of the smooth function $\varphi_{n+1}(\mu)$ crosses the μ-axis exactly once. Without any difficulty, we may conclude that $A_*=A^*$. Now $\varphi_{n+1}(A_*)=0$. Below we will use the induction idea. Note that

$$\lim_{\mu\to 0}\varphi_n(\mu)=0.$$

Now it is easy to derive $\varphi_n(\mu)<0$ on $(0,\mu_n)$ and $\varphi_n(\mu)>0$ on (μ_n,c) for some μ_n. Recall that we have

$${\varphi_{n-1}}'(\mu)=\frac{\varphi_n(\mu)}{\mu^2}\qquad\text{and}\qquad\lim_{\mu\to 0}\varphi_{n-1}(\mu)=0.$$

Therefore, it is easy to conclude that $\varphi_{n-1}(\mu)<0$ on $(0,\mu_{n-1})$ and $\varphi_{n-1}(\mu)>0$ on (μ_{n-1},c), for some μ_{n-1}, and so on. Eventually, we get $\lim_{\mu\to 0}\varphi_1(\mu)=0$, $\varphi_1(\mu)<0$ on $(0,\mu_1)$, and $\varphi_1(\mu)>0$ on (μ_1,c), for some μ_1. By the induction method, we find that there exists some constant $\mu_m=\mu(m)$ such that $\varphi_m(\mu)<0$ if $0<\mu<\mu_m$, $\varphi_m(\mu)=0$ if $\mu=\mu_m$, and $\varphi_m(\mu)>0$ if $\mu_m<\mu<c$. Note that $0<\mu_{m+1}<\mu_m<c$.

Therefore, for all kernel functions in classes (A), (B), and (C), there exists a unique speed $\mu_0=\mu_0(c,\alpha,\theta,K)>0$ such that $0<\mu_0<c$, $\varphi(\mu_0)=\frac{\alpha}{2}-\theta$ and $U(c,\mu_0,0)=\theta$.

Note that the speed is a smooth function of β, where $\beta=\frac{1}{2}-\frac{\theta}{\alpha}$. To see if the speed is increasing with respect to β, differentiating the following equation with respect to β,

$$\int_{-\infty}^{0}\exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right)K(\xi)d\xi=\beta=\frac{1}{2}-\frac{\theta}{\alpha},$$

we have

$$\left[\frac{1}{{\mu_0}^2}\int_{-\infty}^{0}|\xi|\exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right)K(\xi)d\xi\right]\frac{\partial\mu_0}{\partial\beta}=1.$$

Therefore, $\frac{\partial \mu_0}{\partial \beta} = \frac{\alpha}{\varphi'(\mu_0)} > 0$, based on the above analysis. So μ_0 is an increasing function of β. Moreover, it is not difficult to show that

$$\lim_{\beta \to 0} \mu_0(\beta) = 0, \qquad \lim_{\beta \to \frac{1}{2}} \mu_0(\beta) = c.$$

Similarly

$$\frac{\partial \mu_0}{\partial(\theta/\alpha)} = -{\mu_0}^2 / \int_{-\infty}^{0} |\xi| \exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right) K(\xi) d\xi = -\frac{\alpha}{\varphi'(\mu_0)} < 0,$$

$$\left\{\frac{1}{c^2} - \frac{1}{{\mu_0}^2}\frac{\partial \mu_0}{\partial c}\right\} \int_{-\infty}^{0} |\xi| \exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right) K(\xi) d\xi = \left(\frac{{\mu_0}^2}{c^2} - \frac{\partial \mu_0}{\partial c}\right) \frac{\varphi'(\mu_0)}{\alpha} = 0.$$

The proofs of both Lemma 1 and Theorem 1 are now finished. ■

6.3. Summary of influences on traveling wave fronts. Combining all of the proofs of the theorems and Lemma 1, we may give the following summary.

Summary of influence on speed. The speeds of the fronts of the nonlocal model equation (13) with different kinds of synaptic couplings satisfy the estimates

$$0 < \mu_0(c, \alpha, \theta) < c \qquad \text{and} \qquad 0 < \mu_0(c, \alpha, \theta) < \nu_0(\alpha, \theta).$$

Summary of influence on wave profiles. Define the steepness of the front by $\mathcal{S} = U_z(c, \mu_0, 0)$. Note that

$$U_z(c, \mu_0, 0) = \frac{\alpha}{\mu_0} \int_{-\infty}^{0} \exp\left(\frac{c-\mu_0}{c\mu_0}\xi\right) K(\xi) d\xi = \frac{\alpha - 2\theta}{2\mu_0}.$$

The speed μ_0 is an increasing function of c and $\frac{\alpha}{\theta}$ on $(0, \infty)$, and hence $\mathcal{S}$ is a decreasing function of c and $\frac{\alpha}{\theta}$. Therefore, compared with the case where there is no delay, the spatial temporal delay makes the front steep. By the results of Theorem 4, we conclude that

$$\lim_{c \to 0} \mathcal{S} = \infty, \qquad \lim_{c \to \infty} \mathcal{S} = \frac{\alpha - 2\theta}{2\nu_0(\alpha, \theta)},$$

$$\lim_{\frac{\theta}{\alpha} \to 0} \mathcal{S} = \frac{\alpha - 2\theta}{2c}, \qquad \lim_{\frac{\theta}{\alpha} \to \frac{1}{2}} \mathcal{S} = \infty,$$

$$\lim_{\rho \to 0} \mathcal{S} = \frac{\alpha - 2\theta}{2c}, \qquad \lim_{\rho \to \infty} \mathcal{S} = \infty,$$

where, in the first four limits, the kernel function K is fixed, while in the last two limits, the kernel function $K(x)$ is replaced with $\rho K(\rho x)$.

In general, if we keep all of the parameters c, α, and θ fixed, then a positive kernel makes the front increasing, and a negative kernel makes the front decreasing. More precisely, a nonnegative kernel produces a nondecreasing front, a Mexican hat kernel forces the left half of the front to be decreasing while the right half is still increasing, and an upside down Mexican hat kernel forces the middle part of the front to be decreasing while the other parts are still increasing.

Summary of influence on stability. For each kernel function in the three classes, the front is exponentially stable. There is no essential difference in the influence on stability of these fronts.

Acknowledgments. This paper was published under the mercy of Jesus Christ, my Savior and Lord.

The author is very grateful to Professor David Terman at the Department of Mathematics of The Ohio State University for suggesting this problem. He would also like to acknowledge the kind invitation from the Scientific Committee of the Conference on Mathematical Neuroscience (NEUROMATH 2006) held at Principat d'Andorra, September 1–4, 2006. Part of this paper was finished while the author was visiting Professor David Terman at the Mathematical Biosciences Institute of The Ohio State University, August 15 to December 31, 2006, which directly improved this work.

The author is also grateful to Professor Jonathan Bell (University of Maryland, Baltimore, MD) and Professor Eiji Yanagida (Tohoku University, Sendai, Japan) for many fruitful discussions on related topics.

Finally, the author thanks the anonymous referees for their kind suggestions and comments, which made the presentation of this work better.

REFERENCES

[1] S.-I. Amari, *Dynamics of pattern formation in lateral-inhibition type neural fields*, Biolog. Cybernet., 27 (1977), pp. 77–87.

[2] D. G. Aronson and H. F. Weinberger, *Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation*, in Partial Differential Equations and Related Topics (Program, Tulane University, New Orleans, LA, 1974), Lecture Notes in Math. 446, Springer-Verlag, Berlin, New York, 1975, pp. 5–49.

[3] D. G. Aronson and H. F. Weinberger, *Multidimensional nonlinear diffusion arising in population genetics*, Adv. Math., 30 (1978), pp. 33–76.

[4] F. M. Atay and A. Hutt, *Stability and bifurcations in neural fields with finite propagation speed and general connectivity*, SIAM J. Appl. Math., 65 (2005), pp. 644–666.

[5] P. C. Bressloff, *Weakly interacting pulses in synaptically coupled neural media*, SIAM J. Appl. Math., 66 (2005), pp. 57–81.

[6] P. C. Bressloff and S. E. Folias, *Front bifurcations in an excitatory neural network*, SIAM J. Appl. Math., 65 (2004), pp. 131–151.

[7] S. Coombes, G. J. Lord, and M. R. Owen, *Waves and bumps in neuronal networks with axo-dendritic synaptic interactions*, Phys. D, 178 (2003), pp. 219–241.

[8] S. Coombes and M. R. Owen, *Evans functions for integral neural field equations with Heaviside firing rate function*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 574–600.

[9] S. Coombes and M. R. Owen, *Bumps, breathers, and waves in a neural network with spike frequency adaptation*, Phys. Rev. Lett., 94 (2005), 148102.

[10] S. Coombes, *Waves, bumps, and patterns in neural field theories*, Biolog. Cybernet., 93 (2005), pp. 91–108.

[11] R. Curtu and B. Ermentrout, *Pattern formation in a network of excitatory and inhibitory cells with adaptation*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 191–231.

[12] G. B. Ermentrout, *Neural networks as spatio-temporal pattern-forming systems*, Inst. Phys. (Great Britain) Rep. Progr. Phys., 61 (1998), pp. 353–430.

[13] G. B. Ermentrout and J. B. McLeod, *Existence and uniqueness of travelling waves for a neural network*, Proc. Roy. Soc. Edinburgh Sect. A, 123 (1993), pp. 461–478.

[14] J. W. Evans, *Nerve axon equations* I. *Linear approximations*, Indiana Univ. Math. J., 21 (1972), pp. 877–885.

[15] J. W. Evans, *Nerve axon equations* II. *Stability at rest*, Indiana Univ. Math. J., 22 (1972), pp. 75–90.

[16] J. W. Evans, *Nerve axon equations* III. *Stability of the nerve impulse*, Indiana Univ. Math. J., 22 (1972), pp. 577–593.

[17] J. W. Evans, *Nerve axon equations* IV. *The stable and the unstable impulse*, 24 (1975), pp. 1169–1190.
[18] S. E. Folias and P. C. Bressloff, *Breathing pulses in an excitatory neural network*, SIAM J. Appl. Dyn. Syst., 3 (2004), pp. 378–407.
[19] S. E. Folias and P. C. Bressloff, *Stimulus-locked traveling waves and breathers in an excitatory neural network*, SIAM J. Appl. Math., 65 (2005), pp. 2067–2092.
[20] D. Golomb and G. B. Ermentrout, *Continuous and lurching traveling pulses in neuronal networks with delay and spatially decaying connectivity*, Proc. Nat. Acad. Sci., USA, 96 (1999), pp. 13480–13485.
[21] D. Golomb and G. B. Ermentrout, *Effects of delay on the type and velocity of travelling pulses in neuronal networks with spatially decaying connectivity*, Network: Comput. Neural Systems, 11 (2000), pp. 221–246.
[22] Y. Guo and C. C. Chow, *Existence and stability of standing pulses in neural networks:* I. *Existence*, SIAM J. Appl. Dyn. Syst., 4 (2005), pp. 217–248.
[23] Y. Guo and C. C. Chow, *Existence and stability of standing pulses in neural networks:* II. *Stability*, SIAM J. Appl. Dyn. Syst., 4 (2005), pp. 249–281.
[24] A. Hutt and F. M. Atay, *Effects of distributed transmission speeds on propagating activity in neural populations*, Phys. Rev. E, 73 (2006), 021906.
[25] A. Hutt, *Effects of nonlocal feedback on traveling fronts in neural fields subject to transmission delay*, Phys. Rev. E, 70 (2004), 052902.
[26] A. Hodgkin and A. Huxley, *A quantitative description of membrane current and its application to conduction and excitation in nerve*, J. Physiol., 117 (1952), pp. 500–544.
[27] C. K. R. T. Jones, *Stability of the travelling wave solutions of the Fitzhugh-Nagumo system*, Trans. Amer. Math. Soc., 286 (1984), pp. 431–469.
[28] T. Kapitula, N. Kutz, and B. Sandstede, *The Evans function for nonlocal equations*, Indiana Univ. Math. J., 53 (2004), pp. 1095–1126.
[29] D. J. Pinto and G. B. Ermentrout, *Spatially structured activity in synaptically coupled neuronal networks:* I. *Traveling fronts and pulses*, SIAM J. Appl. Math., 62 (2001), pp. 206–225.
[30] D. J. Pinto and G. B. Ermentrout, *Spatially structured activity in synaptically coupled neuronal networks:* II. *Lateral inhibition and standing pulses*, SIAM J. Appl. Math., 62 (2001), pp. 226–243.
[31] D. J. Pinto, R. K. Jackson, and C. E. Wayne, *Existence and stability of traveling pulses in a continuous neuronal network*, SIAM J. Appl. Dyn. Syst., 4 (2005), pp. 954–984.
[32] K. A. Richardson, S. J. Schiff, and B. J. Gluckman, *Control of traveling waves in the mammalian cortex*, Phys. Rev. Lett., 94 (2005), 028103.
[33] W. Rudin, *Functional Analysis*, International Ser. Pure Appl. Math., McGraw–Hill, New York, 1991.
[34] B. Sandstede, *Evans functions and nonlinear stability of travelling waves in neuronal network models*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., (2007), to appear.
[35] D. H. Sattinger, *On the stability of waves of nonlinear parabolic systems*, Adv. Math., 22 (1976), pp. 312–355.
[36] D. H. Terman, G. B. Ermentrout, and A. C. Yew, *Propagating activity patterns in thalamic neuronal networks*, SIAM J. Appl. Math., 61 (2001), pp. 1578–1604.
[37] N. A. Venkov, S. Coombes, and P. C. Matthews, *Dynamic instabilities in scalar neural field equations with space-dependent delays*, Phys. D, (2006), submitted.
[38] H. R. Wilson and J. D. Cowan, *Excitatory and inhibitory interactions in localized populations of model neurons*, Biophys. J., 12 (1972), pp. 1–24.
[39] D.-X. Xia, Z.-R. Wu, S.-Z. Yan, and W.-C. Shu, *Real Analysis and Functional Analysis*, Advanced Educational Publishing House, Beijing, 1988.
[40] L. Zhang, *On stability of traveling wave solutions in synaptically coupled neuronal networks*, Differential Integral Equations, 16 (2003), pp. 513–536.
[41] L. Zhang, *Traveling waves of a singularly perturbed system of integral-differential equations arising from neuronal networks*, J. Dynam. Differential Equations, 17 (2005), pp. 489–522.
[42] L. Zhang, *Dynamics of neuronal waves*, Math. Z., 255 (2007), pp. 283–321.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 3, pp. 645–662

Critical Point Analysis of Transonic Flow Profiles with Heat Conduction*

H. De Sterck†

Abstract. Critical points arising in transonic flow profiles of the steady compressible Euler equations with heat conduction are investigated. Two examples of transonic flow profiles are studied, namely, radial outflow in a gravitational field, and flow in a quasi-one-dimensional nozzle. The stationary flow equations are reformulated in terms of a dynamical system that parametrizes the flow profiles. Adding the heat conduction term introduces a critical point that is of a type different from the well-known sonic critical point that occurs at the transition from subsonic to supersonic flow when there is no heat conduction. This thermal critical point takes over the saddle-point role of the sonic critical point in the flow profile. Both the sonic and the thermal critical points are present in radial outflow profiles, and the type of the sonic critical point is changed from a saddle point to a simple node by the addition of the heat conduction term. In the nozzle case, the sonic point is no longer a critical point of the dynamical system when heat conduction is added. It is illustrated how the results of this analysis can be used for efficient and accurate numerical calculation of transonic flow trajectories and boundary value problems with heat conduction that are of interest in applications like supersonic planetary escape and solar wind models, and in aerospace applications. The analysis also elucidates how many boundary conditions are required for a well-posed transonic boundary value problem with heat conduction and clarifies the mathematical structure of transonic flow profiles with heat conduction that have been calculated numerically in the literature for various applications.

Key words. compressible gas dynamics, transonic flow, heat conduction, critical point

AMS subject classifications. 76H05, 35B38, 65P99

DOI. 10.1137/060677458

1. Introduction.

It is well known that stationary solutions of the compressible Euler equations of gas dynamics may exhibit critical points where the gas makes a transition from the subsonic to the supersonic flow regime [1, 2]. In this paper we investigate the critical points that arise in one-dimensional (1D) transonic flow profiles when heat conduction is added to the Euler equations.

The purpose of this analysis is two-fold.

First, the critical point analysis leads to a better understanding of the mathematical structure of transonic flow profiles with heat conduction that have been calculated numerically in the literature for various applications—for example, supersonic gas escape from planets like Venus, Pluto, and the early Earth [3, 4, 5, 6], and extrasolar planets [7]. It also elucidates how many boundary conditions are required for a well-posed transonic boundary value problem (BVP) with heat conduction.

Second, the results of this critical point analysis can be used for efficient and accurate

*Received by the editors December 11, 2006; accepted for publication (in revised form) by D. Barkley June 20, 2007; published electronically September 28, 2007.
http://www.siam.org/journals/siads/6-3/67745.html
†Department of Applied Mathematics, University of Waterloo, 200 University Avenue West, Waterloo, ON, N2L 3G1, Canada (hdesterck@uwaterloo.ca).

numerical calculation of transonic flow trajectories and BVPs with heat conduction that are of interest in applications like supersonic planetary escape and solar wind models, and in aerospace applications.

In the context of 1D flow profiles, critical points are points at which the ordinary differential equation (ODE) system that governs the stationary flow is singular; i.e., the derivatives of some of the flow variables cannot be determined from the ODE system at the critical points. The stationary flow equations can be reformulated in terms of a dynamical system that parametrizes the flow profiles. Critical (i.e., singular) points of the flow profiles are also critical points of this associated dynamical system in accordance with the standard definition of a critical point in a dynamical system; i.e., they are equilibrium solutions of the dynamical system.

In this paper, two examples of 1D transonic flow profiles are studied, namely, radial outflow in a gravitational field [8], and flow in a quasi-1D nozzle [9]. The properties of the critical points that arise in these kinds of flows are well known for the Euler model without heat conduction. Here, we investigate the critical points that arise when heat conduction is added to the Euler equations.

In the first part of the paper we study such transonic flow solutions for the case of radial spherically symmetric outflow from spherical objects with a gravitational field. The nozzle application is discussed subsequently.

Figure 1 shows a stationary transonic radial flow solution of the spherically symmetrical Euler equations with radial velocity and heat conduction in a gravitational field, which are given by

$$\frac{\partial}{\partial t}\begin{bmatrix} \rho\, r^2 \\ \rho\, u\, r^2 \\ \left(\frac{p}{\gamma-1}+\frac{\rho\, u^2}{2}\right) r^2 \end{bmatrix} + \frac{\partial}{\partial r}\begin{bmatrix} \rho\, u\, r^2 \\ \rho\, u^2\, r^2 + p\, r^2 \\ \left(\frac{\gamma\, p}{\gamma-1}+\frac{\rho\, u^2}{2}\right) u\, r^2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\rho\, GM + 2\, p\, r \\ -\rho\, GM\, u + q_{heat}\, r^2 + \frac{\partial}{\partial r}\left(\kappa\, r^2\, \frac{\partial T}{\partial r}\right) \end{bmatrix}. \tag{1.1}$$

Here, ρ is the fluid mass density, p is the pressure, u is the radial velocity, r is the radial coordinate, t is time, $-\rho\, GM/r^2$ is the gravitational force density, and $\gamma > 1$ is the adiabatic constant. With G being the universal gravity constant and M the mass of the outflow object, we take $GM = 15$ in normalized units for all the test problems in this paper. The right-hand side of the equation contains a heat source term q_{heat} and a heat conduction term with conductivity κ. This set of equations is used as a model for supersonic gas escape from Venus, Pluto, the early Earth, and extrasolar planets [3, 4, 5, 6, 7], and is related to models for the solar wind [8, 10, 11].

The temperature T, sound speed c, entropy S, and radial mass flux F are defined by

$$T = \frac{p}{\rho}, \tag{1.2}$$

$$c^2 = \frac{\gamma\, p}{\rho} = \gamma\, T, \tag{1.3}$$

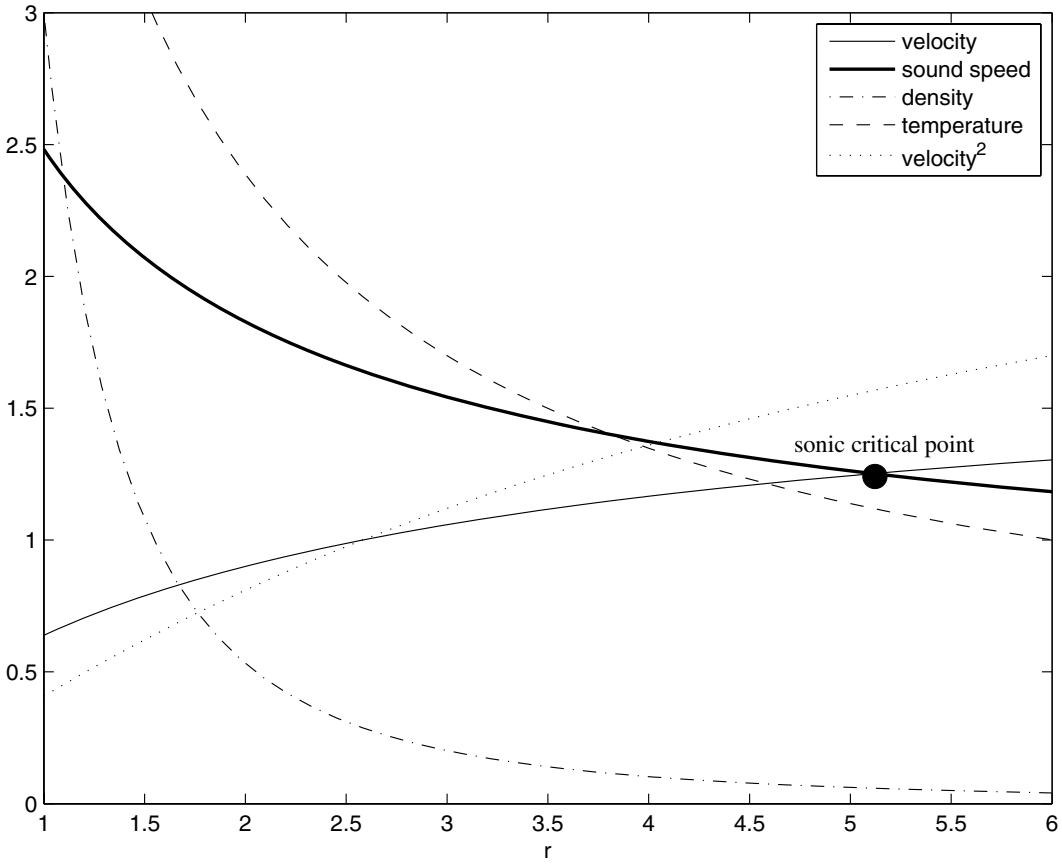


Figure 1. *Radial transonic flow solution for $GM = 15$, $q_{heat}(r) \equiv 0$, and $\kappa(T) = (T/5)^{0.72}$, and with boundary conditions $\rho = 3$, $T = 4.4$, and $dT/dr = -4$ at the inflow boundary $r_a = 1$. The sonic critical point is indicated.*

$$S = \frac{p}{\rho^\gamma}, \tag{1.4}$$

$$F = \rho\, u\, r^2 \tag{1.5}$$

in our choice of nondimensional units. In general, the heat source q_{heat} and heat conductivity κ can be functions of the radius r and/or the gas state variables ρ, p, and u. In particular, we consider radial dependence of the heat source $q_{heat}(r)$ and temperature dependence of the heat conductivity $\kappa(T)$, which are assumptions that are physically relevant for the applications we target [3, 4, 5, 6, 7]. Figure 1 shows a stationary transonic solution to (1.2) with the radial outflow velocity from the planet starting out subsonically near the left boundary. The critical point at the transition to supersonic flow is indicated.

Stationary transonic solutions to the Euler equations are notoriously hard to compute numerically. One of the goals of our critical point analysis is to provide results that allow us to numerically calculate radial transonic flows with heat conduction efficiently and accurately. Traditional techniques employ time-marching strategies of the time-dependent equations [12,

13]. However, explicit time-marching approaches tend to converge very slowly, while implicit techniques are expensive per iteration and may suffer from robustness problems. On the other hand, solving the stationary equations directly has proved difficult due to the presence of critical points of saddle-point type and the associated difficulties with the changing nature of the solution at the critical points, and with the number of boundary conditions to be imposed. In [14] we have proposed a new numerical algorithm for efficiently and accurately calculating radial transonic flows for the Euler equations, which is based on a dynamical systems approach. The results from the present paper allow us to apply these methods to the case where heat conduction is added to the Euler equations.

This paper is structured as follows. In the next section we recall results on critical points for radial Euler flows in a gravitational field without heat conduction, followed by a critical point analysis for the case when heat conduction is added to the Euler model. In section 3 it is shown how the analysis can be extended to the case of a quasi-1D converging-diverging nozzle. Section 4 illustrates how the results on critical points can be utilized for efficient direct calculation of transonic flow trajectories and BVPs for the case of radial outflow in a gravitational field. Conclusions are formulated in section 5.

2. Critical point analysis of radial outflow in a gravitational field. The derivations to be presented in this section start from rewriting the stationary Euler equations in spherical symmetry,

$$\frac{d}{dr}\begin{bmatrix} \rho\, u\, r^2 \\ \rho\, u^2\, r^2 + p\, r^2 \\ \left(\frac{\gamma p}{\gamma-1} + \frac{\rho u^2}{2}\right) u\, r^2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\rho\, GM + 2\, p\, r \\ -\rho\, GM\, u + q_{heat}\, r^2 + \frac{\partial}{\partial r}\left(\kappa\, r^2\, \frac{\partial T}{\partial r}\right) \end{bmatrix}, \tag{2.1}$$

as an ODE system with unknowns $F(r)$, $u(r)$, and $T(r)$ and with decoupled first derivatives as follows:

$$\begin{aligned} \frac{dF}{dr} &= 0, \\ \frac{du}{dr} &= \frac{1}{r^2(u^2-c^2)}\left(2uc^2\left(r - \frac{GM}{2c^2}\right) - (\gamma-1)\frac{q_{heat} r^4 u}{F}\right. \\ &\quad \left. -(\gamma-1)\frac{r^2 u}{F}\frac{d}{dr}\left(\kappa\, r^2\, \frac{dT}{dr}\right)\right), \\ \frac{dT}{dr} &= \frac{\gamma-1}{r^2(u^2-c^2)}\left(T(GM - 2ur^2) - \frac{q_{heat} r^4 (T-u^2)}{F}\right. \\ &\quad \left. -\frac{r^2(T-u^2)}{F}\frac{d}{dr}\left(\kappa\, r^2\, \frac{dT}{dr}\right)\right). \end{aligned} \tag{2.2}$$

It should be noted that the derivations to be presented can also be performed using other state variables—for instance, the density ρ or the entropy S. The choice of variables somewhat influences the effort required to obtain the results and the simplicity of some of the resulting expressions, but in the end the results and conclusions are the same regardless of the choice of variables. The variables in system (2.2) were chosen because they are of direct interest in the applications targeted.

2.1. Euler equations without heat conduction. We first recall results on critical points and their types for the case of vanishing heat conduction $\kappa(T) \equiv 0$. In this case, system (2.2) reduces to a decoupled first-order ODE system, which can be rewritten as an autonomous system by introducing the independent variable s that parametrizes the solution trajectories. This leads to the dynamical system

$$\begin{aligned}
\frac{dr}{ds} &= r^2\,(u^2 - c^2),\\
\frac{dF}{ds} &= 0,\\
\frac{du}{ds} &= 2\,u\,c^2\,\left(r - \frac{GM}{2c^2}\right) - (\gamma - 1)\,q_{heat}\frac{r^4\,u}{F},\\
\frac{dT}{ds} &= (\gamma - 1)\,T\,(GM - 2\,u^2\,r) - (\gamma - 1)\,q_{heat}\frac{r^4}{F}\,(T - u^2).
\end{aligned} \tag{2.3}$$

In general terms, the autonomous system is denoted by

$$\frac{d\mathbf{V}}{ds} = \mathbf{G}(\mathbf{V}), \tag{2.4}$$

with state vector $\mathbf{V} = [F(s)\ u(s)\ r(s)\ T(s)]^T$. The critical points (or equilibrium points) of dynamical system (2.3) satisfy the two conditions

$$\begin{aligned}
u_{crit} &= \sqrt{\gamma\,T_{crit}} = c_{crit},\\
T_{crit} &= \frac{GM}{2\,\gamma\,r_{crit}} + (\gamma - 1)\,\frac{q_{heat}\,r_{crit}^3}{2\,\gamma\,F_{crit}}.
\end{aligned} \tag{2.5}$$

Condition (2.5) identifies the critical point as the well-known sonic critical point. Note that the dynamical system as formulated here is degenerate, in that the four-dimensional (4D) system allows for a two-dimensional (2D) manifold of critical points, as there are only two conditions to be fulfilled instead of four.

The type of the critical point can be derived by investigating the eigenvalues of the Jacobian matrix

$$\left.\frac{\partial \mathbf{G}}{\partial \mathbf{V}}\right|_{\mathbf{V_{crit}}}. \tag{2.6}$$

For the case of vanishing heating source term $q_{heat}(r) \equiv 0$, we obtain

$$\begin{aligned}
\lambda_1 &= 0,\\
\lambda_2 &= 0,\\
\lambda_3 &= GM\sqrt{(-3\gamma + 5)/2},\\
\lambda_4 &= -GM\sqrt{(-3\gamma + 5)/2}.
\end{aligned} \tag{2.7}$$

When $\gamma < 5/3$, this leads to a (degenerate) saddle point. For example, for the outflow problem of hydrogen gas from planetary atmosphere that we target [3, 4, 5, 6, 7], the value $\gamma = 7/5$ for di-atomic gases applies. All the example flow solutions presented in this paper use the value $\gamma = 7/5$. For general heat source functions $q_{heat}(r)$, the expressions for Jacobian (2.6) and its eigenvalues are complicated and the detailed expressions are lengthy and uninsightful. However, using symbolical manipulation software, it can easily be verified that the eigenvalues satisfy

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= 0, \\ \lambda_3 &= a + \sqrt{b}, \\ \lambda_4 &= a - \sqrt{b}. \end{aligned} \tag{2.8}$$

Here, a and b are generic placeholders for nonlinear functions that involve the state variables, parameters like γ and GM, and the heat source function $q_{heat}(r)$ and its radial derivative $dq_{heat}(r)/dr$. Depending on the values of these, the type of the critical point changes. Again, the general conditions that determine the type of the critical point are lengthy and do not provide specific insight. It suffices to say that in many parameter regimes of interest for the applications we target, the critical point turns out to be of (degenerate) saddle-point type.

Figure 2 shows an example of a stationary transonic solution to (1.2) with $\kappa(T) \equiv 0$ and $q_{heat} \equiv 0$. The critical point at the transition to supersonic flow is indicated. In accordance with the discussion above, the critical point turns out to be of saddle-point type.

2.2. Euler equations with heat conduction. We now proceed with the main topic of this paper, namely, the critical point analysis of radial transonic Euler flow with heat conduction, i.e., the case where $\kappa(T) \neq 0$. We start from the ODE system (2.2) and define the additional variable

$$\phi = \kappa\, r^2 \frac{dT}{dr}, \tag{2.9}$$

which is related to the temperature gradient. With the use of this gradient variable ϕ we can rewrite ODE system (2.2) as the fully decoupled first-order system

$$\begin{aligned} \frac{dF}{dr} &= 0, \\ \frac{du}{dr} &= \frac{1}{r^2(u^2-c^2)}\left(2uc^2\left(r - \frac{GM}{2c^2}\right) + \frac{u\phi(u^2-c^2)}{\kappa(T-u^2)}\right. \\ &\qquad \left. -\frac{(\gamma-1)uT(GM-2u^2r)}{T-u^2}\right), \\ \frac{dT}{dr} &= \frac{\phi}{\kappa r^2}, \\ \frac{d\phi}{dr} &= \frac{-\phi F(u^2-c^2)}{(\gamma-1)\kappa r^2(T-u^2)} + \frac{FT(GM-2u^2r)}{r^2(T-u^2)} - q_{heat}r^2. \end{aligned} \tag{2.10}$$

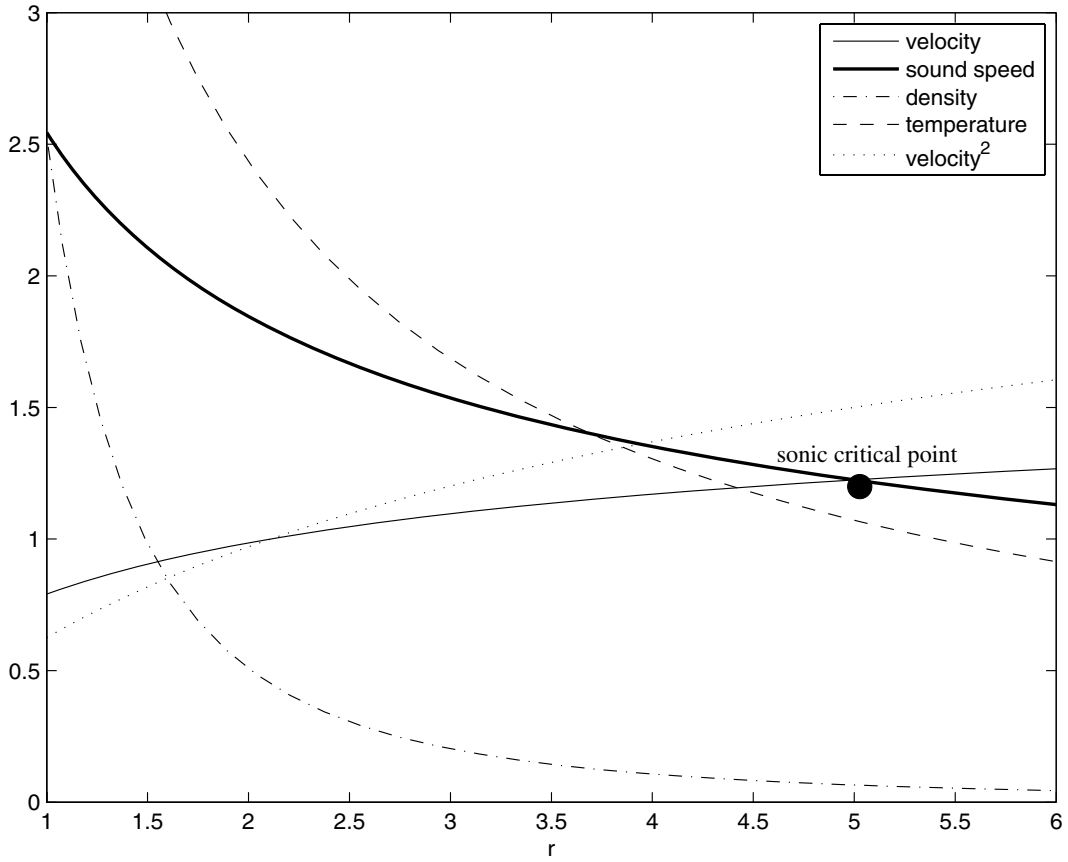


Figure 2. *Radial transonic flow solution for $q_{heat}(r) \equiv 0$ and $\kappa(T) \equiv 0$. The critical radius $r_{crit} = 5$, and the flux $F = 2$. The sonic critical point is indicated.*

This system can be written as an autonomous system by introducing the independent variable s that parametrizes the solution trajectories, resulting in the dynamical system

$$
\begin{aligned}
\frac{dr}{ds} &= -r^2(u^2 - c^2)(u^2 - T), \\
\frac{dF}{ds} &= 0, \\
\frac{du}{ds} &= -2uc^2\left(r - \frac{GM}{2c^2}\right)(u^2 - T) + \frac{\phi u(u^2 - c^2)}{\kappa} \\
&\quad -(\gamma - 1)uT(GM - 2u^2 r), \\
\frac{dT}{ds} &= \frac{-\phi(u^2 - c^2)(u^2 - T)}{\kappa}, \\
\frac{d\phi}{ds} &= \frac{-\phi F(u^2 - c^2)^2}{(\gamma - 1)\kappa} + FT(GM - 2u^2 r)(u^2 - c^2) \\
&\quad +q_{heat} r^4(u^2 - c^2)(u^2 - T).
\end{aligned}
\tag{2.11}
$$

Interestingly, the dynamical system now turns out to have two types of critical points.

The first type of critical point is the well-known sonic critical point. This critical point is four-fold degenerate, as there is now only one condition at this critical point:

$$u_{crit} = \sqrt{\gamma\, T_{crit}} = c_{crit}. \tag{2.12}$$

The eigenvalues of the Jacobian are given by

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= 0, \\ \lambda_3 &= 0, \\ \lambda_4 &= 0, \\ \lambda_5 &= u_{crit}^2 \left(\frac{(\gamma+1)\phi_{crit}}{\kappa(T_{crit})} + 2GM - \frac{4u_{crit}^2 r_{crit}}{\gamma} \right), \end{aligned} \tag{2.13}$$

indicating a critical point of (degenerate) simple node type.

The second type of critical point is characterized by the conditions

$$\begin{aligned} u_{crit} &= \sqrt{T_{crit}} = c_{crit}/\sqrt{\gamma}, \\ \frac{\phi_{crit}}{\kappa} + GM - 2u_{crit}^2 r_{crit} &= 0. \end{aligned} \tag{2.14}$$

We call this type of critical point a thermal critical point ($u_{crit}^2 = T_{crit}$), in order to distinguish it from the sonic critical point type, where $u_{crit}^2 = \gamma\, T_{crit}$. Note that the dynamical system is again degenerate, in that the five-dimensional (5D) system allows for a three-dimensional (3D) manifold of critical points, as there are only two conditions to be fulfilled instead of five.

For this critical point, the eigenvalues of the Jacobian are given by

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= 0, \\ \lambda_3 &= 0, \\ \lambda_4 &= a + \sqrt{b}, \\ \lambda_5 &= a - \sqrt{b}, \end{aligned} \tag{2.15}$$

where a and b are generic placeholders for nonlinear functions that involve the state variables, parameters like γ and GM, the heat source function $q_{heat}(r)$ and its radial derivative $dq_{heat}(r)/dr$, and the heat conduction function $\kappa(T)$ and its temperature derivative $d\kappa(T)/dT$. Full expressions for a and b can be derived using symbolical manipulation software, but they are cumbersome and do not lead to further insight. Again, depending on the values of the parameters, the type of the critical point changes. As before, the general conditions that determine the type of the critical point are lengthy and do not provide specific insight, and it suffices to say that in many parameter regimes of interest for the applications we target, the critical point turns out to be of (degenerate) saddle-point type.

We can thus conclude that, interestingly, adding the heat conduction term to the Euler model changes the type of the sonic critical point from a saddle point to a simple node and

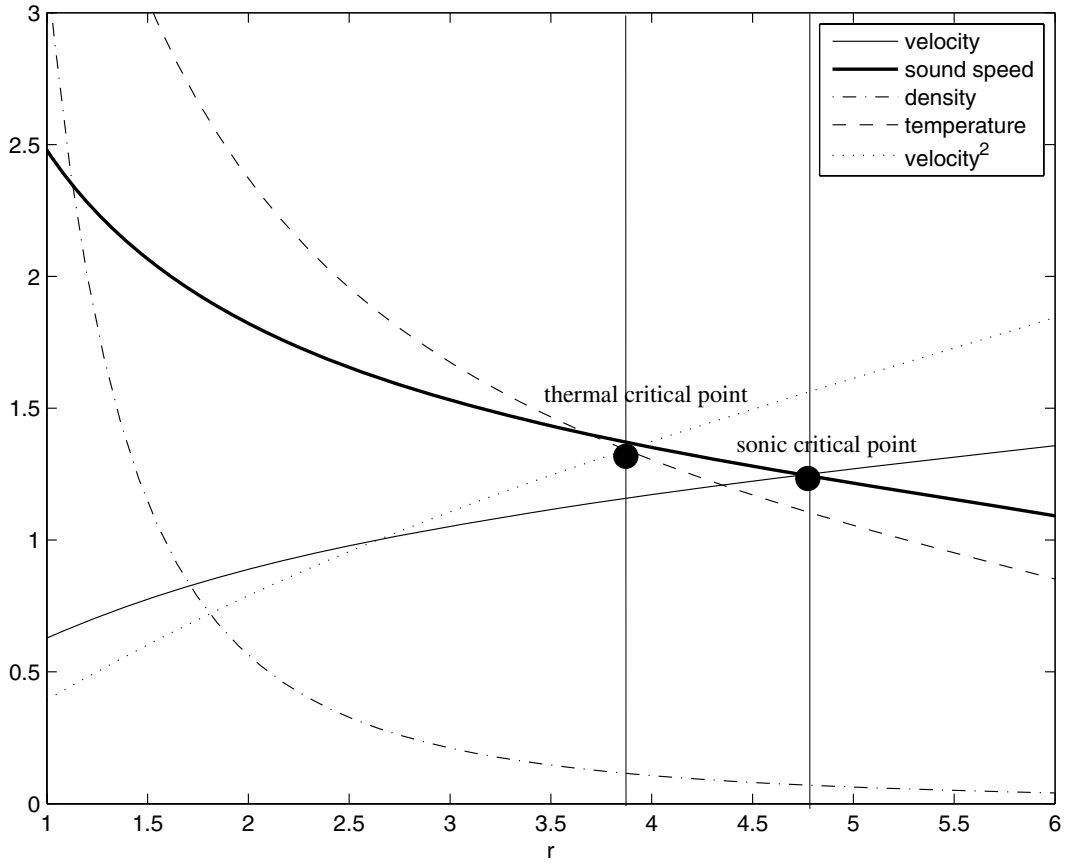


Figure 3. *Radial transonic flow solution for $q_{heat}(r) \equiv 0$ and $\kappa(T) \equiv (T/5)^{0.72}$. The two critical points are indicated. The rightmost critical point is the well-known sonic critical point, while the leftmost critical point is the thermal critical point that is introduced by including the heat conduction term. The location of the thermal critical point $r_{crit} \approx 3.87$, its temperature $T \approx 1.34$, and its velocity $u \approx 1.16$.*

introduces an additional, thermal, critical point into the system, which is of saddle-point type. It turns out that transonic solutions pass through the two types of critical points, as illustrated by the following example flow solution.

Figure 3 shows an example of a stationary transonic solution to (1.2) with $\kappa(T) = (T/5)^{0.72}$, and $q_{heat} = 0$. The exponent 0.72 for the heat conduction function is chosen for this example flow solution due to its relevance for planetary atmosphere calculations [3, 4, 5, 6, 7]. The two critical points are indicated on the figure. The leftmost, thermal, critical point occurs where $u = \sqrt{T}$ and is of saddle-point type. The rightmost critical point is the sonic critical point, where $u = \sqrt{\gamma T}$.

The flow profile of Figure 3 thus has two critical points, but they are of different type (one is thermal, and the other is sonic). Note that this situation is different from the multitransonic accretion profiles studied in [15], in which multiple critical points occur that are all of sonic type for flow equations without heat conduction.

3. Critical point analysis of quasi-1D nozzle flow. In this section we repeat the analysis presented above for the case of a quasi-1D converging-diverging nozzle [9], leading to similar results. Figure 4 shows a de Laval nozzle, in which subsonic flow can be accelerated to supersonic flow in a continuous fashion. If the cross-sectional flow area $A(x)$ varies slowly, the nozzle is long and slender, and cross-flow velocities are small, the flow can be modeled as quasi-1D (see, e.g., [9]), leading to the following equation system:

$$\frac{\partial}{\partial t}\begin{bmatrix} \rho A \\ \rho u A \\ \left(\frac{p}{\gamma-1}+\frac{\rho u^2}{2}\right) A \end{bmatrix} + \frac{\partial}{\partial x}\begin{bmatrix} \rho u A \\ \rho u^2 A + p A \\ \left(\frac{\gamma p}{\gamma-1}+\frac{\rho u^2}{2}\right) u A \end{bmatrix} = \begin{bmatrix} 0 \\ p\frac{dA}{dx} \\ \frac{\partial}{\partial x}\left(\kappa A \frac{\partial T}{\partial x}\right) \end{bmatrix}. \tag{3.1}$$

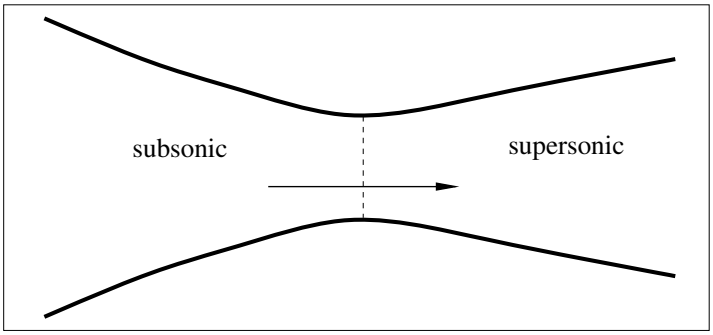


Figure 4. *Transonic flow in a converging-diverging nozzle.*

3.1. Nozzle flow without heat conduction. First, consider the case that the heat conduction coefficient, $\kappa(T)$, vanishes. Then the dynamical system associated with the stationary part of (3.2) can be derived as above, leading to

$$\begin{aligned} \frac{dx}{ds} &= u^2 - c^2, \\ \frac{dF}{ds} &= 0, \\ \frac{du}{ds} &= \frac{\gamma u T}{A}\frac{dA}{dx}, \\ \frac{dT}{ds} &= -\frac{(\gamma-1)u^2 T}{A}\frac{dA}{dx}. \end{aligned} \tag{3.2}$$

The critical points of dynamical system (3.2) satisfy the two conditions

$$\begin{aligned} u_{crit} &= \sqrt{\gamma T_{crit}} = c_{crit}, \\ \frac{dA}{dx}(x_{crit}) &= 0. \end{aligned} \tag{3.3}$$

This dynamical system has, thus, a sonic critical point, which can occur only at the throat of the nozzle, where $dA(x)/dx = 0$. The type of the critical point can be determined by investigating the eigenvalues of the Jacobian matrix, which are given by

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= 0, \\ \lambda_3 &= \sqrt{\frac{\gamma+1}{A}\frac{d^2A}{dx^2}}\,u^2, \\ \lambda_4 &= -\sqrt{\frac{\gamma+1}{A}\frac{d^2A}{dx^2}}\,u^2. \end{aligned} \tag{3.4}$$

This shows that the sonic critical point is a saddle point at the throat of the nozzle (where $d^2A(x)/dx^2 \geq 0$).

3.2. Nozzle flow with heat conduction. When heat conduction is added to the Euler model ($\kappa(T) \neq 0$), we can define the additional variable

$$\phi = \kappa\,\frac{dT}{dx}. \tag{3.5}$$

With the use of this gradient variable ϕ, we can, as before, derive the dynamical system

$$\begin{aligned} \frac{dx}{ds} &= u^2 - T, \\ \frac{dF}{ds} &= 0, \\ \frac{du}{ds} &= \frac{u\,T}{A}\frac{dA}{dx} - \frac{u\,\phi}{\kappa}, \\ \frac{dT}{ds} &= \frac{\phi}{\kappa}\,(u^2 - T), \\ \frac{d\phi}{ds} &= \left(\frac{F\,u^2\,T}{A^2} - \frac{\phi}{A}\,(u^2 - T)\right)\frac{dA}{dx} + \frac{F\,\phi}{A\,\kappa}\,\frac{u^2 - \gamma\,T}{\gamma - 1}. \end{aligned} \tag{3.6}$$

Contrary to the radial outflow case with heat conduction, this dynamical system has only one type of critical point. The critical point is characterized by the conditions

$$\begin{aligned} u_{crit} &= \sqrt{T_{crit}}, \\ \frac{1}{A(x_{crit})}\frac{dA}{dx}(x_{crit}) &= \frac{\phi_{crit}}{\kappa\,T_{crit}} = \frac{1}{T_{crit}}\frac{dT}{dx}(x_{crit}). \end{aligned} \tag{3.7}$$

This critical point is a thermal critical point ($u_{crit}^2 = T_{crit}$). It is interesting to note that the thermal critical point does not occur at the throat of the nozzle. For this critical point, the

eigenvalues of the Jacobian are given by

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= 0, \\ \lambda_3 &= 0, \\ \lambda_4 &= a + \sqrt{b}, \\ \lambda_5 &= a - \sqrt{b}, \end{aligned} \tag{3.8}$$

where a and b are generic placeholders for nonlinear functions that involve the state variables, the parameter γ, the area function $A(x)$ and its first and second spatial derivatives, and the heat conduction function $\kappa(T)$ and its temperature derivative $d\kappa(T)/dT$. Again, full expressions for a and b can be derived using symbolical manipulation software, but they are cumbersome and do not lead to further insight. Also, depending on the values of the parameters, the type of the critical point changes. As before, the general conditions that determine the type of the critical point are lengthy and do not provide specific insight, and it suffices again to say that in many parameter regimes of interest, the critical point turns out to be of (degenerate) saddle-point type.

We can thus conclude that adding the heat conduction term to the Euler model here removes the sonic critical point as a critical point altogether and introduces a new thermal critical point of saddle-point type.

The result is thus similar to the radial outflow case: adding heat conduction to the Euler equations results in the appearance of a thermal critical point. The thermal critical point takes over the role of the sonic critical point as a saddle point in the flow profile.

4. Numerically calculated example flow solutions for radial outflow in a gravitational field. In this section, we illustrate how the critical point properties derived above can be used for efficiently calculating stationary transonic Euler flows numerically for the case of radial outflow in a gravitational field. Note that example flow solutions for the quasi-1D nozzle application can be calculated in a similar manner.

We consider two typical situations. In the first, most simple, case, the state variables at the critical point are known, and a numerical approximation to the critical trajectory through the critical point is sought. In the second case, some state variables are known at the left inflow boundary, and a numerical approximation to the transonic solution that matches the boundary conditions is sought. Such a critical BVP is significantly more complicated to solve, but here too the knowledge about the critical points that was developed aids in efficiently obtaining a numerical approximation. This section also serves to illustrate how the addition of heat conduction influences transonic Euler solutions.

4.1. Example trajectories. We first recall why trajectories that pass through critical points of saddle-point type are difficult to approximate numerically. This can be explained most easily for the simplified case of isothermal Euler flow [8, 11, 14]. In this case the sound speed c is a constant, and a decoupled ODE for $u(r)$ can be derived as

$$\frac{du}{dr} = \frac{2\,u\,c^2\,\left(r - \frac{GM}{2c^2}\right)}{r^2\,(u^2 - c^2)}, \tag{4.1}$$

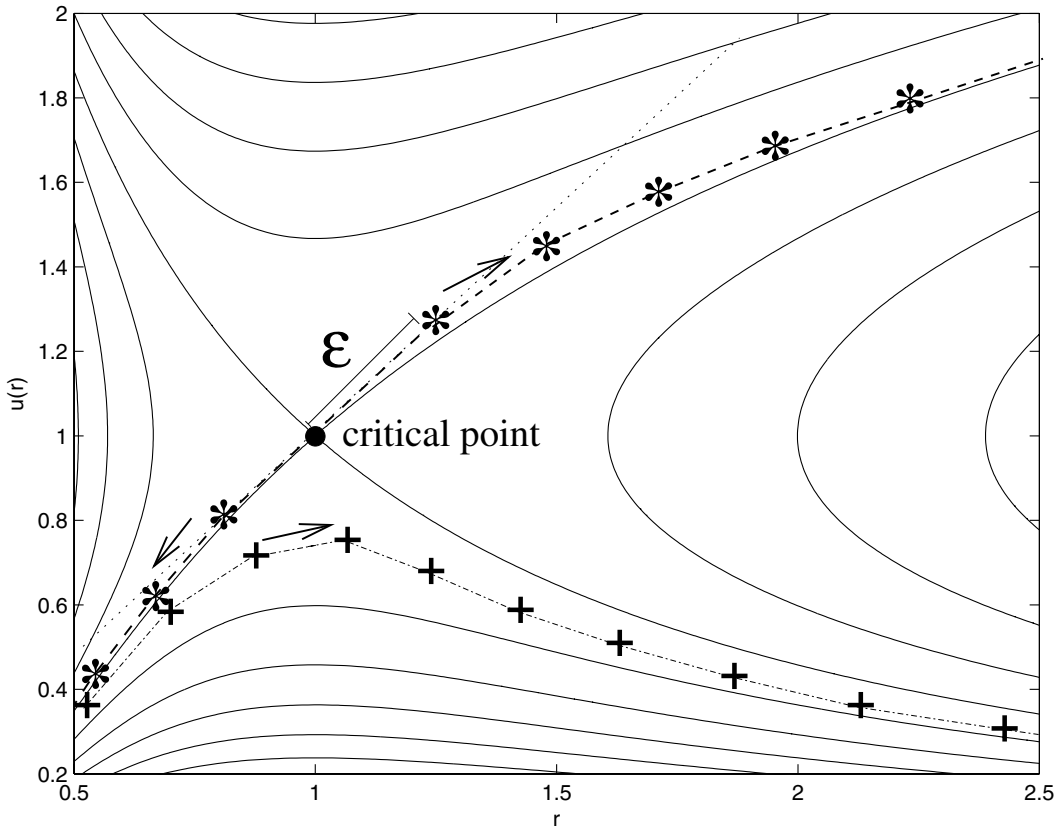


Figure 5. *Solution curves $u(r)$ for the spherically symmetric isothermal Euler equations with $GM = 2$ and $c = 1$. Numerical integration from the left boundary (dash-dotted curve) does not allow approximation of the transonic solution, but integration in two directions outward from the critical point allows numerical approximation of the transonic solution. The arrows indicate the direction of numerical integration.*

which leads to the simplified dynamical system

$$\begin{aligned} \frac{du(s)}{ds} &= -2\,u\,c^2\left(r - \frac{GM}{2c^2}\right), \\ \frac{dr(s)}{ds} &= -r^2\,(u^2 - c^2). \end{aligned} \tag{4.2}$$

This system features a critical point of saddle-point type that is specified by the conditions

$$\begin{aligned} r_{crit} &= \frac{GM}{2c^2}, \\ u_{crit} &= c, \end{aligned} \tag{4.3}$$

or the point $(1, 1)$ in the (r, u) phase plane for the choice of parameters $GM = 2$ and $c = 1$. Figure 5 shows solution trajectories in the (r, u) phase plane.

Assume that one wants to find a numerical approximation for the transonic solution that passes through the critical point. It is easy to see that numerical integration from the left does not lead to a desirable result, as numerical errors cause the approximation to deviate from the transonic curve. However, integration in two directions outward from the critical point (dashed curves) leads to an accurate numerical approximation of the transonic curve. The initial points for the outward integrations should be taken in the direction of the tangent to the transonic solution, which is an eigenvector of the Jacobian matrix $\partial G(V)/\partial V$ at the critical point. This approach for calculating the stable or unstable manifold of a saddle point is a standard dynamical systems technique [16] and can be applied to transonic Euler flow calculation once the critical point location and state are known.

Figure 6 illustrates how this approach can be used for numerical approximation of transonic Euler flow with and without heat conduction. The top panel shows Euler flow without heat conduction, with critical radius $r_{crit} = 5$ and flux $F = 2$. This is the same solution as shown in Figure 2, but now we have added the numerical integration points that were used by our ODE integrator. Numerical integration is outward from the sonic critical point, which is of saddle-point type. We have used an adaptive fourth-fifth-order accurate Runge–Kutta ODE integrator for the result shown, namely, the RK45 Fehlberg method as described, for example, in [17]. This solution has $u^2 = T \approx 1.34$ at $r \approx 3.87$.

The top flow profile has to be compared with the solution presented in the bottom panel, which features heat conduction, and has the same values for temperature and velocity as the top solution at $r \approx 3.87$. This is the same solution as shown in Figure 3 but is repeated here for the purpose of comparison with the top panel. Numerical integration is now outward from the thermal critical point, and it can be seen that the numerical integration does not experience any problem at the sonic critical point, which is now a simple node due to the presence of heat conduction. The top and bottom solutions both feature $u^2 = T \approx 1.34$ at $r \approx 3.87$ but differ otherwise due to the presence of heat conduction in the bottom solution.

4.2. Example BVPs. In Figure 7 we compare two BVPs without and with heat conduction. The top solution has $q_{heat}(r) \equiv 0$ and $\kappa(T) \equiv 0$, while the bottom solution has $q_{heat}(r) \equiv 0$ and $\kappa(T) \equiv (T/5)^{0.72}$. Both flow solutions have $\rho = 3$ and $T = 4.4$ at the inflow boundary. In addition, the bottom solution, with heat conduction, requires one more boundary condition at the inflow boundary in order for the BVP to be well-posed; $dT/dr = -4$ is imposed additionally at the lower boundary. Note the difference in scale of the two plots: the heat conduction brings the location of the sonic critical point much closer to the inflow boundary. The values of the resulting flux variable F, which is the main quantity of interest for this kind of BVP solution in the case of planetary outflow calculations [3, 4, 5, 6, 7], is also influenced significantly by the heat conduction: $F \approx 0.78$ for the BVP without heat conduction, but $F \approx 1.92$ with heat conduction.

Determining the number of boundary conditions necessary for transonic BVPs is not a trivial matter, and numerical calculation of transonic BVPs is difficult due to the a priori unknown location of the critical points. Recently we have proposed a new algorithm for numerical approximation of transonic BVP solutions to the Euler equations. It suffices here to give a brief explanation of how the BVP solutions of Figure 7 were obtained, and full details about the algorithmic approach can be found in [14].

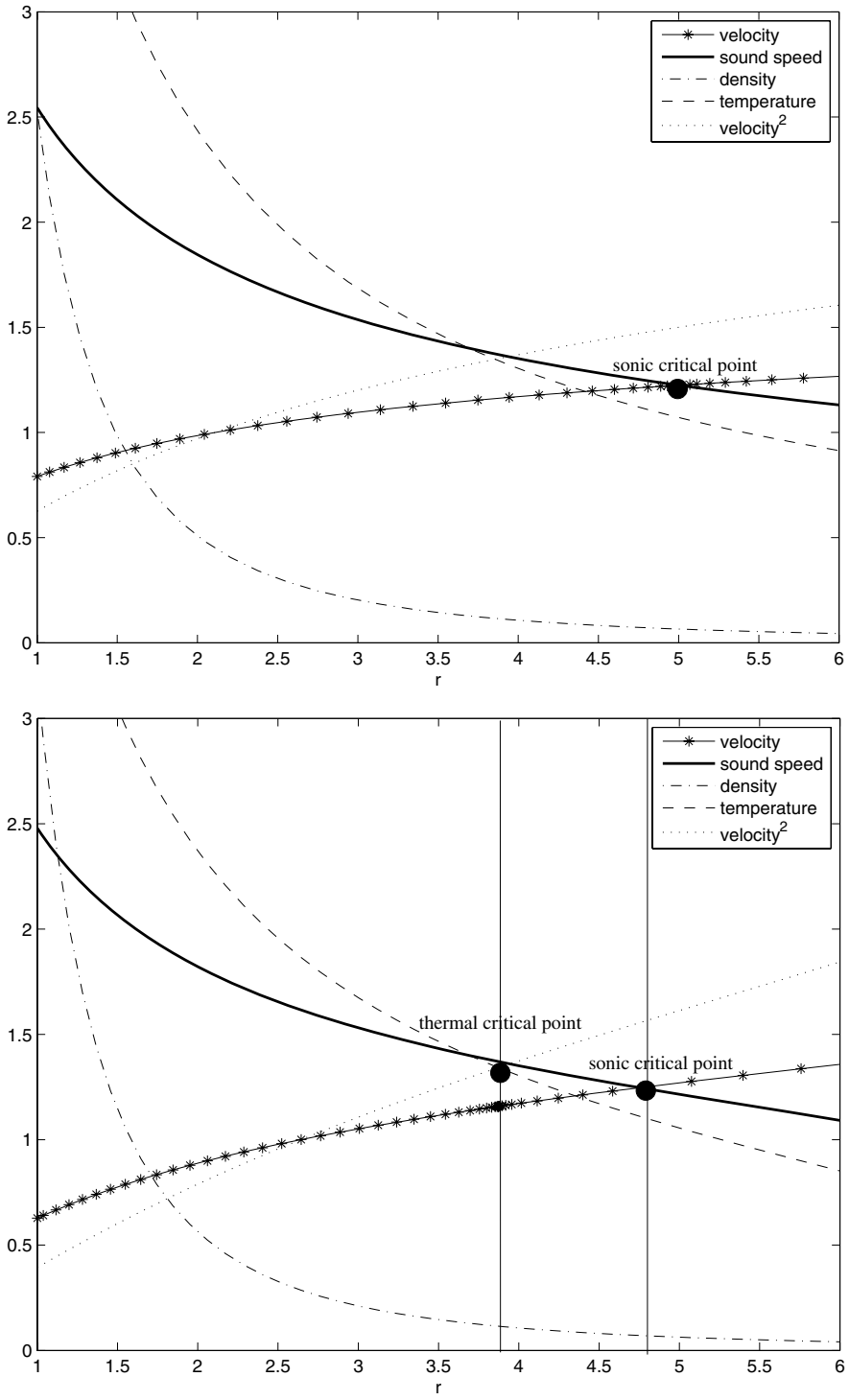


Figure 6. *Radial outflow trajectories without and with heat conduction. Top: Radial transonic flow solution for* $q_{heat}(r) \equiv 0$ *and* $\kappa(T) \equiv 0$. *The sonic critical radius* $r_{crit} = 5$, *and the flux* $F = 2$. *Bottom: Radial transonic flow solution for* $q_{heat}(r) \equiv 0$ *and* $\kappa(T) \equiv (T/5)^{0.72}$. *The thermal critical point has critical radius* $r_{crit} \approx 3.87$, *temperature* $T \approx 1.34$, *and velocity* $u \approx 1.16$.

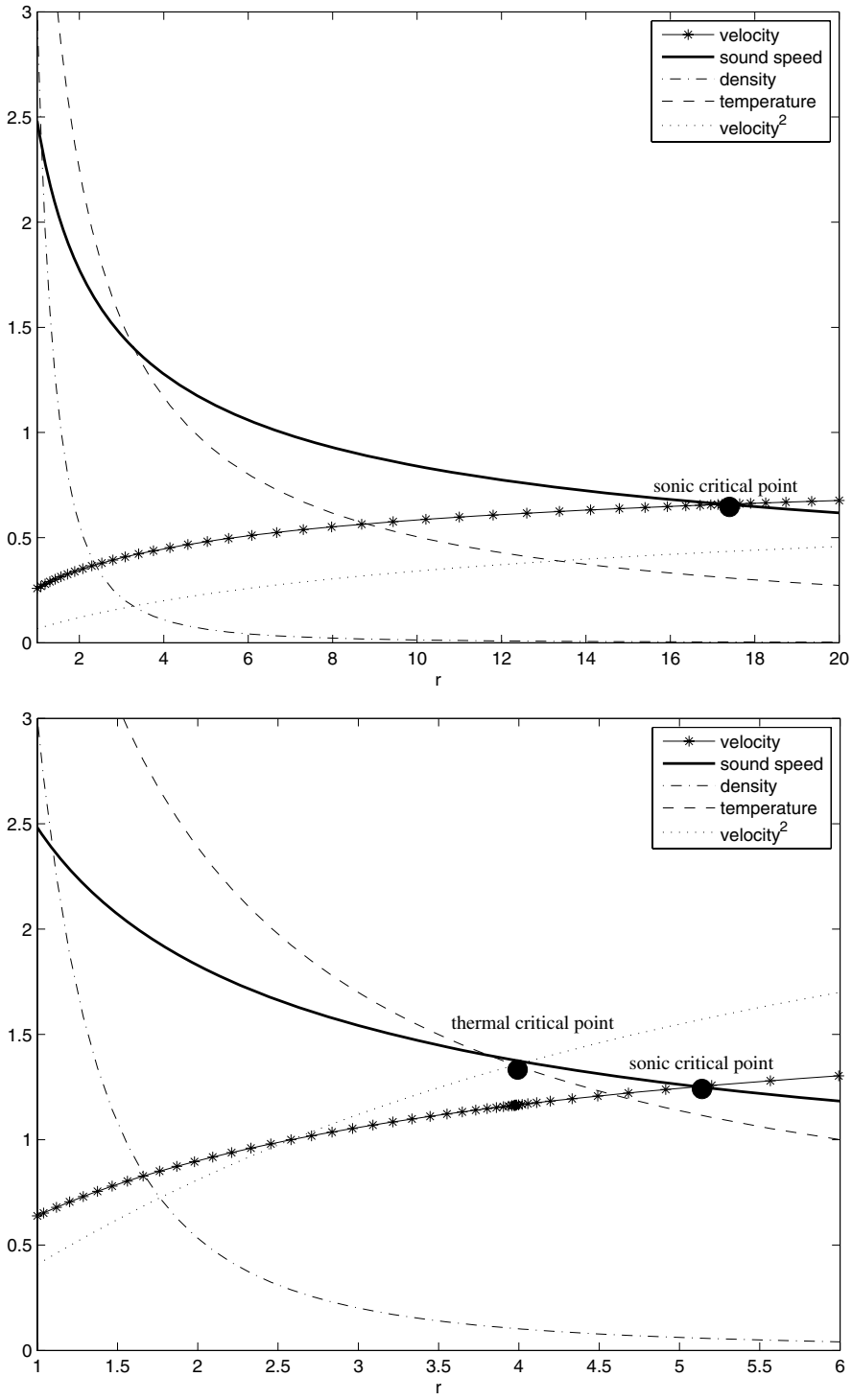


Figure 7. *BVP flow trajectories without (top) and with (bottom) heat conduction. Both flow solutions have* $\rho = 3$ *and* $T = 4.4$ *at the inflow boundary. In addition,* $dT/dr = -4$ *is imposed at the inflow boundary for the bottom solution. The NCP method from* [14] *was used to solve these BVPs numerically.*

Regarding the issue of the number of boundary conditions, the isothermal Euler case is again illustrative. The solution trajectories of (4.1) and (4.2) shown in Figure 5 are uniquely identified by one boundary condition at the left inflow boundary, except for the transonic curve, whose specification does not require any numeric boundary condition value for the velocity u at the left boundary, in the following sense: just the requirement that the solution is transonic, i.e., that the flow is subsonic at the left boundary and supersonic at the right boundary, uniquely determines the solution, and the value of the velocity u at the inflow boundary follows from the transonic flow requirement. In the same way, transonic Euler solutions without heat conduction require only two boundary conditions for 3D system (2.2) (with $\kappa(T) \equiv 0$), in addition to the requirement that the flow solution is transonic, and, similarly, transonic Euler solutions with heat conduction require only three boundary conditions for 4D system (2.10), again in addition to the requirement that the flow solution is transonic.

The BVP solution can then be approximated numerically in the above-described iterative procedure that uses the numerical integration method outward from the critical points, combined with a Newton method that allows the critical point to vary within the manifold of possible critical points, thus driving the iterative approximation toward the particular transonic solution and critical point that match the inflow boundary conditions. See [14] for details about this Newton critical point algorithm. Indeed, in the case of the Euler equations without heat conduction, the two-fold degeneracy of the sonic critical point of saddle-point type (2.5) conveniently matches the number of inflow boundary conditions for the BVP. In a similar way, the results of this paper show that, in the case of the Euler equations with heat conduction, the three-fold degeneracy of the nonsonic critical point of saddle-point type (2.14) matches the number of inflow boundary conditions that specify the transonic BVP solution uniquely.

5. Conclusions. In this paper, we have presented a critical point analysis for transonic flow solutions of the steady compressible Euler equations with heat conduction. By means of reformulating the stationary ODE system as a first-order system that can be rewritten as a dynamical system, it was shown that the addition of the heat conduction term introduces a new critical point that is different from the well-known sonic critical point. This so-called thermal critical point is of saddle-point type in parameter regimes of interest. For radial outflows, the sonic point remains a critical point, and the flow profile thus has two critical points of different type, but the sonic critical point is changed from a saddle point to a simple node by the addition of the heat conduction term. In the quasi-1D nozzle case, the sonic point is not a critical point of the dynamical system anymore when heat conduction is added. Heat conduction adds a term with a second spatial derivative of the temperature, and it can be concluded that adding this heat conduction term changes the critical point properties of the ODE system and the transonic flow solutions significantly.

It was illustrated how the results of this analysis can be used for efficient and accurate numerical calculation of transonic radial outflow trajectories and BVPs using the Newton critical point method that was introduced in [14]. The analysis also elucidates how many boundary conditions are required for a well-posed transonic BVP with heat conduction. The results presented in this paper thus reveal the mathematical structure of stationary transonic solutions of the compressible Euler equations with heat conduction, which are of interest in

applications like supersonic gas escape from Venus, Pluto, and the early Earth [3, 4, 5, 6].

In future work, we will study how the phenomena described in this paper play a role in 2D and 3D flows with heat conduction, where symmetry is relaxed. We expect that the analysis given here can be extended to these higher-dimensional cases, and we speculate that multiple types of critical points may arise there as well. We also intend to investigate whether numerical methods for stationary transonic flow simulation that are based on dynamical systems analysis may be developed for flow problems in multiple spatial dimensions. This will be addressed in future research.

REFERENCES

[1] R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves*, Interscience Publishers, New York, 1948.

[2] R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol.* 2, Interscience Publishers, New York, 1962.

[3] A. J. Watson, T. M. Donahue, and J. C. G. Walker, *The dynamics of a rapidly escaping atmosphere: Applications to the evolution of earth and Venus*, Icarus, 48 (1981), pp. 150–166.

[4] J. F. Kasting and J. B. Pollack, *Loss of water from Venus.* I. *Hydrodynamic escape of hydrogen*, Icarus, 53 (1983), pp. 479–508.

[5] F. Tian and O. B. Toon, *Hydrodynamic escape of nitrogen from Pluto*, Geophys. Res. Lett., 32 (2005), pp. 18201–18204.

[6] F. Tian, O. B. Toon, A. A. Pavlov, and H. De Sterck, *A hydrogen-rich early Earth atmosphere*, Science, 308 (2005), pp. 1014–1017.

[7] F. Tian, O. B. Toon, A. A. Pavlov, and H. De Sterck, *Transonic hydrodynamic escape of hydrogen from extrasolar planetary atmospheres*, Astrophys. J., 621 (2005), pp. 1049–1060.

[8] M. Goossens, *An Introduction to Plasma Astrophysics and Magnetohydrodynamics*, Kluwer, Dordrecht, The Netherlands, 2003.

[9] G. I. Benison and E. L. Rubin, *A time-dependent analysis for quasi-one-dimensional, viscous, heat conducting, compressible Laval nozzle flows*, J. Engineering Mathematics, 5 (1970), pp. 39–49.

[10] R. Keppens and J. P. Goedbloed, *Numerical simulations of stellar winds: Polytropic models*, Astron. Astrophys., 343 (1998), pp. 251–260.

[11] T. Sakurai, *Magnetohydrodynamic solar/stellar wind models*, Computer Physics Reports, 12 (1990), pp. 247–273.

[12] R. J. LeVeque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge University Press, Cambridge, UK, 2002.

[13] B. van Leer, W. T. Lee, and P. L. Roe, *Characteristic Time-Stepping or Local Preconditioning of the Euler Equations*, AIAA paper 91-1552, Reston, VA, 1991.

[14] H. De Sterck and F. Tian, *A fast and accurate algorithm for computing radial transonic flows*, submitted, 2007.

[15] S. Chaudhury, A. K. Ray, and T. K. Das, *Critical properties and stability of stationary solutions in multitransonic pseudo-Schwarzschild accretion*, MNRAS, 373 (2006), pp. 146–156.

[16] B. Krauskopf, H. M. Osinga, E. J. Doedel, M. E. Henderson, J. M. Guckenheimer, A. Vladimirsky, M. Dellnitz, and O. Junge, *A survey of methods for computing (un)stable manifolds of vector fields*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 15 (3) (2005), pp. 763–791.

[17] L. Elden, L. Wittmeyer-Koch, and H. B. Nielsen, *Introduction to Numerical Computation*, Studentlitteratur AB, Lund, Sweden, 2004.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 4, pp. 663–693

When Shil'nikov Meets Hopf in Excitable Systems*

Alan R. Champneys†, Vivien Kirk‡, Edgar Knobloch§, Bart E. Oldeman‡, and James Sneyd‡

Abstract. This paper considers a hierarchy of mathematical models of excitable media in one spatial dimension, specifically the FitzHugh–Nagumo equation and several models of the dynamics of intracellular calcium. A common feature of the models is that they support solitary traveling pulse solutions which lie on a characteristic C-shaped curve of wave speed versus parameter. This C lies to the left of a U-shaped locus of Hopf bifurcations that corresponds to the onset of small-amplitude linear waves. The central question addressed is how the Hopf and solitary wave (homoclinic orbit in a moving frame) bifurcation curves interact in these "CU systems." A variety of possible codimension-two mechanisms is reviewed through which such Hopf and homoclinic bifurcation curves can interact. These include Shil'nikov–Hopf bifurcations and the local birth of homoclinic chaos from a saddle-node/Hopf (Gavrilov–Guckenheimer) point. Alternatively, there may be barriers in phase space that prevent the homoclinic curve from reaching the Hopf bifurcation. For example, the homoclinic orbit may bump into another equilibrium at a so-called T-point, or it may terminate by forming a heteroclinic cycle with a periodic orbit. This paper presents the results of detailed numerical continuation results on different CU systems, thereby illustrating various mechanisms by which Hopf and homoclinic curves interact in CU systems. Owing to a separation of time scales in these systems, considerable care has to be taken with the numerics in order to reveal the true nature of the bifurcation curves observed.

Key words. FitzHugh–Nagumo equation, homoclinic, Hopf, calcium dynamics, excitable media, traveling waves

AMS subject classifications. 34C15, 34C23, 34C37, 35B32, 37G10, 37G15

DOI. 10.1137/070682654

1. Introduction.

So-called *excitable* systems of reaction-diffusion equations are used to model a variety of biophysical processes, including neuronal information processing, heart electro-physiology, and the processes by which cells signal to one another; see, e.g., [17, 28]. One of the early successes in mathematical biology was Hodgkin and Huxley's work on modeling a squid giant nerve axon [27], leading to the equations that bear their name. Later in the 1960's FitzHugh [18] and Nagumo [32] came up with a simplified version, which has since come to be considered the canonical excitable system.

The FitzHugh–Nagumo model is called excitable because, in the absence of diffusion, a small perturbation from the stable equilibrium causes straightforward relaxation back to the equilibrium, whereas a large enough perturbation causes a sudden, large-amplitude burst

*Received by the editors February 14, 2007; accepted for publication (in revised form) by J. Guckenheimer May 2, 2007; published electronically October 5, 2007. This work was supported in part by the New Zealand Institute of Mathematics and Its Applications, the Weierstrass Institute Berlin, the NIH, EPSRC, and by NSF grants DMS-0305968 and DMS-0605238.

http://www.siam.org/journals/siads/6-4/68265.html

†Department of Engineering Mathematics, University of Bristol, Bristol BS8 1TR, UK (a.r.champneys@bristol.ac.uk).

‡Department of Mathematics, University of Auckland, Private Bag 92019, Auckland, NZ (kirk@math.auckland.ac.nz, be.oldeman@auckland.ac.nz, j.sneyd@auckland.ac.nz).

§Department of Physics, University of California, Berkeley, CA 94720 (knobloch@px1.berkeley.edu).

followed by slow relaxation. The power of the FitzHugh–Nagumo model arises from the fact that a vast array of physical and biological systems displays phenomenologically similar behavior. On the face of it, the propagation of a wave of electric potential down a nerve axon and the spread of a forest fire share little in common. Mathematical models are powerful because they can abstract the essential behaviors from each of these situations and show how similar they really are.

One recent application of excitable system theory has been to the study of the propagation of waves of increased calcium concentration both within and between cells [42, 4]. In the last 20 years it has become clear that the dynamics of calcium inside cells is responsible for the control of a wide array of cellular processes, and for this reason calcium waves and oscillations have been studied intensively by both experimentalists and theoreticians. Right from the earliest models it has been recognized that calcium waves propagate by an excitable mechanism essentially the same as that in the FitzHugh–Nagumo model and thus can be partially understood by the analysis of such simpler excitable systems.

Here, we write the FitzHugh–Nagumo equations in the form

$$\begin{aligned} v_t &= \Delta v_{xx} + f_\alpha(v) - w + p, \\ w_t &= \varepsilon(v - \gamma w), \end{aligned} \tag{1.1}$$

where x is a real spatial variable, t represents time, and $f_\alpha \equiv v(v-1)(\alpha - v)$. In the original FitzHugh–Nagumo model the variable v represents the plasma membrane electric potential, with diffusion constant Δ, and w is a phenomenological variable representing the combined inactivation effects of potassium and sodium ion channels. The small parameter ε represents the ratio of the time scale of the membrane potential (which is a fast process) to that of the channels (which act on a much longer time scale). The form of f_α is chosen so that, depending on the values of the threshold, α, or applied current, p, there is the possibility of either bistability between different spatially homogeneous states or excitability of the background state. (We use p to denote the applied current, rather than the more usual I, so as to maintain consistency with the later models discussed here.)

For Δ sufficiently large the background equilibrium state becomes destabilized and the dynamics supports pulse-type traveling waves which are solitary structures that move with a constant wave speed s. Equations (1.1) in the moving frame take the form

$$\begin{aligned} \dot{v} &= d, \\ \dot{d} &= \frac{1}{\Delta}(sd - v(v-1)(\alpha - v) + w - p), \\ \dot{w} &= \frac{\varepsilon}{s}(v - \gamma w), \end{aligned} \tag{1.2}$$

where the dot denotes differentiation with respect to the new variable $z \equiv x + st$. In these equations the background equilibrium is of saddle type, and the traveling pulse in the system (1.1) corresponds to a homoclinic orbit of this saddle in the moving frame.

In what follows we consider (1.2) for the parameter values

$$\Delta = 5.0, \quad \alpha = 0.1, \quad \gamma = 1.0, \quad \varepsilon = 0.01 \tag{1.3}$$

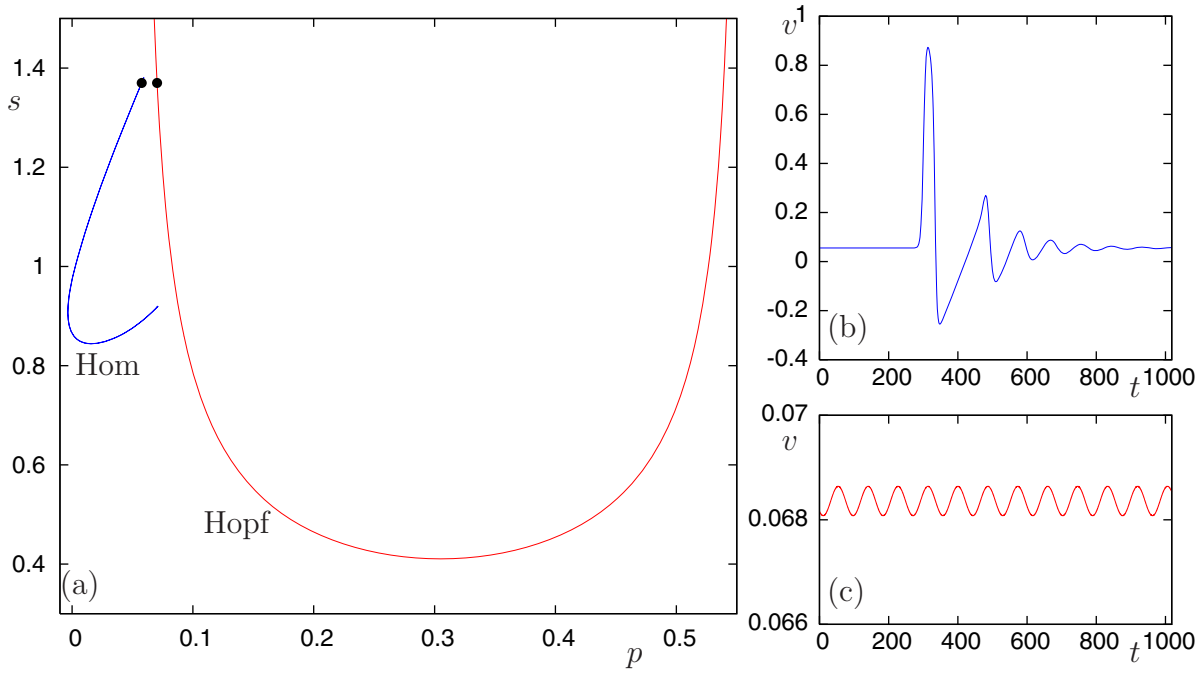


Figure 1. (a) *Partial bifurcation set for the traveling wave FitzHugh–Nagumo equations* (1.2) *for the parameter values given in* (1.3)*, illustrating the U-shaped curve of Hopf bifurcations and C-shaped curve on which homoclinic orbits occur. The large dots mark the* (p, s) *values corresponding to the time series given in panels* (b) *and* (c)*. Panel* (b) *shows a time series for a homoclinic orbit near the top of the C branch, at* $(p, s) = (0.0575862, 1.37)$*. Panel* (c) *shows a time series for a nearby periodic orbit, at* $(p, s) = (0.0703753, 1.37)$*; this periodic orbit is created in a Hopf bifurcation on the U branch.*

with the wave speed, s, and applied current, p, allowed to vary. For these fixed values there is a unique equilibrium point, which is of saddle type with a one-dimensional unstable manifold and a two-dimensional stable manifold provided that s is outside the red curve, $s = s_H(p)$, labeled Hopf in Figure 1. On the red curve, a Hopf bifurcation occurs and for $s > s_H(p)$ the equilibrium is completely unstable. Since the equilibrium no longer has a stable manifold, it is impossible for a homoclinic orbit to occur inside the red curve; hence the corresponding solitary pulses in the PDE cannot exist for parameter values inside the U-shaped curve. In fact, careful path-following techniques for homoclinic orbits [12] reveal that solitary pulse solutions lie on a C-shaped curve to the left of the U in the (p, s) parameter plane, which is also depicted in Figure 1. For more on computations of such traveling wave solutions in various different forms of the FitzHugh–Nagumo model, see [11, 36] and references therein. The C-shape of the pulse curve implies that more than one solitary pulse exists for a range of values of p, with two different wave speeds. In general though, it is known that only pulses on the upper, faster branch of the C can be stable.

Although it might seem that this structure, of a C-shaped curve of homoclinic orbits and a U-shaped curve of Hopf bifurcations, would be peculiar to the FitzHugh–Nagumo model, it turns out that a much wider array of excitable models share this same basic feature. Indeed, in every excitable model we have examined so far, including the FitzHugh–Nagumo model, many

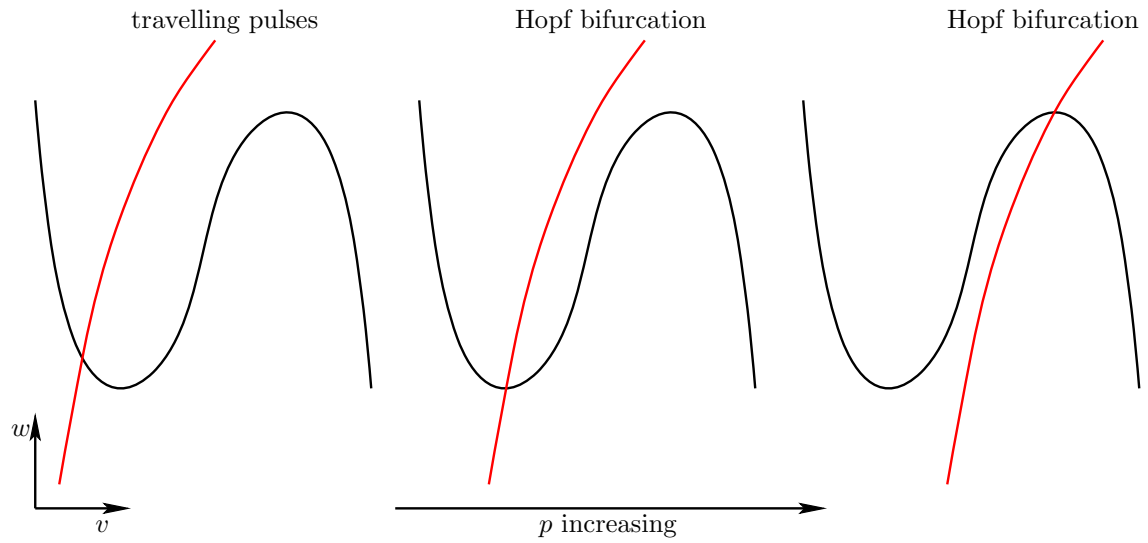


Figure 2. *Nullclines of a generic excitable system for increasing values (left to right) of the bifurcation parameter, p. When the red nullcline intersects the N-shaped nullcline just to the left of the minimum, a traveling pulse can exist. Hopf bifurcations occur as the red nullcline passes through the turning points of the N-shaped nullcline. This generic behavior, as p increases, leads to a generic CU structure in the (p, s) two-parameter bifurcation set.*

excitable models of calcium waves, and the Hodgkin–Huxley equations, this same fundamental structure appears. In the following we refer to such systems as *CU systems*.

The reason for this CU structure in excitable systems can be understood intuitively from consideration of nullclines (Figure 2). First consider an excitable system (such as the FitzHugh–Nagumo equations) with two important nullclines: an N-shaped nullcline and another roughly linear nullcline such as that shown by the red nullcline in Figure 2. As the bifurcation parameter, p, increases, the two nullclines move relative to one another, so that, effectively, the red nullcline moves across and intersects the N-shaped nullcline, first at the local minimum and then at the local maximum, generating a Hopf bifurcation at each of these distinguished intersections. The two Hopf bifurcations will exist for all wave speeds, s, above a minimum value, and thus the curve of Hopf bifurcations in the (p, s) plane will have a generic U-shape. (Depending on the exact shape of the nullclines, this U can have additional loops.) Furthermore, in a system of this kind, traveling pulses typically occur for some interval of the bifurcation parameter where the two nullclines intersect to the left of the local minimum, and appear in a saddle-node bifurcation as p increases, giving rise to a C-shape in the (p, s) plane. Thus, the traveling pulse and the two Hopf bifurcations occur sequentially as p increases and the CU structure is observed. Any excitable system that has nullclines of these general shapes (which includes most models of calcium waves, as well as the FitzHugh–Nagumo and Hodgkin–Huxley models) will have this CU structure.

In addition to the FitzHugh–Nagumo model, we study here a hierarchy of models of intracellular calcium waves, originally developed as descriptions of calcium waves in pancreatic acinar cells [44, 24, 42]. A typical model for intracellular calcium dynamics in one spatial dimension has the form

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} + J_{\mathrm{IPR}}(G, c, p)(c - c_e) - J_{\mathrm{serca}} + \delta(J_{\mathrm{in}} - J_{\mathrm{pm}}),$$
$$\frac{dc_e}{dt} = -J_{\mathrm{IPR}}(G, c, p)(c - c_e) + J_{\mathrm{serca}},$$
$$\frac{dG}{dt} = k_3(p, c)G,$$

where c denotes the concentration of free calcium in the cell cytoplasm, c_e denotes the concentration of free calcium in the endoplasmic reticulum (ER), and G denotes a set of gating variables that describe the time-dependent behavior of the calcium release channel, the so-called inositol trisphosphate receptor (IPR). The variables G are exactly analogous to the variables m, n, and h in the Hodgkin–Huxley model, or to the variable w in the FitzHugh–Nagumo model. The parameter p denotes the concentration of the second messenger inositol trisphosphate (IP_3). IP_3 is produced as a result of stimulation by hormones or neurotransmitters, and is the second messenger that controls the release of calcium from the ER, via the IPR. In all the calcium wave models we study here, p will be treated as a bifurcation parameter. In general, it can be manipulated experimentally, by changing the level of stimulation applied to the cell. In some cell types p is also a dynamic variable, with its rates of production and degradation controlled by calcium. Although we do not consider any such models here, preliminary computations show them to have the same CU structure, and thus our results should be applicable to that class of models also.

The J's denote various calcium fluxes into and out of the cytoplasm and ER. For instance, J_{serca} denotes the flux of calcium through ATPase calcium pumps on the membrane of the ER. These pumps move calcium up its concentration gradient from the cytoplasm to the ER, consuming ATP in the process. J_{in} is a leak of calcium from the outside of the cell into the cytoplasm, and J_{pm} is the flux of calcium through plasma membrane ATPase calcium pumps, while J_{IPR} is the flux of calcium through the IPR. Note that J_{IPR} depends on G, c, and p, as these are the three major controllers of the IPR.

Within this overall class of calcium wave models there is huge variety. For a start, there are many different expressions for each of the various flux terms; the particular one chosen depends on many factors but is most strongly influenced by the available experimental data and by the goal of the model. In addition, there may be many gating variables, G, or only a few. Although complicated Markov models of the IPR, with many different states, are sometimes used to derive the differential equations for G, these complex models can often be simplified by assuming fast transitions between various states, leading to simpler models with fewer gating variables. Here, we study a variety of models, ranging from a model with a single gating variable to a model with seven gating variables. Readers interested in the process of model construction and simplification are referred to the original papers. Here we shall merely state the model equations and their parameters.

An aim of this paper is to identify the ways in which the C-shaped curve of homoclinic orbits can terminate, either at or before it reaches the U-shaped curve of Hopf bifurcations. For example, in Figure 1, it may seem that the computation of the curve of homoclinic bifurcations shown in the figure has been abnormally terminated, since topological considerations demand that a bifurcation curve cannot end "in mid air" unless it is at a codimension-two point. In fact, as we shall see in section 3.1 below, there is a rational explanation of this

apparent termination, which to our knowledge has not been reported in the literature. More generally, a detailed understanding of how the C and U curves interact is crucial for an understanding of wave structure in these excitable models. For lower values of p in Figure 1 the model supports a stable traveling pulse solution, but as p increases, these pulse solutions become unstable and, inside the U curve, turn into periodic plane waves. Thus, in order to understand how pulse waves turn into periodic waves, we must first understand the ways in which the C curve can terminate, either at or before the U curve. Since, in all the models we consider here, the parameter p is one that can be controlled experimentally, it is plausible that theoretical predictions about the transition from pulse to periodic waves will be able to be tested experimentally. However, such considerations are left for future work, as a proper mathematical understanding is a prerequisite for the development of testable predictions.

For each model, we shall carry out a detailed numerical bifurcation analysis in order to elucidate the way in which the homoclinic curve terminates. We will show that homoclinic branches can terminate in quite different ways in different systems. First, the mechanism involved depends on whether the Hopf bifurcation is subcritical or supercritical at the point where the homoclinic bifurcation curve approaches, but even with this knowledge there are several possibilities, and the only way to determine what is going on is to use numerical techniques. These investigations have helped us identify five different codimension-two mechanisms by which a branch of homoclinic bifurcations of an equilibrium may be terminated in a CU system: a Shil'nikov–Hopf bifurcation; an equilibrium-to-equilibrium heteroclinic cycle (the so-called T-point); a local saddle-node/Hopf bifurcation; a codimension-two heteroclinic cycle between an equilibrium and a periodic orbit; and a tangency in a codimension-one equilibrium-to-periodic heteroclinic cycle. Termination via a heteroclinic cycle between an equilibrium and a fold of limit cycles is also possible, but this mechanism has not yet been seen in a CU system.

The rest of the paper is organized as follows. Section 2 gives a brief overview of homoclinic bifurcation theory and in particular of some codimension-two mechanisms which can terminate a branch of homoclinic orbits as it approaches a Hopf bifurcation. Relevant references are given to places in the literature where full unfoldings may be found, or, in the case of at least two of the mechanisms, to forthcoming analysis by some of the present authors that will provide just such detail. Section 3 contains the main contribution of this paper, which is the presentation of careful numerical continuation results that reveal the specifics of the bifurcation diagrams for each of the six example systems. Finally, section 4 attempts to draw general lessons from these numerical results. Tentative explanations are also offered for previously reported "anomalous" Shil'nikov–Hopf bifurcations, where a homoclinic curve appears to pass straight through the Hopf curve, in defiance of theory, a "ghost" phenomenon that appears to be due to the slow-fast nature of excitable systems.

2. Homoclinic bifurcations.

This section introduces notation and reviews results about homoclinic bifurcations from the existing literature. See, for instance, [29, 39] and references therein for full details.

2.1. Codimension-one Shil'nikov homoclinic bifurcation.

Consider a continuous-time dynamical system

$$\dot{u} = f(u, p), \quad u \in \mathbb{R}^n, \quad p \in \mathbb{R}^m,$$

and assume there is a hyperbolic equilibrium point $u_0(p)$ which for a codimension-one set of parameter values has an isolated homoclinic orbit H_1 that connects its stable and unstable manifolds. The dynamical behavior of orbits near the homoclinic orbit depends on, among other things, the eigenvalues of the flow linearized about the equilibrium. The *leading* eigenvalues of the equilibrium are defined as the eigenvalues of $f_u(u_0(p), p)$ that have positive and negative real parts closest to zero. From now on, we assume that the leading eigenvalues of the equilibrium involved in the homoclinic bifurcation consist of a real, positive eigenvalue $\lambda > 0$ and a pair of complex conjugate eigenvalues $-\mu \pm i\omega$, $\mu > 0$. This is precisely the eigenvalue structure characteristic of the equilibrium to the left (of the left arm) of the U-shaped Hopf curve in all of the examples that follow.

The saddle quantity $\sigma = \lambda - \mu$ is defined as the sum of the real parts of the leading stable and the leading unstable eigenvalues. If $\sigma < 0$ then the homoclinic bifurcation upon varying a parameter is tame and involves finitely many periodic orbits in a neighborhood of the homoclinic orbit in both phase and parameter spaces. The case of prime interest here is the so-called Shil'nikov behavior that occurs when $\sigma > 0$. This case involves shift dynamics on infinitely many symbols which implies, in particular, the existence of infinitely many periodic orbits that have arbitrarily long periods of spiraling behavior close to u_0. Moreover, there are sequences of nearby parameter values at which there exist double-pulse homoclinic orbits H_2 that make two large excursions in a neighborhood of the primary homoclinic orbit. In addition there are infinitely many N-pulse homoclinic orbits for all positive integers N.

A number of codimension-two mechanisms have been observed by which the dynamics in a neighborhood of H_1 changes qualitatively; see [11] for a partial review. For example, the saddle quantity can change sign as we vary parameters along a homoclinic bifurcation curve in the parameter plane. Such a transition is called a Belyakov bifurcation, after the Russian mathematician who first described this situation [3]. A schematic representation of what occurs in such a codimension-two bifurcation is depicted in the lower portion of Figure 3. Another kind of Belyakov bifurcation [2] occurs when the complex pair of leading eigenvalues associated with the equilibrium undergoing a homoclinic bifurcation becomes real, so that a saddle-focus becomes a saddle. This latter transition, which is depicted in the upper part of Figure 3 is the one that we encounter in this paper. Associated with this type of Belyakov transition is the birth of multipulse homoclinic orbits; these orbits enter and then leave a neighborhood of the equilibrium one or more times before converging back to the equilibrium. Codimension-one curves of multipulse homoclinic orbits emanate from the codimension-two Belyakov point.

Neither of these codimension-two bifurcations produces an end-point on the C-shaped homoclinic branch. The question we wish to address is just how this termination happens as the C approaches the U in the systems we investigate. Naively, one might expect that this would occur via the so-called Shil'nikov–Hopf bifurcation (see Figure 4) that occurs as the complex eigenvalues of the Shil'nikov equilibrium cross the imaginary axis. However, as summarized in the next section, there are other possibilities as well.

2.2. Possible codimension-two termination mechanisms. Figure 4 depicts partial unfoldings of six different codimension-two mechanisms by which a branch of homoclinic orbits to an equilibrium may terminate as it approaches a Hopf bifurcation curve in a parameter plane. We describe below the dynamics corresponding to each of the panels in the figure. Note

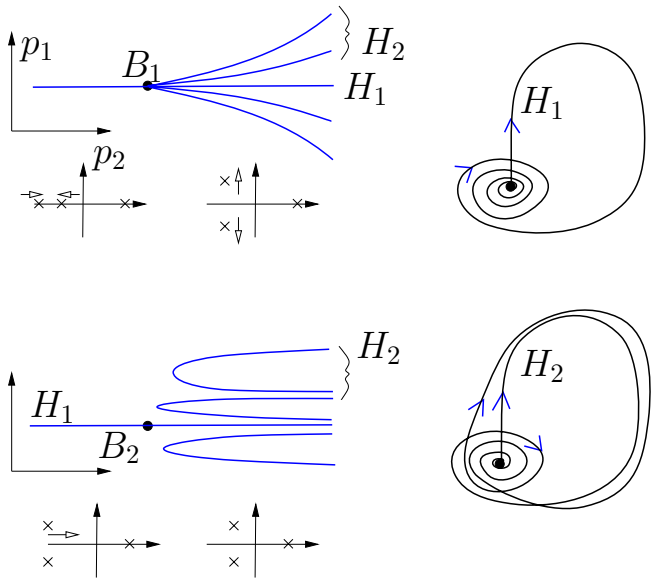


Figure 3. *Partial unfoldings of two Belyakov transitions in the (p_1, p_2) parameter plane. Here p_1 unfolds the existence of a primary homoclinic orbit H_1, whereas variation of p_2 causes the homoclinic bifurcation to change from the tame case to the Shil'nikov case. In the upper panel, the Belyakov point B_1 is caused by a pair of nonsemisimple eigenvalues coalescing on the real axis and becoming complex. In the lower panel the saddle quantity σ changes sign. In both cases, infinitely many N-pulse homoclinic orbits are created in the neighborhood of the B-points, represented here by a few curves of double-pulse orbits H_2. There are also infinitely many curves of period-doublings and folds of periodic orbits that are not depicted.*

that this list is not exhaustive but is specific to the case found in the CU systems considered here where there is an equilibrium that has a one-dimensional unstable manifold to the left of the U-shaped Hopf curve. In section 3 below we shall find examples of at least four of these mechanisms in the CU systems we consider.

The six cases are as follows:

(a) A Shil'nikov–Hopf bifurcation. This codimension-two point occurs when the equilibrium to which the homoclinic orbit is connected undergoes a supercritical Hopf bifurcation. On the far side of the Hopf bifurcation, the parameter curve of homoclinic orbits transforms into a heteroclinic cycle that connects the equilibrium to the small-amplitude limit cycle born at the Hopf bifurcation. There is also a parabolic curve on which homoclinic tangencies to the periodic orbit exist. Inside the parabola there are shift dynamics on infinitely many symbols. This situation was first analyzed by Belyakov [1], with more complete treatments appearing in the work of Hirschberg and Knobloch [26] and Deng and Sakamoto [13].

(b) A fold-Hopf (or saddle-node/Hopf) bifurcation occurs when a pure imaginary pair and a zero eigenvalue occur in the linearization about the equilibrium. See [29] and references therein for details of the unfolding of this codimension-two local bifurcation. There are several cases, depending on the signs of certain nonlinear terms in the two-dimensional normal form. In the so-called $(+,-)$ case, it is inevitable that two small-amplitude homoclinic orbits originate from this codimension-two bifurcation [5, 19],

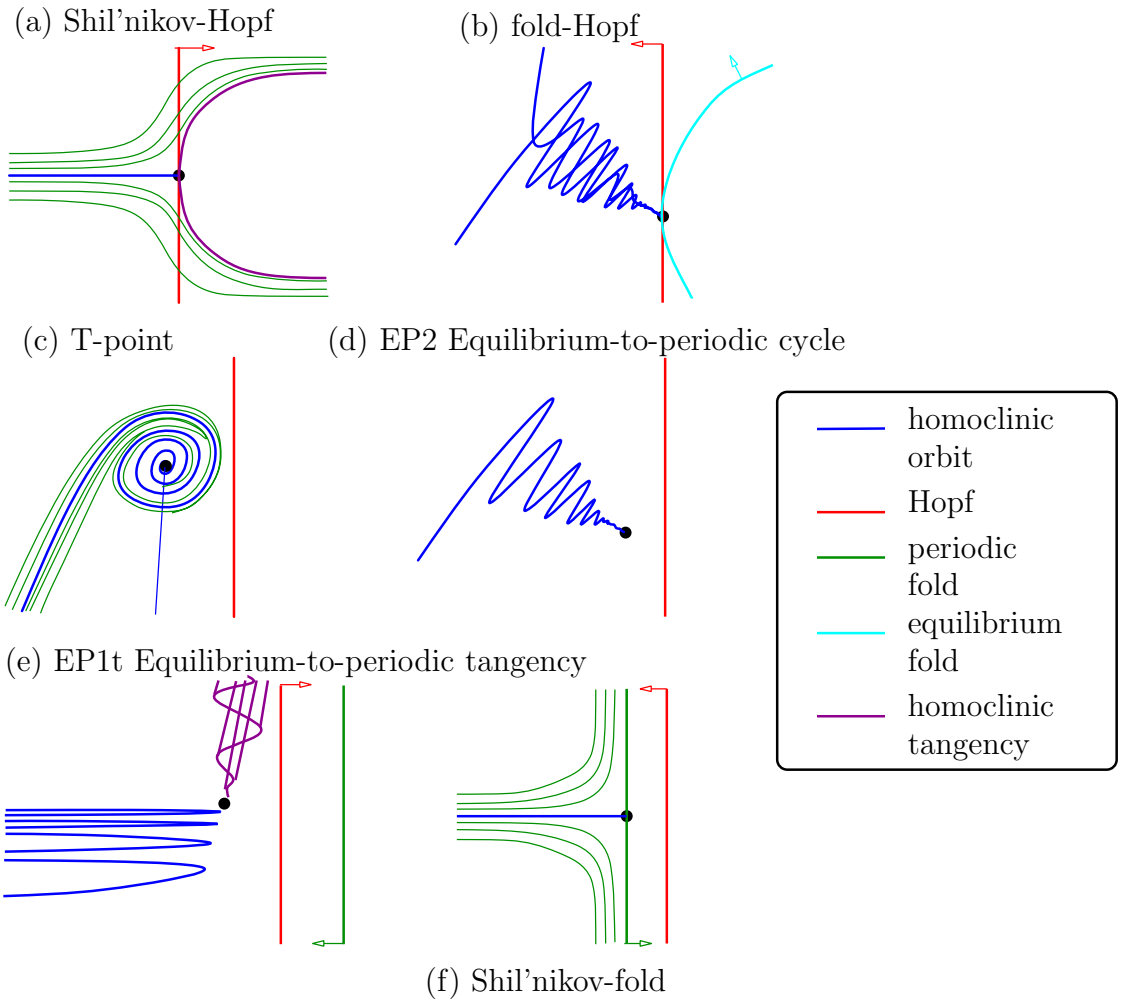


Figure 4. *Qualitative unfolding diagrams of six different codimension-two mechanisms that could cause the termination of the homoclinic branch (blue curve) in a CU system, as it approaches the (red) Hopf curve. The arrows in the figure denote the direction in which the periodic orbit bifurcates from the Hopf bifurcation or from other key bifurcation curves. Note that in full unfoldings of each case there are many more bifurcation curves than those depicted. See text for more details.*

doing so with a characteristic entwined wiggling in the parameter space; see [8].

(c) A T-point [25] (or Bykov point [6, 7]). This is an equilibrium-to-equilibrium codimension-two heteroclinic cycle, and two branches of primary homoclinic orbits bifurcate from it. If one equilibrium is a saddle-focus with a negative saddle index and the other has real eigenvalues, then one curve of homoclinic orbits spirals in parameter space, while the other approaches the T-point monotonically. Such a T-point may or may not appear close to a Hopf bifurcation in parameter space.

(d) Equilibrium-to-periodic heteroclinic cycle (EP2 point). At such a point there exists a heteroclinic cycle consisting of a connection from an equilibrium to a periodic orbit and a connection from the periodic orbit back to the equilibrium. In the case that the

unstable manifold of the equilibrium is one-dimensional, we require that the periodic orbit has a three-dimensional unstable manifold and an $(n-2)$-dimensional stable manifold. Then the heteroclinic connection *from* the equilibrium *to* the periodic orbit will itself be of codimension two while the connection back the other way is generic. Note that, unlike each of the other codimension-two mechanisms we review here which can occur in $\mathbb{R}^3$, an EP2 point requires four phase space dimensions. The asymptotics of how the curve of homoclinic orbits approaches such an EP2 point were studied in [34]. The rate and manner of convergence (i.e., direct or wiggly approach) depend in part on the size of the Floquet multiplier of the periodic orbit and whether the Floquet multipliers are real or complex. A complete unfolding of other codimension-one bifurcation curves nearby is unknown.

(e) Equilibrium-to-periodic heteroclinic tangency (EP1t). In contrast to an EP2 point, here the periodic orbit has a two-dimensional unstable manifold, and so the heteroclinic connection *from* the equilibrium *to* the periodic orbit is of codimension one, whereas the connection from the periodic orbit *back to* the equilibrium remains generic. A codimension-two point is reached when the generic connection forms a tangency. This case was partially studied in [34], with more details to appear in [9]. In fact, such a point does not represent an end-point of a curve of homoclinic orbits; there is an infinite number of homoclinic curves, each of which undergoes a turning point in the parameter plane close to the codimension-two point. Each successive curve corresponds to an orbit with an additional loop near the periodic orbit. Also arising from the codimension-two point are curves of homoclinic tangencies to the periodic orbit that have a complex topology.

(f) A blue-sky catastrophe of homoclinic orbits. This occurs when the homoclinic bifurcation curve ends at a curve of fold bifurcations of periodic orbits. One natural way for such a codimension-two point to arise would be if the saddle-node periodic orbit appears in the stable manifold of the equilibrium to the left of the red Hopf curve in the case where the Hopf bifurcation is subcritical (the small arrows in Figure 4 indicate the direction of bifurcation). The resulting barrier prevents the existence of a homoclinic orbit, and the homoclinic branch must terminate via the orbit wrapping itself infinitely many times around the emerging structurally unstable periodic orbit. This case is currently being studied [10]. There may be other similar termination mechanisms involving homoclinic wrapping around nonhyperbolic periodic orbits, for example, orbits undergoing a period-doubling bifurcation, but we are not aware of any theoretical analyses.

It is pertinent to point out that homoclinic curves that appear to terminate "in mid air" in a parameter plane may in fact not imply the existence of a codimension-two bifurcation point at all. For example, we shall see in several of the examples below cases where the homoclinic bifurcation curve turns around sharply and doubles back on itself as it approaches (but does not quite reach) the Hopf bifurcation. The homoclinic orbit typically gains an extra loop as it does so, so that globally a single-pulse orbit becomes a double-pulse orbit. The sharpness of the turn in the parameter space means that the bifurcation curve appears to terminate, but there is in fact no terminating bifurcation. However, as we shall see, such sharp turns occur naturally as a consequence of some of the other mechanisms listed above, notably the EP1t case.

3. Numerical continuation results. Throughout we use the continuation routines HomCont [11, 12] that are embedded in the software AUTO [15] in order to trace curves of homoclinic orbits in two parameters. AUTO is a numerical bifurcation analysis package which uses pseudo-arclength continuation in combination with the solution of two-point boundary value problems with Gauss–Legendre collocation. We use the other capabilities of AUTO [14, 15] to trace curves of local bifurcations of equilibria and periodic orbits and to identify codimension-two bifurcation points. The periodic and homoclinic orbits under consideration in this paper are discretized with a total of `NTST` $\times$ `NCOL` points, where `NTST` is the number of mesh intervals and `NCOL` the number of so-called Gauss collocation points per mesh interval.

The general method that was used to produce the bifurcation diagrams was as follows: an equilibrium point was determined either analytically or numerically for certain parameter values. This equilibrium was then continued in one parameter, until a Hopf bifurcation was detected. AUTO can both continue this Hopf bifurcation in two parameters and continue a periodic orbit from the Hopf bifurcation in one parameter. The period of the periodic orbit grows to infinity as the homoclinic orbit is approached. Then the HomCont algorithm can phase-shift such a periodic orbit so that the beginning and end-points are close to the equilibrium and then continue the homoclinic orbit in two parameters using projection boundary conditions.

Note that cases (a) and (e) in Figure 4 involve interaction of the primary homoclinic curve with homoclinic tangencies to periodic orbits. There are currently no implemented general boundary value methods for continuation of homoclinic tangencies to periodic orbits in arbitrary dimensions; although see [45]. Instead, we can use the result that homoclinic tangencies are accompanied by a sequence of parameter values at which fold bifurcations occur [21, 20]. Therefore, by following folds of high-period periodic orbits that approach a homoclinic tangency, we can compute approximations to such global bifurcation curves.

The systems studied in this paper are all highly stiff (with time scales differing by up to 13 orders of magnitude in some cases). This stiffness complicates the computations. For instance, the computation of Floquet multipliers of periodic solutions in AUTO is not very reliable once the system has very small or very large multipliers; see [31]. Of particular significance for the systems in this paper is the observation that detection of homoclinic bifurcations can be difficult in stiff systems. For instance, there may exist periodic orbits with extremely high period that can be misdetected as implying the existence of homoclinic orbits. In such circumstances it is sometimes necessary to increase the number of mesh intervals (`NTST`) to values up to 600 to retain convergence, while a number of about 30 is usually sufficient for nonstiff systems. Similarly, it was sometimes found to be necessary to increase the truncated time interval T over which we compute approximations to homoclinic orbits up to 10^6 to avoid, but still not completely eliminate, false positive results.

3.1. Model 1: FitzHugh–Nagumo. We begin by considering the bifurcation set of the traveling wave FitzHugh–Nagumo system (1.2) with the parameter values (1.3), as discussed in the introduction. Note that equations (1.2) have a symmetry. Specifically, in the case $\gamma = 1$, $\alpha = 0.1$, the equations are equivariant under the transformation

$$v \to \frac{11}{15} - v, \quad w \to \frac{11}{15} - w, \quad d \to -d, \quad p \to \frac{11}{15}\left[1 - \frac{38}{225}\right] - p, \tag{3.1}$$

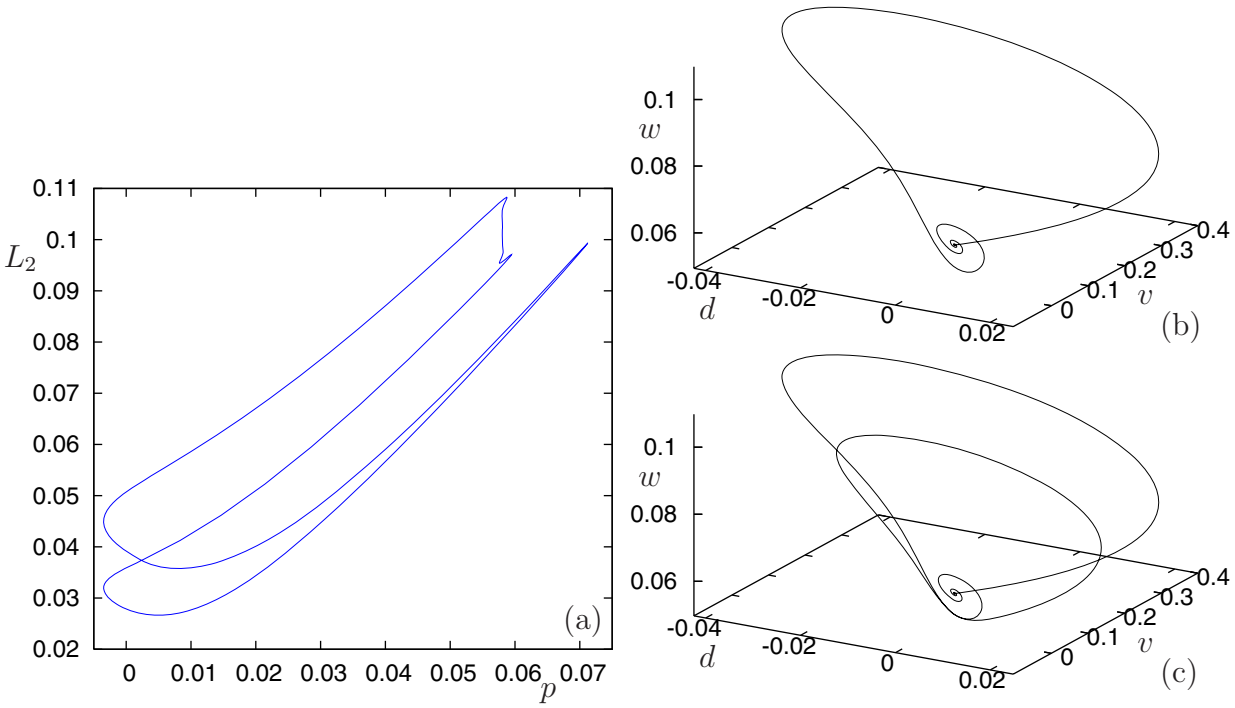


Figure 5. *Another projection of the "homoclinic banana" for the FitzHugh–Nagumo system* (1.2) *with the parameters* (1.3) *(cf. the C curve in Figure* 1*). Panels* (b) *and* (c) *show two different homoclinic orbits at* $p \approx 0.06$, $s \approx 0.894386$ *near the lower end of the banana. Note the gross difference in phase space between these two orbits. We (loosely) refer to the orbit in* (b) *as a single-pulse orbit and the orbit in* (c) *as a double-pulse orbit.*

which means that the bifurcation set in the (p, s) plane is symmetric under reflection through the line $p = 0.3047$. Thus there are actually two C-shaped curves of homoclinic bifurcations, one to the left of the Hopf U as shown, and a mirror image to the right of the U (not shown). We discuss only the leftmost C here; results are analogous for the other C.

Detailed numerics reveal that the homoclinic locus that appears to be a single C-curve in Figure 1 is in reality a pair of curves which join together in two fold-like structures in the parameter plane, forming a closed curve; when the homoclinic branch is followed numerically in the parameters p and s, the branch turns around at one end of the C and then appears to retrace its path to the other end of the C, where it turns around again and returns to its starting point. On the scale of Figure 1, the C appears to be a single curve, and even when zooming in the separation between the curves is hard to see because the distance between curves approaches the numerical accuracy of AUTO/HomCont. However, as shown in Figure 5, the two curves can be distinguished by plotting the numerically computed L_2-norm along the homoclinic locus. We call the thin closed loop of homoclinic orbits in Figure 1 a *homoclinic banana*. Thus we see that the homoclinic locus in the FitzHugh–Nagumo system does not terminate at each end in a codimension-two point, as would seem the case from a first glance at the bifurcation set. This is an example of the *apparent* termination mechanism discussed at the end of section 2.2. We note that the ends of the C curve have different

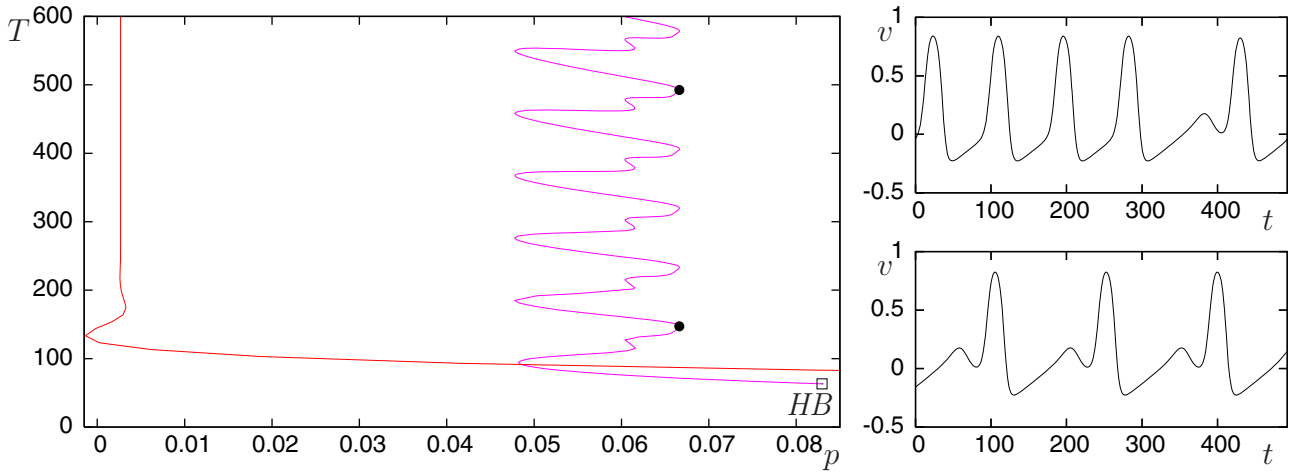


Figure 6. *Bifurcation diagram and time series for the FitzHugh–Nagumo model with $s = 1$. Left: Bifurcation diagram showing the period T as a function of p. The red curve continues beyond the right arm of the Hopf U, eventually terminating in a second homoclinic bifurcation at $p \approx 0.61$. The purple curve arises from a Hopf bifurcation on the left arm of the Hopf U (bifurcation point marked HB). Right: Time series for the periodic orbits with period $T = 492.476$ and $T = 147.184$, indicated by the large dots in the bifurcation diagram.*

shapes when the L_2-norm is plotted against p, with the upper end being blunt and involving a number of turning points, while the lower end has, within numerical accuracy, just one sharp turn. Preliminary calculations suggest the existence of *canards*, that is, rapid growth from small scale to large scale oscillations in a thin wedge of parameter space, near the upper end of the C curve, and this may explain the difference in structure at the two ends.

Near both ends of the C curve, the Hopf bifurcation is subcritical, and the periodic orbit created in the Hopf bifurcation appears on the left side of the U. Numerical continuation of this periodic orbit for fixed $s = 1$ produces the purple wiggly curve shown in Figure 6. The form of this curve indicates that the periodic orbit approaches a set of homoclinic tangencies to a limit cycle (see, e.g., [20]). There are four accumulation points for the saddle-node bifurcations of periodic orbits on the purple curve in Figure 6, presumably corresponding to four separate parameter values at which there are tangencies between the stable and unstable manifolds of a hyperbolic limit cycle. We note here that the rate of converge in p of each sequence of folds to a fixed value is extremely rapid. As each group of four consecutive folds is traversed the orbit undergoing this wiggling acquires an extra loop near the limit cycle (see the two subpanels to the figure which depict orbits with 1 and 5 oscillations per period near the limit cycle, respectively). The limit cycle in question (shown as a red curve in Figure 6) can itself be continued into a Shil'nikov homoclinic orbit. At different parameter values we might thus expect to see an interaction between homoclinic orbits to an equilibrium and homoclinic tangencies to a periodic orbit.

Figure 6 suggests there is a lot more structure to the bifurcation set for the FitzHugh–Nagumo equations than is shown in Figure 1. Figure 7 shows some of the additional bifurcations that occur between the C and the U. Note, in particular, the degenerate Hopf bifurcation marked with a square and the curve of saddle-node bifurcations of a periodic orbit

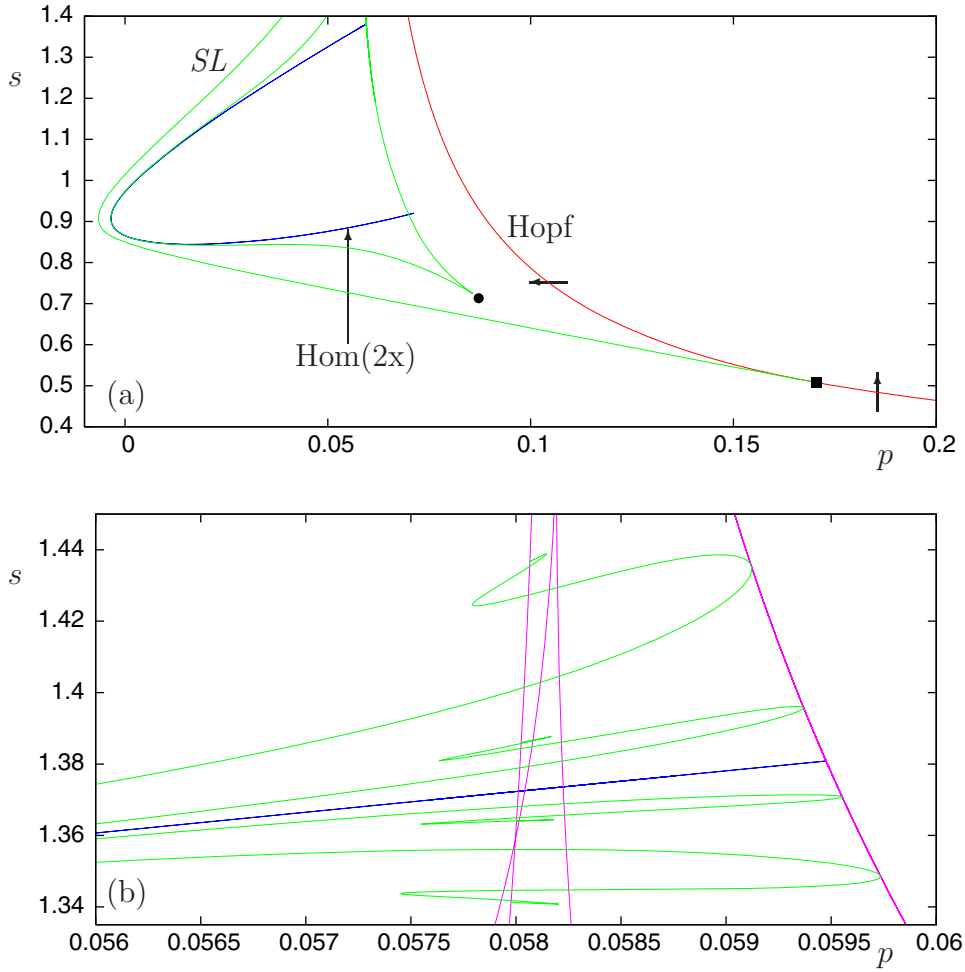


Figure 7. *A more detailed bifurcation set for the FitzHugh–Nagumo system.* (a) *The curve labeled Hom* $(2x)$ *is the "homoclinic banana" described in the text. The rightmost curve is part of the U-shaped Hopf curve, with the square marking a degenerate Hopf bifurcation and the large dot marking a cusp bifurcation. All other curves denote saddle-node bifurcations of limit cycles (SL). The arrows indicate the direction in which periodic orbits appear on the various sections of the Hopf bifurcation curve.* (b) *A zoom near the top end of the homoclinic curve. The blue curve delineates the homoclinic banana. The green curves show saddle-node bifurcations of limit cycles associated with the unfolding of the codimension-one Shil'nikov bifurcation. The purple curves also show saddle-node bifurcations of periodic orbits, but these are associated with the folds on branches of periodic orbits that approach a homoclinic tangency of a periodic orbit (two such saddle-node bifurcations were marked with large dots in Figure 6).*

that emanates from this point. Also, some of the saddle-node bifurcations of periodic orbits seen in Figure 6 have been continued in two parameters and their loci are shown in Figure 7. We distinguish between two types of saddle-node bifurcations in the zoom in Figure 7(b). Of particular interest are the purple curves which correspond to the fold bifurcations in Figure 6 that approach the homoclinic tangencies. Thus the purple curves are approximations to loci of homoclinic tangencies to periodic orbits. The sharp turning point in the blue homoclinic curve

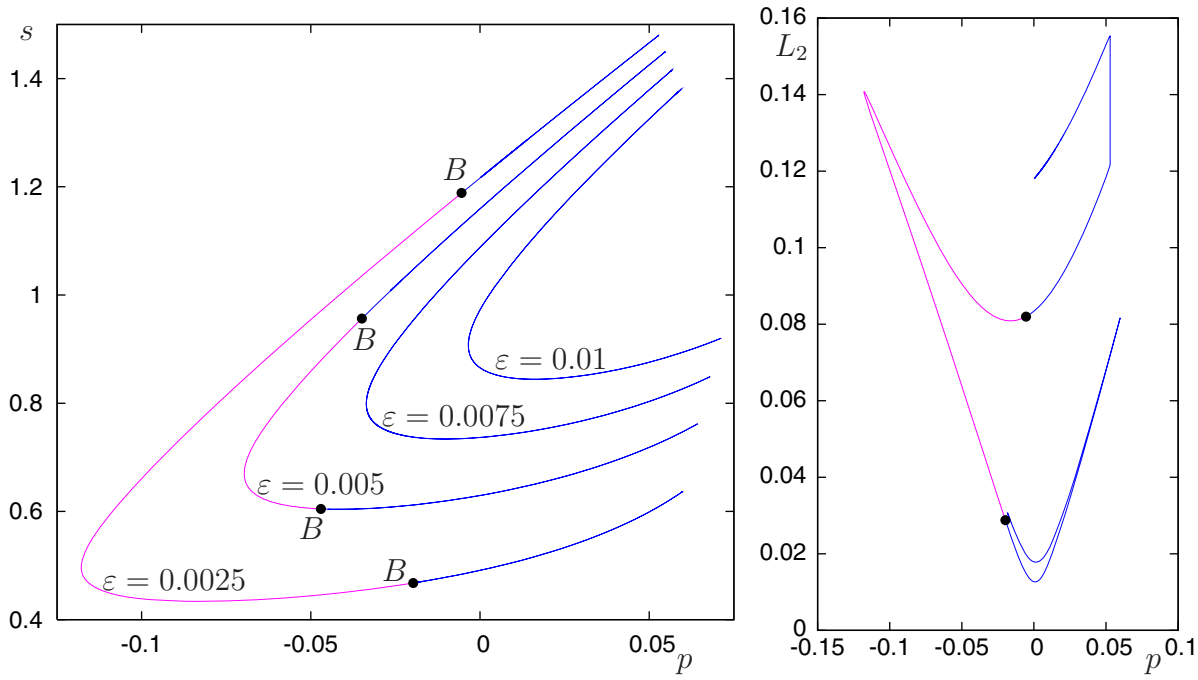


Figure 8. *Homoclinic bifurcation curves in parameter space for different values of ε in the FitzHugh–Nagumo system. Belyakov transitions (marked B) are present on the curves corresponding to $\varepsilon = 0.0025$ and $\varepsilon = 0.005$ (the banana-split case), but not the remaining curves (the banana case). On the left side of the Belyakov points the eigenvalues of the equilibrium are real rather than complex. Right: L_2-norm for the homoclinic orbits occurring along the banana-split curve in the case $\varepsilon = 0.0025$. This panel demonstrates that each point on the C-shaped curve between the Belyakov points corresponds to a single homoclinic bifurcation curve, while to the right of each Belyakov point there are two curves of homoclinic bifurcations.*

occurs very close to the parameter value at which the homoclinic orbit and homoclinic tangency interact. This is an example of an EP1t equilibrium-to-periodic heteroclinic tangency, which is case (e) of Figure 4.

It is also instructive to look for Belyakov transitions. It is straightforward to check numerically that, for the parameter values used above, the saddle-quantity of the homoclinic orbits on the C-curve is always positive and the stable leading eigenvalues are always complex. Thus homoclinic bifurcations of the equilibrium in this system always display Shil'nikov-type behavior, and no Belyakov transitions occur. However, for different values of the parameter ε the behavior changes. As is depicted in Figure 8, for smaller values of ε there exist two Belyakov transitions on the C-shaped homoclinic bifurcation curve. In between these two Belyakov points, the eigenvalues of the equilibrium are real and there exists just a single homoclinic bifurcation curve. On the right-hand side of each Belyakov point the C-curve is in fact a pair of homoclinic bifurcation curves just as for the homoclinic banana above; a single-pulse homoclinic orbit deforms into a double-pulse homoclinic orbit, with the double-pulse orbit then terminating at the Belyakov point. Thus the homoclinic banana is now split. Similar behavior was found in [33] and [30]. Using HomCont we were able to find the critical value of ε where the behavior changes and the two Belyakov points appear. This codimension-

three point, denoting the transition between the banana and the *banana-split* case, occurs at $\varepsilon = 0.006018385$ for $p = -0.05314325$ and $s = 0.07101599$.

In summary, we have gained some insight into the termination of the homoclinic bifurcation curve in the traveling-wave FitzHugh–Nagumo equations; namely, the C curve does not terminate at each end but instead folds back on itself, with the single-pulse homoclinic orbit deforming into a double-pulse orbit in the process. One way of explaining the sharp turning points of the homoclinic curve is through the numerical evidence we have presented for the existence of equilibrium-to-periodic heteroclinic tangencies (EP1t points). The intricate dynamics near the EP1t point, particularly the interaction between fold curves and homoclinic tangencies, will be discussed further in [9]. Note though that things are made even more subtle by the slow-fast nature of the FitzHugh–Nagumo model and the consequent occurrence of canard-like behavior.

3.2. Model 2: Sneyd, LeBeau, and Yule three variable model. In [41] a model of intracellular calcium wave propagation in pancreatic acinar cells was constructed and studied. Two versions of this model were considered; after passage to traveling wave coordinates one model has three phase space variables and the other four. In this section we look at the former (SLY-3), and in the next section at the latter (SLY-4). For more details on the construction of these models, the interested reader is referred to the original paper.

The SLY-3 model equations are given by

$$\begin{aligned} \dot{c} &= d, \\ D_c\dot{d} &= sd - k_f\left(\frac{ph\varphi_1}{\varphi_1 p + \varphi_{-1}}\right)^4 + \frac{V_p c^2}{K_p^2 + c^2} - J_{\text{leak}}, \\ s\dot{h} &= \varphi_3(1-h) - \left(\frac{\varphi_1\varphi_2 p}{\varphi_1 p + \varphi_{-1}}\right)h. \end{aligned} \tag{3.2}$$

Here c denotes the calcium concentration, s is the wave speed introduced when moving to traveling wave coordinates, and p represents the concentration of inositol trisphosphate (IP_3), a second messenger that can be experimentally controlled, at least to a certain extent. All other quantities in the equations are either constant or functions of c; see Appendix A for details. We are interested in the bifurcation set in the (p, s) parameter plane. Much of the bifurcation set for this model was described already in [41], with some further details in [37, 8]. This example is presented here as an illustration of two of the five mechanisms (i.e., T-point and fold-Hopf) by which homoclinic loci terminate near Hopf bifurcations in CU systems. Several new relevant bifurcation results are also included.

There are four main homoclinic branches in this system. One of these, named branch A following the notation of [41], is shown in Figure 9 together with the locus of Hopf bifurcations. We regard this system as a CU system even though the two arms of the Hopf U cross each other, forming a loop as shown in Figure 9. The Hopf bifurcation is degenerate at $s \approx 4.858$ and at $s \approx 6.7502$, and the direction in which the bifurcating periodic orbit appears on each section of the Hopf bifurcation curve is indicated in the figure.

At the upper end of the C-shaped curve, branch A terminates at a T-point, as described in [41]. An interesting result about the dynamics near the T-point was proved in [35], where

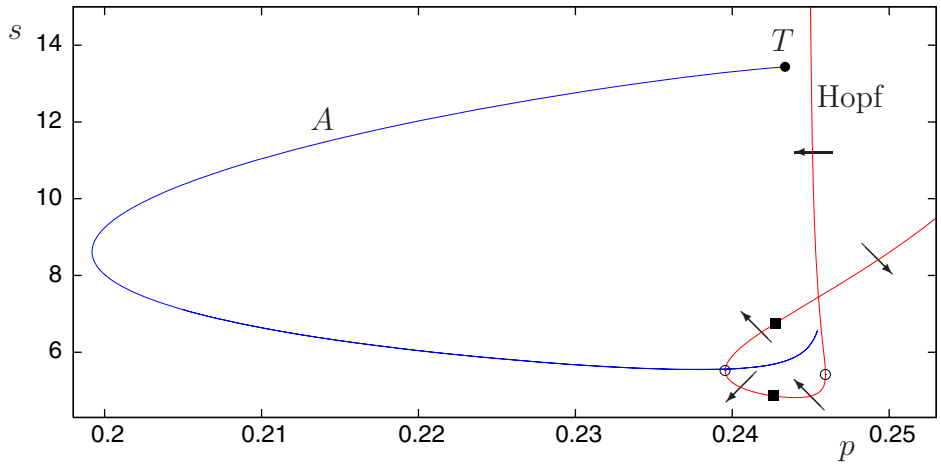


Figure 9. *Partial bifurcation set for the SLY-3 model, equations (3.2), showing the CU nature of the bifurcation curves. The two arms of the Hopf U cross one another, forming a loop, with the arrows denoting the directions in which periodic orbits appear in the Hopf bifurcation. The open dots on the Hopf locus denote saddle-node Hopf bifurcations, and the squares mark degenerate Hopf bifurcations. The C-shaped homoclinic bifurcation curve (labeled A following the notation of [41]) terminates at the top at a T-point (labeled T) and turns around sharply at the lower end in the same way as the lower end of the C curve in the FitzHugh–Nagumo equations.*

it was shown that stable pulses in the underlying PDEs arise from the gluing together of two unstable fronts. At the lower end of the C-shaped curve, the branch of homoclinic bifurcations makes a tight turn, and the homoclinic orbit deforms continuously from a single-pulse orbit to a double-pulse orbit, in a similar way to that seen near the lower end of the C curve in the FitzHugh–Nagumo model. The branch of double-pulse orbits then doubles back all the way to the T-point, terminating at or very near the T-point. Thus, once again we find the C-shaped curve is a homoclinic banana, although in this case one end of the banana is attached to the T-point. We note that the banana-nature of branch A is a new result. In [41] it was conjectured that branch A terminates near the Hopf bifurcation when the amplitude of the homoclinic orbit goes to zero. On the other hand, [37] showed that branch A had a sharp turning point at the lower end of the C curve but was able to trace the doubled-back curve only as far as the large bend in the C at $(p, s) \approx (0.2, 8.5)$. This is an example of the way in which the extreme stiffness of the CU models considered in this paper can result in misleading numerical results, even when great care is taken with the computations.

Three further curves of homoclinic bifurcations (branches B to D) are shown along with branch A in Figure 10. Branch B starts at the T-point; in Figure 10 it leaves the diagram in the direction of increasing s and will not concern us further. Branch C starts at a saddle-node Hopf point and does not terminate; it also leaves the bifurcation set in the direction of large s. Branch D starts at the same saddle-node Hopf point as branch C (it wiggles around branch C; see [8]) and then follows branch A, finally terminating at the T-point.

There are other curves of homoclinic bifurcations besides the four branches, A to D, identified above. For instance, if we follow one of the double-pulse orbits that lies close to branch A, we find that it, too, originates from the T-point. In the other direction these

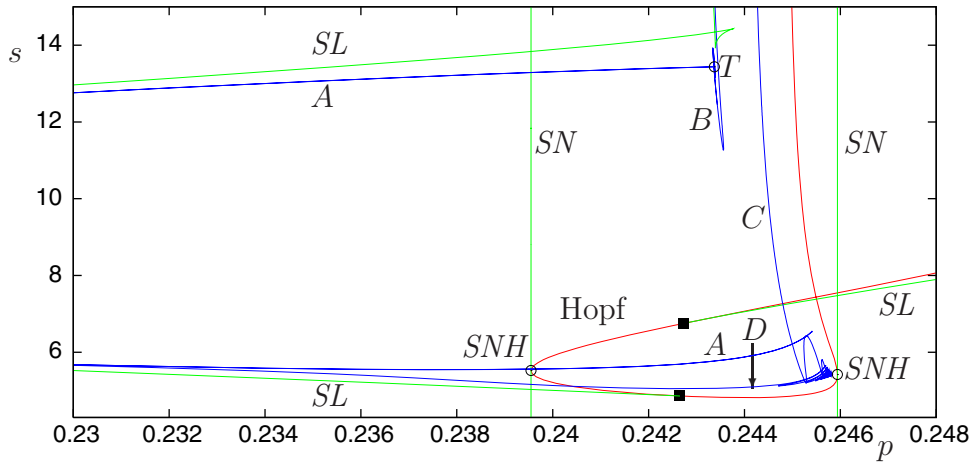


Figure 10. *A more detailed bifurcation set for the SLY-3 model. This figure is an enlargement of part of Figure 9. Depicted are the homoclinic bifurcations A, B, C, and D, the two saddle-node bifurcations of equilibria SN, the two saddle-node Hopf bifurcations SNH, the T point, and the two saddle-node bifurcations of limit cycles SL which terminate at the degenerate Hopf bifurcations marked with squares.*

branches follow branch A toward the SNH point but turn around sharply just before this point and return to the T-point again.

3.3. Model 3: Sneyd, LeBeau, and Yule four variable model. The SLY-4 model, studied in [41], is given by the following system of equations:

$$\begin{aligned} \dot{c} &= d, \\ D_c\dot{d} &= sd - k_f y^4 + \frac{V_p c^2}{K_p^2 + c^2} - J_{\text{leak}}, \\ s\dot{x} &= p\varphi_1 x - \varphi_{-1} y - \varphi_2 y, \\ s\dot{y} &= -(p\varphi_1 x - \varphi_{-1} y) + \varphi_3(1 - x - y). \end{aligned} \tag{3.3}$$

The variable h in SLY-3 is related to x and y in SLY-4 by $h = x + y$, with all other variables, parameters, and expressions in SLY-4 the same as for SLY-3, as detailed in Appendix A.

A partial bifurcation set for this system is shown in Figure 11. The bifurcations seen are similar to those in the FitzHugh–Nagumo system. In particular, the C-shaped curve is a homoclinic banana as for the FitzHugh–Nagumo equations. There is also a saddle-node bifurcation of periodic orbits emanating from a degenerate Hopf bifurcation, and homoclinic bifurcations of periodic orbits. Figure 11(b) shows bifurcations of one branch of periodic orbits associated with SLY-4; compare with Figure 6.

This example does not exhibit any new mechanisms for termination of the homoclinic branches, but does show that two closely related models (SLY-3 and SLY-4) can have completely different termination mechanisms for the branch of homoclinic bifurcations. Why this should be so is not clear. Nevertheless, it is apparent that minor changes in model structure, ones that have little to no significant effect on model behavior over the vast majority of parameter space, can cause significant qualitative changes in certain sensitive areas of parameter

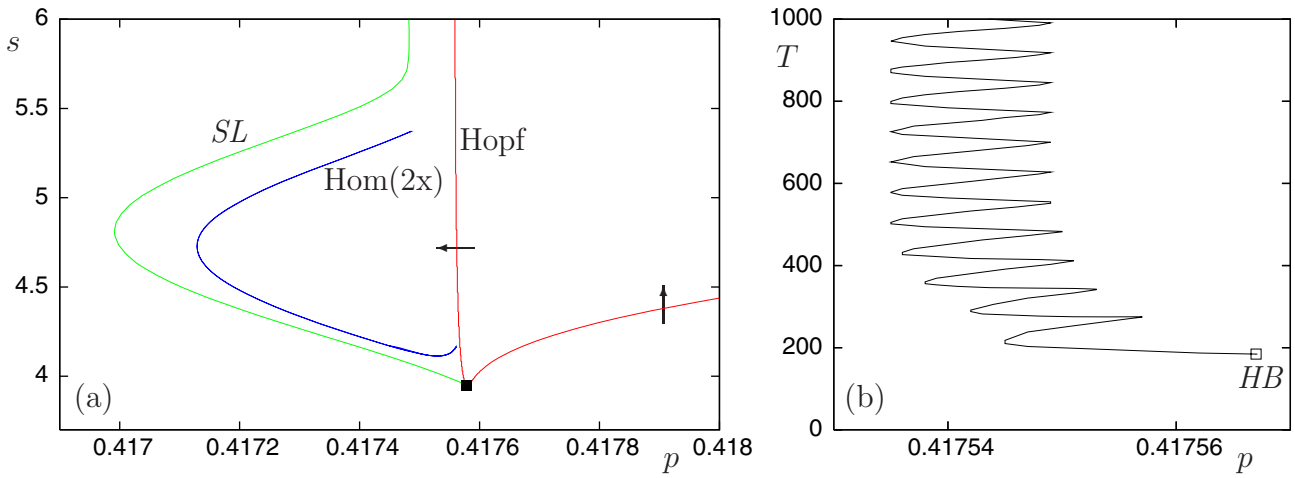


Figure 11. (a) *A partial bifurcation set for the SLY-4 model, equations* (3.3). *The curve labeled Hom* $(2x)$ *represents a homoclinic banana. The square on the Hopf curve marks a degenerate Hopf bifurcation. All other curves are saddle-node bifurcations of limit cycles. The arrows through the Hopf locus indicate the direction in which periodic orbits appear on the various sections of the Hopf bifurcation curve.* (b) *Bifurcation diagram for the SLY-4 model, for* $s = 4.17$, *showing the period* T *of the periodic orbits that emanate from the Hopf bifurcation in panel* (a).

space. Although it is unlikely that such changes in the bifurcation structure would be reflected in experimental results, or lead to testable predictions, it is still necessary to understand the complete range of possibilities in order to develop a proper mathematical understanding of the CU structure.

3.4. Model 4: A nine-dimensional calcium model. The equations we consider in this section come from a model for calcium wave propagation in pancreatic acinar cells developed in [40] and [43] and further studied in [38]. After passage to a moving frame of reference, we obtain the traveling wave equations for the model:

$$
\begin{aligned}
\dot{c} &= d,\\
D_c\dot{d} &= sd - J,\\
s\dot{c_e} &= \gamma(J_{\text{serca}} - (k_f P_{\text{IPR}} + \nu_1 P_{\text{ryr}} + J_{\text{er}})(c_e - c)),\\
s\dot{R} &= \varphi_{-2}O - \varphi_2 pR + (k_{-1} + l_{-2})I_1 - \varphi_1 R,\\
s\dot{O} &= \varphi_2 pR - (\varphi_{-2} + \varphi_4 + \varphi_3)O + \varphi_{-4}A + k_{-3}S,\\
s\dot{A} &= \varphi_4 O - (\varphi_{-4} + \varphi_5)A + (k_{-1} + l_{-2})I_2,\\
s\dot{I_1} &= \varphi_1 R - (k_{-1} + l_{-2})I_1,\\
s\dot{I_2} &= \varphi_5 A - (k_{-1} + l_{-2})I_2,\\
s\dot{w} &= k_{\text{cm}}(w_\infty - w)/w_\infty,
\end{aligned}
\tag{3.4}
$$

where $J \equiv (k_f P_{\text{IPR}} + \nu_1 P_{\text{ryr}} + J_{\text{er}})(c_e - c) - J_{\text{serca}} - J_{\text{mito}} + \delta(J_{\text{in}} - J_{\text{pm}})$. Here c_e denotes the concentration of calcium ions in the ER, R denotes the fraction of receptors in various states,

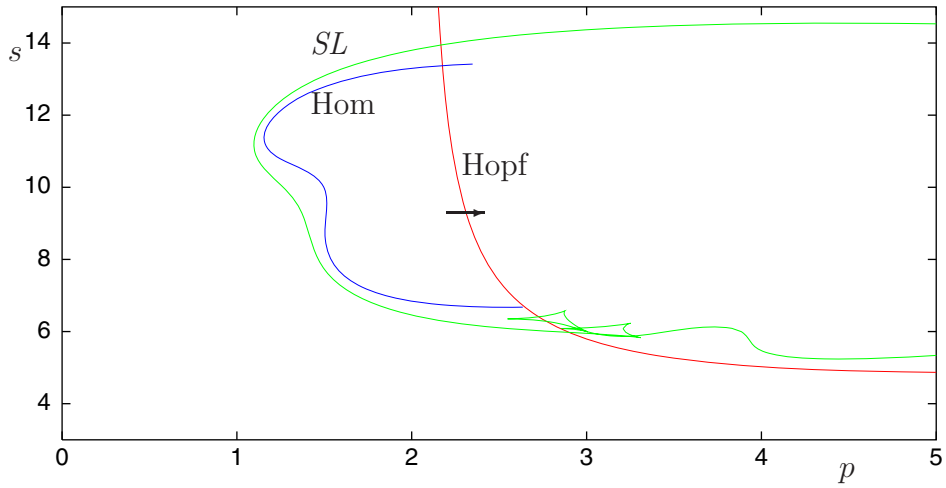


Figure 12. *Partial bifurcation set for the nine-dimensional calcium model, equations (3.4). The C curve of homoclinic bifurcations appears to cross the Hopf bifurcation curve at one end, but this is believed to be a numerical artifact as explained in the text. At the other end, the C curve turns around sharply at the first in a sequence of sharp turns which are not visible in this figure. There are no degenerate points on the Hopf locus; the saddle-node bifurcation of limit cycles (curve marked SL) does not terminate in this figure.*

and $\delta = 0.1$ controls the magnitude of transmembrane fluxes relative to the trans-ER fluxes without resting $[Ca^{2+}]$. The other parameters and constants are given in Appendix A.

Figure 12 shows a partial bifurcation set for this system of equations. At the lower end of the C-shaped homoclinic locus, the homoclinic branch has a sharp turning point, with the branch doubling back on itself. Unlike the case in the earlier example systems in which the turning point caused a homoclinic banana, here we have something topologically very different; see Figure 13. Here the homoclinic branch follows an (infinite) sequence of turning points that accumulate on two extreme values of (p, c). The turning points actually separate into six sets of points with four intermediate sets of folds between the extrema (which are the only sets that are easily visible in Figure 13(a)).

As we move through one complete cycle of six folds, the homoclinic orbit gains an extra large loop. As this process continues the orbit transforms from a primary Shil'nikov homoclinic orbit (Figure 13(b)) to something that increasingly resembles a heteroclinic cycle between the equilibrium point and a periodic orbit (Figure 13(c)). Note that Figure 13(a) shows the L_2-norm computed by AUTO as a function of p; in a plot of c versus p all the separate folds of the branch fall on top of one another to within numerical accuracy. The AUTO L_2-norm is defined as $(1/T)\int_0^T \|u\|^2 dt$, where T is a fixed large time over which the homoclinic orbit is computed. Since the equilibrium has a higher L_2 value than the periodic orbit, the L_2-norm for fixed T *decreases* in Figure 13(a) as more turns are added around the periodic orbit. Behavior of this type is explained in [9]. Each of the accumulation points of the folds of the homoclinic locus corresponds to an EP1t point, and the global connectedness of the locus is due to the topology of the intersection between the stable manifold of the equilibrium and the unstable manifold of the periodic orbit.

In Figure 12 the top end of the homoclinic locus seems to pass straight through the Hopf

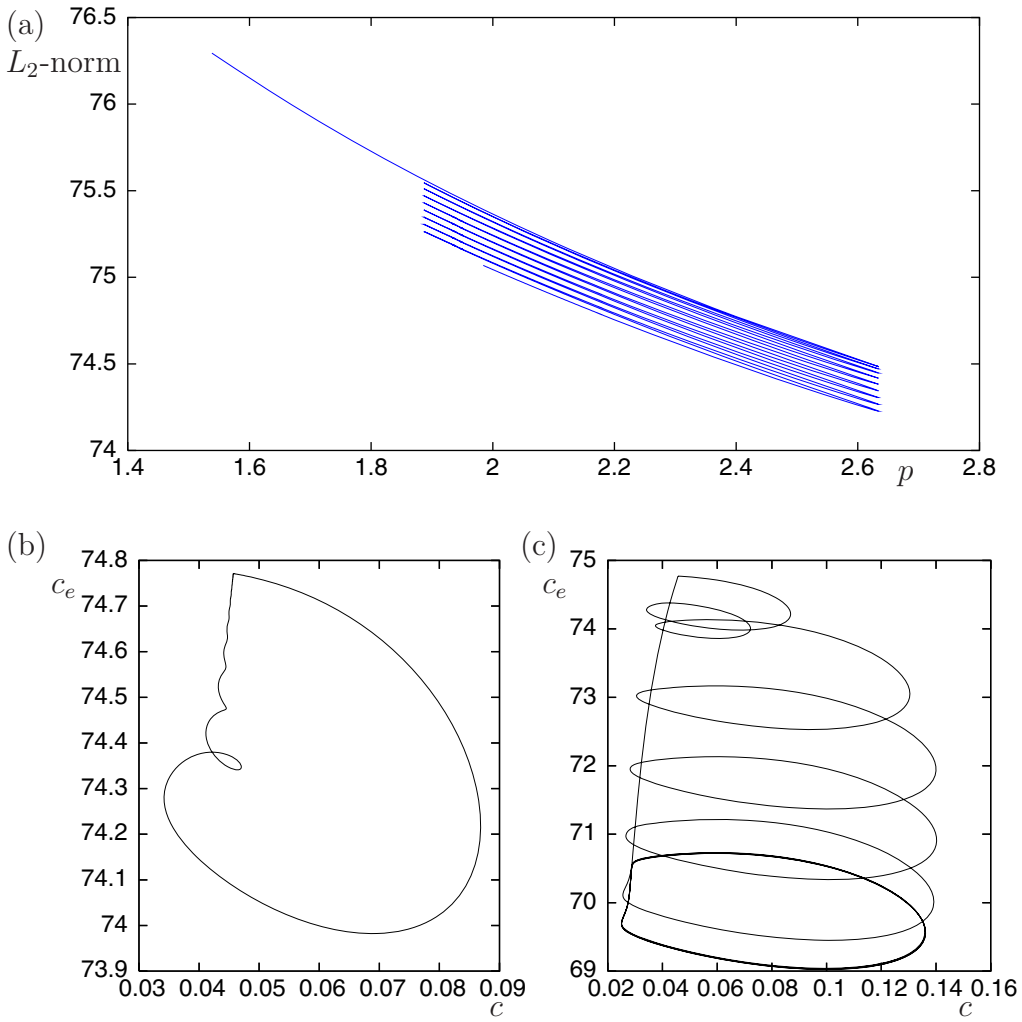


Figure 13. (a) *Detail along the lower part of the homoclinic branch in Figure 12, showing the L_2-norm of the homoclinic orbit as a function of p. The turning points on the homoclinic branch initially approximately alternate between two locations, with p values given approximately by* 2.63466, 1.88681, 2.63383, 1.88681, 2.63383, *and* 1.88681, *and then follow a repeated sequence of six values of p:* 2.63383, 1.93415, 1.93954, 1.88681, 1.95014, 1.94495. *Not all turning points can be distinguished in the figure. Below: Homoclinic orbits near the lower end of the homoclinic branch.* (b) *The homoclinic orbit before going through any turning points, at $(p, s) = (2.4, 6.68342)$;* (c) *the homoclinic orbit at approximately the same place in parameter space, at $(p, s) = (2.4, 6.68342)$ after going through* 17 *outer turning points.*

bifurcation curve and terminate at $p \approx 2.3$! However, we believe this to be a numerical artifact and that the homoclinic curve actually terminates at a Shil'nikov/Hopf bifurcation when the homoclinic locus reaches the Hopf locus. This numerical anomaly is due to the slow-fast nature of the system; a detailed discussion is deferred to section 3.6.

3.5. Model 5: A four-dimensional simplification. Assuming fast equilibration between all the receptor states except R and A, we can reduce (3.4) to a model with only a single gating variable, but one whose steady-state behavior is equivalent to (3.4):

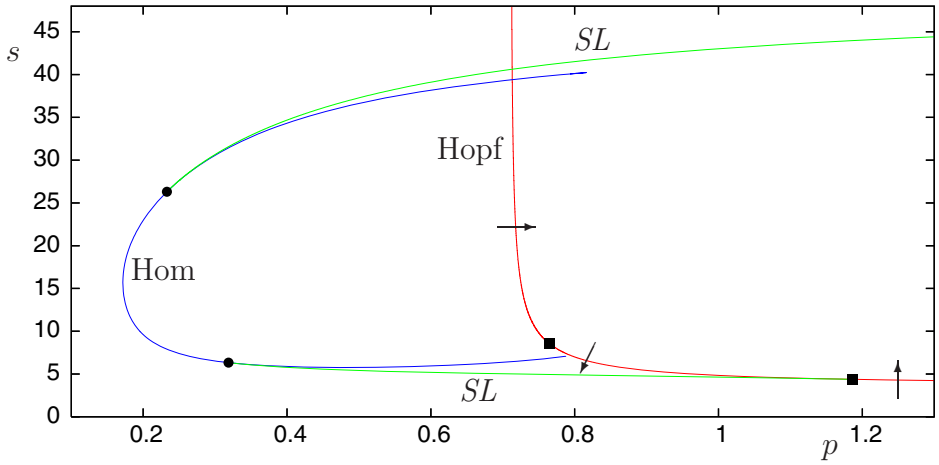


Figure 14. *Partial bifurcation set for the four-dimensional calcium model, equations (3.5). The homoclinic bifurcation curve appears to cross the Hopf locus at the top end, but this is believed to be a numerical artifact, as discussed in the text. At the lower end the homoclinic locus goes through a sequence of sharp turning points, eventually converging at a codimension-two point at $p = 0.7682935$. The two curves of saddle-node bifurcations of periodic orbits (marked SL) terminate on the homoclinic bifurcation curve at the large dots where the leading eigenvalues of the corresponding equilibrium resonate; the saddle quantity of the saddle vanishes at these points. The lower SL bifurcation terminates at its right end at a degenerate Hopf bifurcation, indicated by a square. A second degenerate Hopf bifurcation lies to the left, just above the lower end of the C curve; the SL curve emanating from this degenerate Hopf bifurcation is not shown. The arrows indicate the direction of the Hopf bifurcation as in Figure 7.*

$$
\begin{aligned}
\dot{c} &= d,\\
D_c\dot{d} &= sd - ((J_{\mathrm{IPR}} + J_{\mathrm{er}})(c_e - c) - J_{\mathrm{serca}} + \delta(J_{\mathrm{in}} - J_{\mathrm{pm}})),\\
\dot{c_t} &= \delta(J_{\mathrm{in}} - J_{\mathrm{pm}}),\\
s\dot{R} &= -\varphi_5 A + (k_{-1} + l_{-2})(1 - R).
\end{aligned}
\tag{3.5}
$$

Here the variable c_t is related to the variables in (3.4) via the equation $c_t = sc_e/\gamma - D_c d + sc$. All other variables, parameters, and constants in these equations have the same meaning as in (3.4).

Figure 14 shows a partial bifurcation set for (3.5). The top end of the homoclinic branch has similar behavior to that seen in the nine-dimensional model, with the numerical approximation to the homoclinic locus crossing the Hopf bifurcation curve. Again, we believe this to be a numerical artifact but return to a fuller discussion of the phenomenon in section 3.6.

Behavior near the lower end of the homoclinic curve is different in this system than in the nine-dimensional model. As before, the branch initially turns around sharply, and there follows a series of further sharp turns, but the turning points converge to a single point rather than to two (see Figure 15). Again, as in the nine-dimensional model, the homoclinic orbit looks increasingly like a heteroclinic cycle between the equilibrium and a periodic orbit as we move along the folded homoclinic branch (see Figures 15(b) and 15(c)). We conjecture that the turning points in the homoclinic bifurcation curve are organized here by a codimension-two heteroclinic bifurcation involving the equilibrium and the periodic orbit. The oscillatory

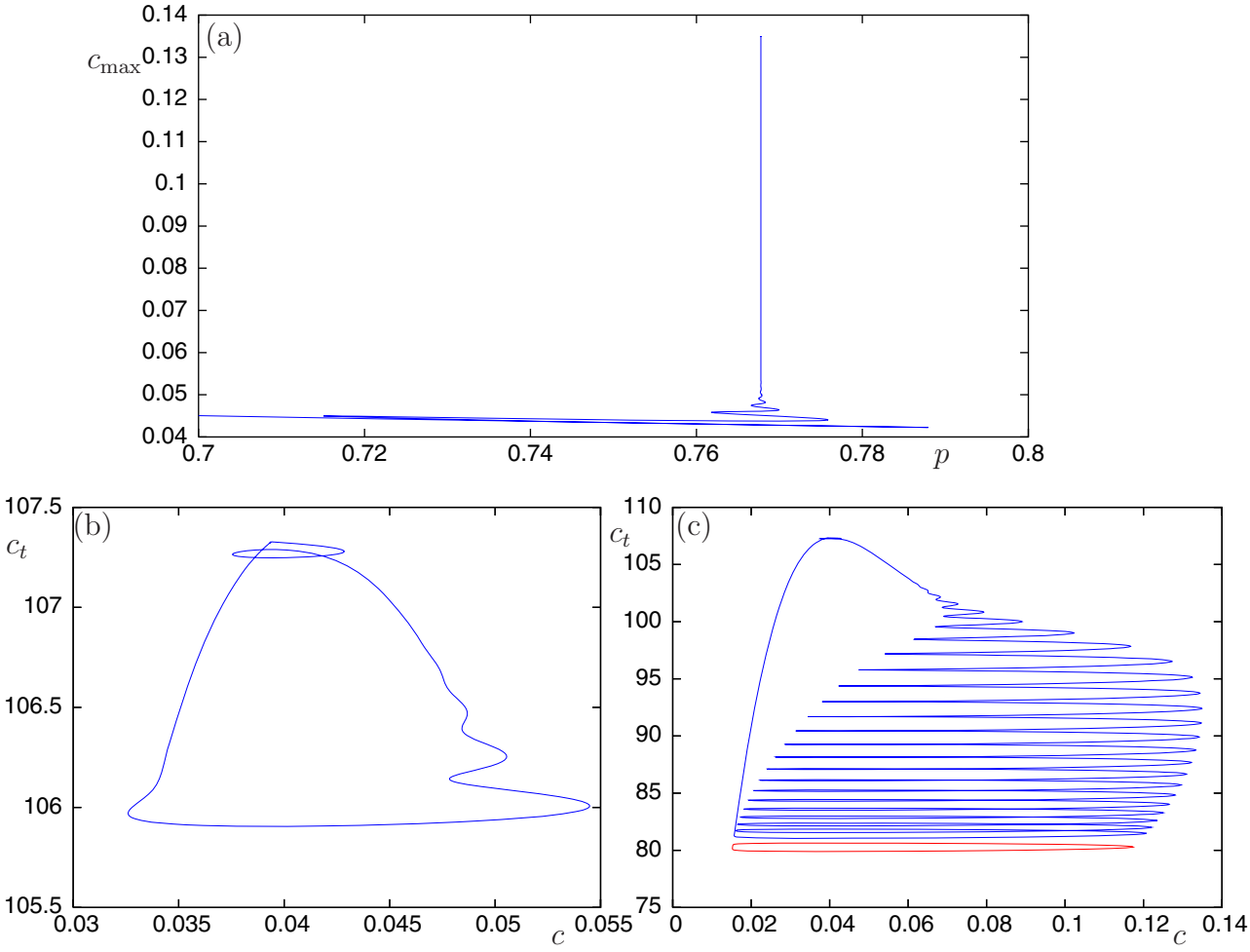


Figure 15. (a) *The lower part of the homoclinic bifurcation curve in the four-dimensional calcium model, equations* (3.5)*, showing the maximum value of* c *on the homoclinic orbit,* $c_{\max}$*, as a function of* p*. After many turns* $c_{\max}$ *stabilizes at* 0.134922*.* (b) *The homoclinic orbit at* $p = 0.7682935$ *after* 15 *turns of the homoclinic bifurcation curve.* (c) *The homoclinic orbit at* $p = 0.7682935$ *after many turns of the homoclinic bifurcation curve. Also plotted (red curve) is a periodic orbit that bifurcates from the right arm of the U-shaped Hopf locus. This orbit was obtained by continuing in* p *for* s *fixed from the Hopf bifurcation point at* $(p, s) = (22.04131, 6.878683)$*.*

approach to this codimension-two point is reminiscent of an EP2 point (case (d) in Figure 4). However, at the time of writing it is not clear whether this is the case, or whether we have another example of the phenomenon in Figure 13, explained by the presence of several EP1t points, with remarkably thin wiggles at large periods.

The chief difference between the EP1t and EP2 heteroclinic cycles is the dimension of the stable manifold of the periodic orbit. Indeed, the rate of convergence of the turning points near either an EP2 or an EP1t point was found in [34] to depend on the Floquet multipliers of the periodic orbit. However, due to the stiffness of the equations we have been unable to calculate these multipliers accurately enough to compare numerics with the theory. (AUTO

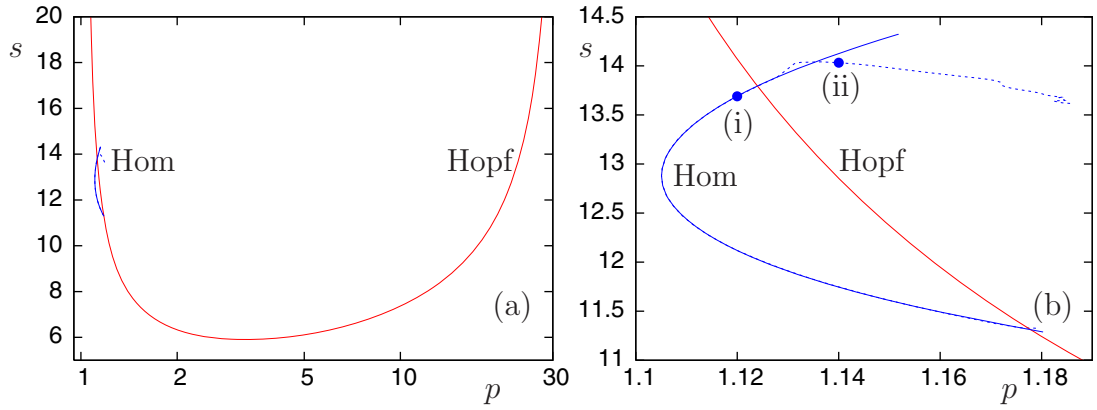


Figure 16. (a) *Partial bifurcation set for the five-dimensional model, equations* (3.6). (b) *An enlargement of part of* (a). *Homoclinic and Hopf bifurcations are labeled. The points labeled* (i) *and* (ii) *in* (b) *correspond to the parameters used for the phase portraits in Figure* 17. *In* (b) *two different algorithms are used to compute the homoclinic curve—projection boundary conditions (solid line) and fixed period-*5000 *periodic orbit (dashed line). To the left of the Hopf line these two curves are overlaid to the accuracy depicted.*

calculations give the Floquet multipliers of the periodic orbit as $1, 0.28, 8 \times 10^{-7}, -1 \times 10^{6}$, but these multipliers cannot be trusted since the signs are clearly wrong. However, from these calculations it appears that the periodic orbit, as followed from the Hopf bifurcation, does not undergo any bifurcations and has consistently two multipliers strictly inside the unit circle, suggesting an EP1t point.)

A significant difference between this case and the nine-dimensional model is that in these equations the Hopf bifurcation is subcritical near the lower end of the homoclinic branch, producing a periodic orbit to the left of the Hopf locus, although it is not clear whether this has any bearing on the question of how the homoclinic branch terminates.

3.6. Model 6: A five-dimensional variant. The model studied in this section is another simplification of the nine-dimensional model described in section 3.4; instead of using a six-state model of the IPR as in (3.4), all but three of the states are assumed to be in instantaneous equilibrium, and a fast time-scale reduction is applied to simplify the model. The reduced, five-dimensional model is given by the following set of equations:

$$
\begin{aligned}
\dot{c} &= d, \\
D_c\dot{d} &= sd - [(k_{\mathrm{f}}P_{\mathrm{IPR}} - J_{\mathrm{er}})(c_e - c) - J_{\mathrm{serca}} + \delta(J_{\mathrm{in}} - J_{\mathrm{pm}})], \\
s\dot{c_e} &= -\gamma\left[(k_{\mathrm{f}}P_{\mathrm{IPR}} - J_{\mathrm{er}})(c_e - c) - J_{\mathrm{serca}}\right], \\
s\dot{O} &= \phi_2 pR - \theta_1 O + (k_{-1} + l_{-2})I_2 - \theta_2 O, \\
s\dot{R} &= \theta_1 O - \phi_2 pR.
\end{aligned}
\tag{3.6}
$$

All parameters and constants in these equations are given in Appendix A. This system of equations was studied in [22, 23], where a partial bifurcation diagram was obtained (Figure 16). In this figure, the homoclinic bifurcation curve appears to overshoot the Hopf bifurcation curve at both ends by a significant amount, in a similar way to that seen near the upper ends of

the C-shaped curves in Figures 12 and 14. (Note that the Hopf bifurcation is supercritical for all values of p and s in this figure.) Since there is only one equilibrium solution to (3.6), there must be a homoclinic-Hopf bifurcation of the equilibrium solution at the point where the homoclinic and Hopf bifurcation curves coincide. Just to the left of the Hopf bifurcation locus, we generically expect that the homoclinic orbit will enter the equilibrium solution along a direction tangent to the center manifold. The dynamics associated with such a homoclinic-Hopf bifurcation in a three-dimensional system is described in [26], while [13] discusses the case in three or higher dimensions. In both papers it is shown that the homoclinic bifurcation curve terminates at the Hopf locus. To understand why the numerically computed homoclinic bifurcation curve shown in Figure 16 overshoots the Shil'nikov–Hopf point it is necessary to look a little closer at the algorithms used to obtain the numerical approximation to the homoclinic locus. We do so below but note that much of this discussion was also contained in [23].

Homoclinic bifurcation curves like those shown in Figure 16 can be obtained using any algorithm that sets up well-posed boundary conditions after truncation to a (large) finite time interval $t \in (0, T)$. For example, HomCont poses projection boundary conditions based on eigenvectors of the adjoint linearized problem about the equilibrium point. The dimensions of the stable and unstable subspaces are fixed at the beginning of the computation. In a case such as ours, where these dimensions change midcomputation (as the Hopf locus is crossed), HomCont will continue to compute an orbit that satisfies the specified boundary conditions even though the boundary conditions are no longer appropriate. Thus, if using HomCont to compute the approximate homoclinic locus for (3.6), we must ignore any part of the computed curve located inside the U-shaped Hopf locus; the computed curve in this region apparently approximates a heteroclinic connection from the equilibrium to the Hopf periodic orbit rather than a homoclinic bifurcation of the equilibrium. More details on the HomCont algorithm can be found in [12].

An alternative approximation to the homoclinic locus may be obtained using periodic boundary conditions, and following a periodic orbit of fixed high period (say, period 5000 or higher). This method is known to be less accurate than using projection boundary conditions, but nevertheless converges as $T \to \infty$ to the correct result. The approximate homoclinic locus for (3.6) obtained with this method is indicated by the dashed curve in Figure 16(b); like the case with projection boundary conditions, this curve also crosses the Hopf locus at each end. However, examination of the phase portraits at various points on the curve suggests that the locus of homoclinic bifurcations of the equilibrium terminates at the Hopf locus, just as expected from a Shil'nikov–Hopf bifurcation. Panels (a) and (b) in Figure 17 show the period 5000 orbit approximating the homoclinic orbit at a point to the left of the Hopf bifurcation (point (i) in Figure 16(b)), while panels (c) and (d) show the orbit at point (ii) in Figure 16(b). It is clear from the figure that after the Hopf bifurcation, instead of simply spiraling in toward the equilibrium point, the computed orbit spirals *out* toward the periodic orbit (shown as a red curve in Figure 16). Thus we are actually computing an approximation to a heteroclinic connection from the equilibrium to the periodic orbit created in the Hopf bifurcation, rather than a true homoclinic orbit to the equilibrium. In this case, it appears that high-period orbits persist beyond the Hopf bifurcation even though the homoclinic bifurcation terminates at the Hopf locus.

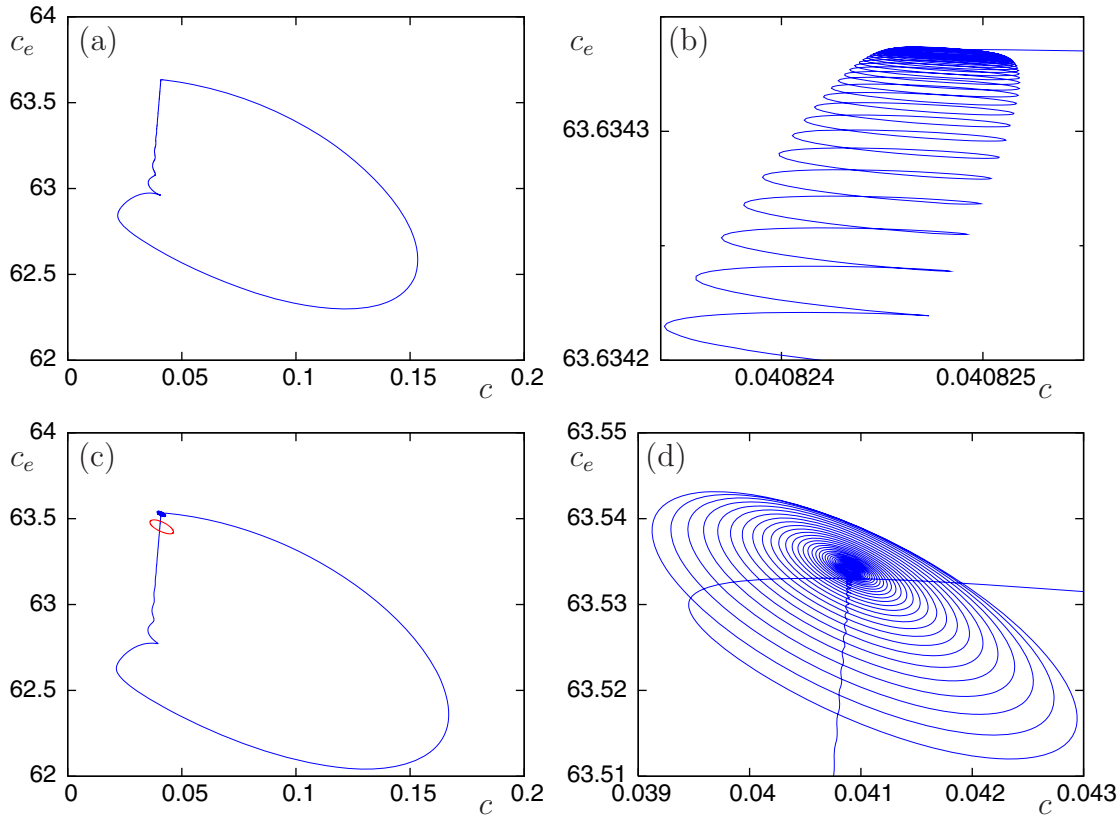


Figure 17. *Phase portraits at the points labeled* (i) *and* (ii) *in Figure 16, computed with AUTO by following an orbit of fixed period (period* 5000*) with AUTO constant* `NTST` = 1000. (a) *At point* (i)*;* (b) *enlargement of* (a) *near the equilibrium.* (c) *At point* (ii). *The red curve is the periodic orbit created in the Hopf bifurcation;* (d) *enlargement of* (c) *near the equilibrium.*

From detailed examination of the phase portraits for this model it is evident that the true homoclinic locus terminates at the Hopf bifurcation at both the upper and lower ends. These are just regular Shil'nikov–Hopf bifurcation points (case (a) of Figure 4). We conjecture that exactly the same process occurs in the nine-dimensional and four-dimensional models (3.4) and (3.5), at the upper end of the C-branch, where the homoclinic locus appears to cross the Hopf bifurcation curve. These equations are even more stiff than the present five-dimensional model, and even detailed examination of the phase portraits fails to reveal that the computation of the homoclinic branch has become spurious upon crossing the Hopf U. The difficulty in these cases appears to be the presence of a slow manifold in the associated phase space. Most models of calcium dynamics, including all the CU-systems studied here, have slow manifolds that are important in determining the dynamics. (A comprehensive review of models of calcium waves is given in [16], while a detailed discussion of slow manifolds in this context, as well as in the context of more generic excitable system theory, can be found in [28].)

In the three models (3.4), (3.5), and (3.6) the slow manifold is one-dimensional and passes

through the single equilibrium, with the motion on the slow manifold being toward the equilibrium. A homoclinic orbit in these systems typically leaves the equilibrium along the unstable manifold and makes a global excursion, before being drawn into a tiny tube around the slow manifold. It then follows the slow manifold toward the equilibrium until it gets close enough that the dynamics on the center manifold takes over. If the equilibrium solution is attracting within the center manifold, the orbit will then approach the equilibrium solution in a direction tangent to the center manifold, thereby closing the homoclinic connection. If, however, the equilibrium has undergone a supercritical Hopf bifurcation and is unstable, it can take a remarkably long time before the trajectory along the slow manifold reaches sufficiently close to the equilibrium to experience this instability. As a result, if one uses a high-period periodic orbit approximation to the homoclinic orbit, with insufficiently high period, the computed orbit may not reach the center manifold at all before leaving the equilibrium. In such a case, the dominant feature of the center manifold dynamics, namely, outward spiraling from the equilibrium, will not be observed.

Based on the above discussion, we conjecture that the mechanism for the termination of the upper end of the C-shaped curves in (3.4) and (3.5) and both ends of the C-curve in (3.6) is simply the Shil'nikov–Hopf bifurcation.

4. Discussion. We have studied a hierarchy of detailed models of excitable systems, focusing in particular on those that model calcium dynamics. The main aim of the work has been to understand the connection between two different dynamical regimes for which traveling-wave–type behavior is observed in these systems, and in particular to look at the interaction between small-amplitude waves born in a Hopf bifurcation in the traveling wave equations, and solitary pulses represented by homoclinic orbits in the traveling wave equations. We have found that this interaction can occur in a variety of different ways, despite broad similarities in behavior in the models we have analyzed. It seems that there are many different ways in which Shil'nikov chooses to meet Hopf!

Specifically, we have sought codimension-two mechanisms for the termination of homoclinic bifurcation curves as they approach a Hopf bifurcation. In our examples we have identified cases of T-points, saddle-node/Hopf local bifurcations, and equilibrium-to-periodic heteroclinic bifurcations. We also found Shil'nikov–Hopf bifurcations in a number of cases; these bifurcations were unusually numerically delicate because of the degree of stiffness in the model equations, with the consequence that in each case the homoclinic bifurcation curve appeared at first analysis to pass through the Hopf locus instead of terminating there, as theory would predict. We argued that this behavior is a numerical artifact and gave a plausibility argument as to why even very careful numerics near Shil'nikov–Hopf bifurcations can be misleading whenever slow manifolds are present. Proper multiple time-scale analysis of this behavior is left for future work.

The stiffness of the models we studied most likely plays an important role in other terminating mechanisms besides the Shil'nikov–Hopf bifurcation. For instance, we found that rapid, canard-like growth of oscillations occurs near the blunt end of the homoclinic banana in the FitzHugh–Nagumo equations. An investigation of the role of canards in calcium models is the subject of ongoing work.

A number of our models contained an apparent codimension-two termination mechanism, where a codimension-one homoclinic bifurcation curve appeared to terminate "in mid air." In

each case, it turned out that the bifurcation curve was not actually terminating but rather going through a sharp turn and doubling back on itself, but with the doubled-back curve indistinguishable from the original curve on the scale of the bifurcation set; typically the homoclinic orbit gained an extra loop in the process so that a single-pulse orbit became a double-pulse orbit. Such sharp turns require no special explanation in terms of bifurcation theory but are important within our taxonomy of termination mechanisms since they arise naturally as a consequence of some of the other mechanisms, as described earlier.

We do not, at present, fully understand the significance of the sub- or supercriticality of the Hopf bifurcation in these systems. A better understanding may follow from forthcoming theoretical work on homoclinic orbits to degenerate Hopf points [10].

We have not considered here the PDE stability of the traveling wave structures we have identified. Generally in these types of models it is the upper branch of the C-curve of homoclinic bifurcations that is of physiological interest as this corresponds (in the PDE) to stable pulses while the other branches correspond to unstable pulses. The significance of the details of our codimension-two termination mechanisms for the PDE dynamics is unclear. For instance, given the strong contractions due to the slow-fast nature of the dynamics, the precise details of the spurious passage through the Shil'nikov–Hopf bifurcations may not be important; something that is exponentially close to a homoclinic orbit on the "wrong side" of the Hopf bifurcation may in fact imply the existence of an almost steady traveling pulse for exponentially long times. It would be of interest, therefore, to see if the Hopf bifurcation in these cases might not be "felt" at all in PDE computations of stable traveling pulses over finite time scales, in the same way that its presence was almost undetectable when tracing paths of homoclinic bifurcations to the traveling wave ODEs.

The detailed numerics we have presented are for two distinct types of excitable systems, the FitzHugh–Nagumo equations, which might be regarded as the canonical excitable system, and a class of models of the dynamics of intracellular calcium. We expect that many of the features we have found will also be relevant to other classes of excitable systems—for example, models of heart tissue and neural oscillators.

Appendix A. Parameters and formulae for model definitions. The following constants and quantities were used in numerical simulations of models 2 and 3 (SLY-3 and SLY-4):

k_f	J_{leak}	k_1	k_2	k_3	k_{-1}	V_p	K_p	r_2	r_4	r_6	r_{-2}	R_1	R_3	R_5	D_c
28	0.2	0	0.53	1	0.88	1.2	0.18	100	20	0	0	6	50	1.6	25

$$\begin{aligned}\varphi_{-1} &= (k_{-1} + r_{-2})R_3/(c + R_3),\\ \varphi_1 &= (k_1 R_1 + r_2 c)/(R_1 + c),\\ \varphi_2 &= (k_2 R_3 + r_4 c)/(R_3 + c),\\ \varphi_3 &= (k_3 R_5 + r_6 c)/(R_5 + c).\end{aligned}$$

The following constants and quantities were used in numerical simulations of models 4, 5, and 6 (nine-, four-, and five-dimensional models):

J_{inbase}	k_f	J_{er}	γ	V_{serca}	K_{serca}	V_{pm}	K_{pm}	δ	D_c	ν_1	V_{mito}
0.2	0.32	0.002	5.405	120.0	0.18	28.0	0.425	0.1	20	0.04	0

In the four-dimensional model 5 we used the values $D_c = 25$ and $k_f = 0.4$ instead of the above values.

k_{ap}	k_{am}	k_{bp}	k_{bm}	k_{cp}	k_{cm}
1500	28.8	1500	385.9	1.75	0.1

k_1	k_{-1}	k_2	k_{-2}	k_3	k_{-3}	k_4	k_{-4}	l_2	l_6	L_1	L_5	l_4	L_3
0.64	0.04	37.4	1.4	0.11	29.8	4	0.54	1.7	4707	0.12	54.7	1.7	0.025

$$K_a = (k_{am}/k_{ap})^{\frac{1}{4}}, \qquad K_b = (k_{bm}/k_{bp})^{\frac{1}{3}},$$

$$K_c = k_{cm}/k_{cp}, \qquad w_\infty = \frac{1 + (K_a/c)^4 + (c/K_b)^3}{1 + 1/K_c + (K_a/c)^4 + (c/K_b)^3},$$

$$J_{mito} = V_{mito}\frac{c^3}{1+c^2}, \qquad P_{ryr} = w\frac{1 + (c/K_b)^3}{1 + (K_a/c)^4 + (c/K_b)^3},$$

$$P_{IPR}(O, A) = (O/10 + 9A/10)^4 \text{ (model 4)}, \qquad P_{IPR}(O) = O^4\theta_3 \text{ (model 6)},$$

$$S = 1 - R - O - A - I_1 - I_2, \qquad J_{in} = J_{inbase} + 0.05p,$$

$$l_{-2} = l_2k_{-1}/(k_1L_1) \text{ (models 4 and 5)}, \qquad l_{-2} = 0.8 \text{ (model 6)},$$

$$l_{-6} = k_{-4}l_6/(k_4L_5) \text{ (models 4 and 5)}, \qquad l_{-6} = 11.4 \text{ (model 6)},$$

$$l_{-4} = k_{-2}l_4/(k_2L_3) \text{ (models 4 and 5)}, \qquad l_{-4} = 2.5 \text{ (model 6)},$$

$$\varphi_1(c) = \frac{(k_1L_1 + l_2)c}{L_1 + c(1 + L_1/L_3)}, \qquad \varphi_2(c) = \frac{k_2L_3 + l_4c}{L_3 + c(1 + L_3/L_1)},$$

$$\varphi_{-2}(c) = \frac{k_{-2} + l_{-4}c}{1 + c/L_5}, \qquad \varphi_3(c) = \frac{k_3L_5}{L_5 + c},$$

$$\varphi_4(c) = \frac{(k_4L_5 + l_6)c}{L_5 + c}, \qquad \varphi_{-4}(c) = \frac{L_1(k_{-4} + l_{-6})}{L_1 + c},$$

$$\varphi_5(c) = \frac{(k_1L_1 + l_2)c}{L_1 + c},$$

$$\Phi_2 = \varphi_{-2}/\varphi_2, \qquad \Phi_4 = \varphi_{-4}/\varphi_4,$$

$$A = pR/(p + \Phi_2\Phi_4 + p\Phi_4), \qquad J_{IPR} = k_fA^4,$$

$$J_{pm}(c) = \frac{V_{pm}c^2}{K_{pm}^2 + c^2}, \qquad J_{serca}(c, c_e) = \frac{V_{serca}c}{c_e(K_{serca} + c)},$$

$$\theta_1(c) = \frac{0.5 + 0.5c^2}{0.36 + 20c^2}, \qquad \theta_2(c) = \frac{8.1c^2}{0.2 + 5c^2},$$

$$\theta_3(c) = \frac{0.00074 + 8.6c^2}{0.74 + 13c^2}, \qquad I_2 = 1 - R - O \text{ (model 6 only)}.$$

Acknowledgments. The authors acknowledge helpful conversations with Elan Gin, Jerry Marsden, Jens Rademacher, Björn Sandstede, and Martin Wechselberger and thank Georg Gottwald for pointing out the reason the CU-structure is a common feature of excitable systems.

REFERENCES

[1] L. A. Belyakov, *A case of the generation of a periodic motion with homoclinic curves*, Mat. Zametki, 15 (1974), pp. 336–341 (in Russian).

[2] L. A. Belyakov, *The bifurcation set in a system with a homoclinic saddle curve*, Mat. Zametki, 28 (1980), pp. 910–916 (in Russian).

[3] L. A. Belyakov, *Bifurcation of systems with homoclinic curve of a saddle-focus with saddle quantity zero*, Mat. Zametki, 36 (1984), pp. 838–843 (in Russian).

[4] M. J. Berridge, M. D. Bootman, and H. L. Roderick, *Calcium signalling: Dynamics, homeostasis and remodelling*, Nat. Rev. Mol. Cell. Biol., 4 (2003), pp. 517–529.

[5] H. W. Broer and G. Vegter, *Subordinate Shil'nikov bifurcations near some singularities of vector fields having low codimension*, Ergodic Theory Dynam. Systems, 4 (1984), pp. 509–525.

[6] V. V. Bykov, *Bifurcations of dynamical systems close to systems with a separatrix contour containing a saddle-focus*, in Methods of the Qualitative Theory of Differential Equations, E. A. Leontovich-Andronova, ed., Gor'kov. Gos. Univ., Gorki, Russia, 1980, pp. 44–72.

[7] V. V. Bykov, *The bifurcations of separatrix contours and chaos*, Phys. D, 62 (1993), pp. 260–299.

[8] A. R. Champneys and V. Kirk, *The entwined wiggling of homoclinic curves emerging from saddle-node/Hopf instabilities*, Phys. D, 195 (2004), pp. 77–105.

[9] A. R. Champneys, V. Kirk, E. Knobloch, B. Oldeman, and J. Rademacher, *Unfolding a Tangent Equilibrium-to-Periodic Heteroclinic Cycle*, in preparation, 2007.

[10] A. R. Champneys, V. Kirk, and E. Knobloch, *Blue Sky Catastrophes of Homoclinic Orbits*, in preparation, 2007.

[11] A. R. Champneys and Yu. A. Kuznetsov, *Numerical detection and continuation of codimension-two homoclinic bifurcations*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 4 (1994), pp. 785–822.

[12] A. R. Champneys, Yu. A. Kuznetsov, and B. Sandstede, *A numerical toolbox for homoclinic bifurcation analysis*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 6 (1996), pp. 867–887.

[13] B. Deng and K. Sakamoto, *Šil'nikov-Hopf bifurcations*, J. Differential Equations, 119 (1995), pp. 1–23.

[14] E. J. Doedel, A. R. Champneys, T. F. Fairgrieve, Y. A. Kuznetsov, B. Sandstede, and X. Wang, *AUTO* 97*: Continuation and Bifurcation Software for Ordinary Differential Equations*, http://indy.cs.concordia.ca/auto/.

[15] E. J. Doedel, R. C. Paffenroth, A. R. Champneys, T. F. Fairgrieve, Yu. A. Kuznetsov, B. E. Oldeman, B. Sandstede, and X. Wang, *AUTO* 2000*: Continuation and Bifurcation Software for Ordinary Differential Equations (with HomCont)*, http://cmvl.cs.concordia.ca/auto/.

[16] M. Falcke, *Reading the patterns in living cells—the physics of Ca^{2+} signaling*, Adv. Phys., 53 (2004), pp. 255–440.

[17] C. Fall, E. S. Marsland, J. M. Wagner, and J. J. Tyson, eds., *Computational Cell Biology*, Springer-Verlag, New York, 2002.

[18] R. FitzHugh, *Impulses and physiological states in theoretical models of nerve membrane*, Biophys. J., 1 (1961), pp. 445–446.

[19] P. Gaspard, *Local birth of homoclinic chaos*, Phys. D, 62 (1993), pp. 94–122.

[20] P. Gaspard and X.-J. Wang, *Homoclinic orbits and mixed-mode oscillations in far-from-equilibrium systems*, J. Statist. Phys., 48 (1987), pp. 151–199.

[21] N. K. Gavrilov and L. P. Shilnikov, *On three-dimensional systems close to systems with a structurally unstable homoclinic curve* II, Math. USSR-Sb., 19 (1973), pp. 139–156.

[22] E. Gin, *A Bifurcation Analysis of Calcium Buffering*, M.Sc. thesis, The University of Auckland, Auckland, New Zealand, 2005.

[23] E. Gin, V. Kirk, and J. Sneyd, *A bifurcation analysis of calcium buffering*, J. Theoret. Biol., 242 (2006), pp. 1–15.

[24] D. R. Giovannucci, J. I. Bruce, S. V. Straub, J. Arreola, J. Sneyd, T. J. Shuttleworth, and D. I. Yule, *Cytosolic Ca^{2+} and Ca^{2+}-activated Cl^- current dynamics: Insights from two functionally distinct mouse exocrine cells*, J. Physiol., 540 (2002), pp. 469–484.

[25] P. A. Glendinning and C. Sparrow, *T-points: A codimension two heteroclinic bifurcation*, J. Statist. Phys., 43 (1986), pp. 479–488.

[26] P. Hirschberg and E. Knobloch, *Šil'nikov-Hopf bifurcation*, Phys. D, 62 (1993), pp. 202–216.

[27] A. L. Hodgkin and A. F. Huxley, *A quantitative description of membrane current and its applications to conduction and excitation in nerve*, J. Physiol., 117 (1952), pp. 500–544.
[28] J. Keener and J. Sneyd, *Mathematical Physiology*, Springer-Verlag, New York, 1998.
[29] Yu. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Springer-Verlag, New York, 1995.
[30] Yu. A. Kuznetsov, O. De Feo, and S. Rinaldi, *Belyakov homoclinic bifurcations in a tritrophic food chain model*, SIAM J. Appl. Math., 62 (2001), pp. 462–487.
[31] K. Lust, *Improved numerical Floquet multipliers*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 11 (2001), pp. 2389–2410.
[32] J. S. Nagumo, S. Arimoto, and S. Yoshizawa, *An active pulse transmission line simulating nerve axon*, Proc. IRE, 50 (1962), pp. 2061–2070.
[33] B. E. Oldeman, A. R. Champneys, and B. Krauskopf, *Homoclinic branch switching: A numerical implementation of Lin's method*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 10 (2003), pp. 2977–2999.
[34] J. D. M. Rademacher, *Homoclinic orbits near heteroclinic cycles with one equilibrium and one periodic orbit*, J. Differential Equations, 218 (2005), pp. 390–443.
[35] M. M. Romeo and C. K. R. T. Jones, *The stability of travelling calcium pulses in a pancreatic acinar cell*, Phys. D, 177 (2003), pp. 242–258.
[36] M. Krupa, B. Sandstede, and P. Szmoylan, *Fast and slow waves in the FitzHugh-Nagumo equation*, J. Differential Equations, 133 (1997), pp. 49–97.
[37] J. Sim, *Saddle-Node/Hopf Bifurcations in a Physiological Model*, M.Sc. thesis, The University of Auckland, Auckland, New Zealand, 2002.
[38] D. Simpson, V. Kirk, and J. Sneyd, *Complex oscillations and waves of calcium in pancreatic acinar cells*, Phys. D, 200 (2005), pp. 303–324.
[39] L. P. Shilnikov, A. L. Shilnikov, D. V. Turaev, and L. O. Chua, *Methods of Qualitative Theory in Nonlinear Dynamics, Part* II, World Scientific, Singapore, 2001.
[40] J. Sneyd and J. F. Dufour, *A dynamic model of the type-2 inositol trisphosphate receptor*, Proc. Natl. Acad. Sci. USA, 99 (2002), pp. 2398–2403.
[41] J. Sneyd, A. LeBeau, and D. Yule, *Travelling waves of calcium in pancreatic acinar cells: Model construction and bifurcation analysis*, Phys. D, 145 (2000), pp. 158–179.
[42] J. Sneyd and K. Tsaneva-Atanasova, *Modeling calcium waves*, in Understanding Calcium Dynamics: Experiments and Theory, Lecture Notes in Phys. 623, Springer-Verlag, Berlin, 2003, pp. 179–199.
[43] J. Sneyd, K. Tsaneva-Atanasova, J. I. E. Bruce, S. V. Straub, D. R. Giovannucci, and D. I. Yule, *A model of calcium waves in pancreatic and parotid acinar cells*, Biophys. J., 85 (2003), pp. 1392–1405.
[44] A. V. Tepikin and O. H. Petersen, *Mechanisms of cellular calcium oscillations in secretory cells*, Biochem. Biophys. Acta, 1137 (1992), pp. 197–207.
[45] K. Yagasaki, *Numerical detection and continuation of homoclinic points and their bifurcations for maps and periodically forced systems*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 8 (1998), pp. 1617–1627.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 4, pp. 694–727

The Onset of Oscillations in Microvascular Blood Flow*

John B. Geddes†, Russell T. Carr‡, Nathaniel J. Karst†, and Fan Wu‡

Abstract. We explore the stability of equilibrium solution(s) of a simple model of microvascular blood flow in a two-node network. The model takes the form of convection equations for red blood cell concentration, and contains two important rheological effects—the Fåhræus–Lindqvist effect, which governs viscosity of blood flow in a single vessel, and the plasma skimming effect, which describes the separation of red blood cells at diverging nodes. We show that stability is governed by a linear system of integral equations, and we study the roots of the associated characteristic equation in detail. We demonstrate using a combination of analytical and numerical techniques that it is the relative strength of the Fåhræus–Lindqvist effect and the plasma skimming effect which determines the existence of a set of network parameter values which lead to a Hopf bifurcation of the equilibrium solution. We confirm these predictions with direct numerical simulation and suggest several areas for future research and application.

Key words. blood flow, microvascular network, instability, bifurcation

AMS subject classifications. 37-xx, 92-xx

DOI. 10.1137/060670699

1. Introduction.

Periodic or oscillatory dynamics in biological systems are common. Examples include the pacemaker in the heart, breathing patterns, and periodic fluctuations in leukocyte production in leukemia [15]. These types of phenomena require models consisting of nonlinear equations, most often in the form of nonlinear differential equations [14]. In fact, the modeling and analysis of complex physiological signals is a very active area for the application and development of dynamical systems theory [16].

Another example of fluctuating dynamics in biology is microvascular blood flow. Nobel prize winner August Krogh noted the heterogeneity of blood flow in the webbed feet of frogs in the early 1920's [24]. In *The Anatomy and Physiology of Capillaries* he wrote [23]

> In single capillaries the flow may become retarded or accelerated from no visible cause; in capillary anastomoses the direction of flow may change from time to time.

As more techniques were developed for measuring events in microcirculation, more fluctuations were reported. The servo-null pressure measurement system of Wiederhielm et al. [35] and Johnson and Wayland's [20] dual slit red cell velocity measurement device showed that

*Received by the editors September 25, 2006; accepted for publication (in revised form) by L. Fauci May 11, 2007; published electronically October 17, 2007. This work was supported in part by a grant from the National Institutes of Health (5 R01 HL067789-04), the undergraduate research opportunity program at the University of New Hampshire, and the Dean of Faculty's research fund at Olin College.
http://www.siam.org/journals/siads/6-4/67069.html
†Franklin W. Olin College of Engineering, Needham, MA 02492 (john.geddes@olin.edu, nathaniel.karst@students.olin.edu).
‡Department of Chemical Engineering, University of New Hampshire, Durham, NH 03824 (rtc@cisunix.unh.edu, fw@cisunix.unh.edu).

fluctuations practically always occur in microvessels. These fluctuations are usually interpreted as evidence of "biological control" of the flow (precapillary sphincters, vasomotion, etc.). In fact the sine qua non cause for observed oscillations is often vasomotion [32].

However, oscillations may not be due solely to "biological control." In 1973 Y. C. Fung suggested that the observed dynamics might be due to statistical variations in the properties of cells and vessels [13]. Fung hypothesized that small stochastic variations in cell size or membrane stiffness could lead to apparently random fluctuations in flow and pressure in simple networks. He outlined a program of study of blood cell properties, vessel mechanics, blood rheology, and network geometry to be completed before his hypothesis could be tested; in many ways this research program is still active today [28].

The 1990's saw several attempts to apply the ideas of nonlinear dynamics to the microcirculation. Time series analysis of fluctuations suggested that the oscillations may be due to deterministic chaos. Yip, Nolstein-Rathlou, and Marsh [36] proposed that pressure oscillations in the kidney were due to nonlinearities in the tubuloglomerular feedback control mechanism including multiple time delays. Cavalcanti and Ursino [6] developed computer simulations of vasomotion in microvessels. They included both passive and active response to pressure which resulted in oscillations in vessel diameter or blood flow. Griffith [17] measured fluctuations in pressure and flow in the rabbit ear, again attributing vasomotion as the cause of the fluctuations. He also estimated the correlation dimension from the time series data to be between two and three and suggested that the dynamics were due to vasomotion control. More recently, Mollica, Jain, and Netti [26] interpreted the heterogeneity in tumor blood flow in terms of collapsible conduits for blood flow and Parthimos et al. [27] analyzed fluctuations in arteriol diameter and red cell velocity in the microcirculation of rats and measured small, positive Lyapunov exponents.

While the importance of vasomotion cannot be denied, there is growing evidence that fluctuations in microvascular blood flow can be due to inherent instabilities in the capillary network. Kiani et al. [21] found spontaneous oscillations in blood velocity in hamsters, even under experimental conditions which rule out myogenic or vasomotive effects. They also formulated a mathematical model of network blood flow and found good qualitative agreement with their experimental results. Their blood flow model is based on following slugs of red blood cells through the capillary network; the red blood cells move with a velocity which is determined by the hydraulic resistance of the network. In a network consisting of almost four hundred vessels (the topology and geometry were based on direct in vivo observations), spontaneous oscillations were found in 30% of them; the other vessels had steady flow and hematocrit.

Carr and LeCoin [4] reformulated the model as a system of partial differential equations for blood hematocrit. They found that blood velocity, hematocrit, and nodal pressures can oscillate spontaneously in the absence of biological control in small networks with about fifteen vessels. They found evidence of Hopf bifurcations and limit cycles but were unable to determine which parameters controlled the dynamics. In 2005, Carr, Geddes, and Wu [5] demonstrated that oscillations were possible in a network consisting of only two nodes and four vessels—the so-called two-node network. While the dimensionless parameters governing the onset of instability were discovered, finding unstable parameter values was an exercise in educated guessing. In addition, the influence of the various rheological properties of blood

was poorly understood.

In this paper, we focus our attention on the two-node network and analyze in detail the onset of instability via Hopf bifurcation. In section 2 we review the major rheological properties of blood, including the Fåhræus–Lindqvist effect and the plasma skimming effect. We also introduce two simple, parametric models for these effects which make the analysis tractable. In section 3 we introduce the model of the two-node network and discuss the equilibrium solutions. In section 4 we derive the linearized equations close to steady state and the resulting characteristic equation that governs stability. In section 5 we show that Hopf bifurcations are possible, and we describe their dependence on three dimensionless parameters. In section 6 we develop a method for mapping the Hopf bifurcation points back to parameter space, thus allowing us to predict the network geometries that lead to instability, which we confirm via direct numerical simulation. We also investigate the relative importance of the Fåhræus–Lindqvist effect and the plasma skimming effect using two lumped parameters. Finally, in section 7 we offer some closing thoughts and remarks.

2. Blood rheology in microvessels.

Blood is a concentrated suspension containing red blood cells, white blood cells, and platelets. These components are suspended in plasma, an aqueous solution containing a variety of ions and macromolecules. We focus our attention on the red blood cells, which are biconcave discs with a typical diameter of about $8\mu m$ and thickness of $2\mu m$. Normal blood has a volume concentration of red blood cells (hematocrit) of about 45%.

The microcirculation consists of vessels with diameters ranging from $10\mu m$ to $100\mu m$, and is responsible for heat and mass exchange with the surrounding tissues; every cell in the body is within $100\mu m$ of a capillary. In this section we review two important rheological effects which we include in our model. The Fåhræus–Lindqvist effect captures the hematocrit- and diameter-dependent viscosity of blood when flowing through a single vessel. The plasma skimming effect describes the way red blood cells are distributed at a diverging node. We also propose two simple parametric versions of these which we use in this study in order to make the analysis tractable.

2.1. The Fåhræus–Lindqvist effect.

In the mid-nineteenth century, Jean Leonard Marie Poiseuille turned from his studies on the microcirculation of the frog to a series of detailed experiments on the flow of liquids in small glass capillaries [33]. Poiseuille demonstrated experimentally that the resistance to flow, defined as the ratio of the pressure drop to the volumetric flow rate, in a glass tube of circular cross-section is proportional to the length of the tube and inversely proportional to the fourth power of the diameter. In modern notation, Poiseuille's law reads

$$R = \frac{128 L \mu}{\pi D^4},$$

where R is the resistance to flow, D is the diameter of the tube, L is the length of the vessel, and the constant of proportionality μ is a measure of the viscosity of the fluid. The first theoretical derivation of Poiseuille's law seems to have been published in 1860 by Hagenbach, but numerous authors appear to have made similar derivations during the same time period [33].

In 1931 Fåhræus and Lindqvist [11] conducted a set of experiments on blood flow through narrow capillary tubes with diameters ranging from $30\mu m$ to $300\mu m$. They demonstrated

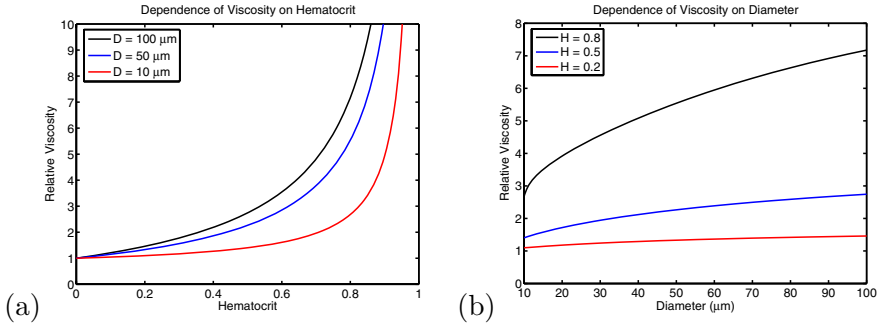


Figure 1. *Relative viscosity as a function of hematocrit and diameter from* (2.1)–(2.3). (a) *Viscosity dependence on hematocrit for diameters of* $10\mu m$, $50\mu m$, *and* $100\mu m$. (b) *Viscosity dependence on diameter for hematocrit values of* $H = 0.2$, $H = 0.5$, *and* $H = 0.8$.

that the viscosity of blood, computed according to Poiseuille's law, decreases with decreasing tube diameter, and they hypothesized that this was due to the red blood cells acting as a suspension. While not explicitly remarked upon in their paper, the viscosity of blood also depends upon hematocrit.

The dependence of blood viscosity on tube diameter (and hematocrit) has been confirmed by numerous investigators. In an empirical tour de force, Pries, Neuhaus, and Gaehtgens [30] compiled a database of viscosity measurements in tubes with diameters ranging from $3\mu m$ to $2000\mu m$ and with a range of hematocrits from 0 to 0.93. They also conducted a new set of experiments using a capillary viscometer and combined all of the data into an empirical relationship for the relative viscosity of blood $\mu(H, D)$,

$$\mu(H, D) = 1 + (\mu_{0.45} - 1)\frac{(1-H)^C - 1}{(1-0.45)^C - 1}, \tag{2.1}$$

where H is the hematocrit and $\mu_{0.45}$ is the relative viscosity of blood at $H = 0.45$ which depends on diameter,

$$\mu_{0.45} = 220e^{-1.3D} + 3.2 - 2.44e^{-0.06D^{0.645}}. \tag{2.2}$$

The parameter C also depends on diameter according to

$$C = \frac{1}{1 + 10^{11}D^{12}} + (0.8 + e^{-0.075D})\left(-1 + \frac{1}{1 + 10^{11}D^{12}}\right), \tag{2.3}$$

and in both cases D is measured in microns. An example of the relative viscosity dependence on hematocrit is shown in Figure 1a for diameter values of $10\mu m$, $50\mu m$, and $100\mu m$. Notice that in each case the relative viscosity is unity at zero hematocrit and increases monotonically with increasing hematocrit and that increasing the diameter leads to an increase in relative viscosity. An example of the relative viscosity dependence on diameter is shown in Figure 1b for diameters in the range of $10\mu m$ to $100\mu m$ and hematocrit values of $H = 0.2$, $H = 0.5$, and $H = 0.8$.

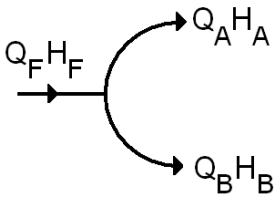


Figure 2. *Diverging node. The feed vessel has diameter D_F, flow Q_F, and hematocrit H_F. Daughter vessels A and B have diameters D_A and D_B, flows Q_A and Q_B, and hematocrits H_A and H_B, respectively.*

2.2. The plasma skimming effect. August Krogh introduced the term plasma skimming in 1921 in order to explain the disproportionate distribution of red blood cells observed at microvascular bifurcations in vivo [24]. In the absence of plasma skimming, the hematocrit entering a side branch would equal that of the feed vessel, or, alternatively, the flow ratio of red blood cells entering a side branch would equal the volumetric flow ratio. Numerous authors, however, have demonstrated both in vitro and in vivo that the red blood cell flow ratio is a nonlinear function of the volumetric flow ratio which implies that the hematocrit ratio is not unity but depends on the volumetric flow ratio. In order to be clear, consider the diverging bifurcation shown in Figure 2, which consists of a feed vessel F and two daughter vessels A and B. Denote the hematocrit and flow in the feed and daughter vessels as H_F, Q_F, H_A, Q_A, H_B, and Q_B, respectively. In general, the hematocrit ratio H_A/H_F is a function of the flow ratio $Q = Q_A/Q_F$ and is parameterized by the feed hematocrit H_F and the diameters of the vessels D_F, D_A, and D_B [12].

Many attempts to derive or measure the so-called plasma skimming function have been made. In vitro studies using plastic particles were conducted by Bugliarello and Hsiao [3] and Chien et al. [7], while Dellimore, Dunlop, and Canham [8] and Fenton, Carr, and Cokelet [12] conducted in vitro experiments using human blood. Klitzman and Johnson [22] performed in vivo experiments using hamsters, while Pries, Ley, and Gaehtgens [29] examined the distribution of red cells at sixty-five arteriol bifurcations in the rat mesentry. Some general conclusions can be drawn from these studies: red cells are not distributed in proportion to the volume flow; there is a critical fractional flow, Q_0, to a side branch below which the latter receives no blood cells; and side branch hematocrit is not always equal to feed hematocrit for $Q = 0.5$. Pries, Ley, and Gaehtgens [29] fitted experimental data to the piecewise plasma skimming function

$$\frac{H_A Q_A}{H_F Q_F} = \begin{cases} 0, & Q < Q_0, \\ \frac{e^r (Q-Q_0)^p}{e^r (Q-Q_0)^p + (1-Q-Q_0)^p}, & Q_0 \le Q \le 1 - Q_0, \\ 1, & Q > 1 - Q_0, \end{cases} \tag{2.4}$$

and determined the dependence of the dimensionless fitting parameters on the network parameters as

$$r = -\frac{6.96}{D_F} \ln\left(\frac{D_A}{D_B}\right),$$

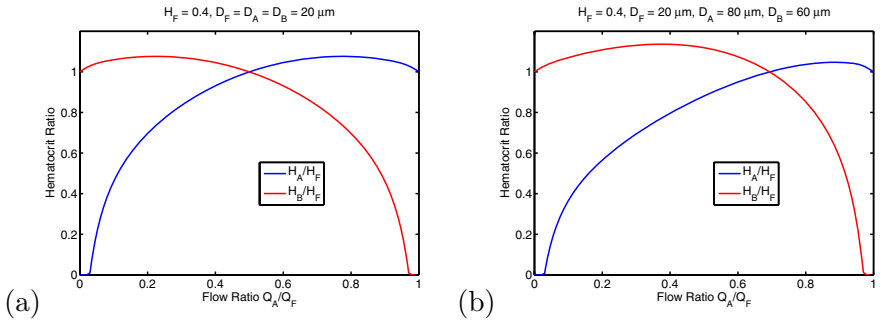


Figure 3. *Plasma skimming function from* (2.4). (a) $H_F = 0.4$, $D_F = D_A = D_B = 20\mu m$. (b) $H_F = 0.4$, $D_F = 20\mu m$, $D_A = 80\mu m$, $D_B = 60\mu m$.

$$p = 1 + 6.98\frac{(1 - H_F)}{D_F},$$
$$Q_0 = \frac{0.4}{D_F},$$

where all of the diameters are measured in microns. More recently, Enden and Popel [10] conducted three-dimensional simulations of flow in a T-type bifurcation and found good agreement with the various experimental studies mentioned above.

In Figure 3 we plot Pries, Ley, and Gaehtgens's plasma skimming function for a couple of different parameter values in order to highlight the key features. We choose to plot the hematocrit ratio as a function of flow ratio for reasons that will become clear in future sections. Figure 3a shows both the hematocrit ratio in branch A and the hematocrit ratio in branch B as functions of the flow ratio $Q = Q_A/Q_F$ for the following parameter values: $H_F = 0.4$, $D_F = 20\mu m$, and $D_A = D_B = 20\mu m$. As a result of the equal diameters in the daughter branches, the parameter $r = 0$, which implies that the hematocrit ratio in both branches is unity for $Q = 0.5$. In addition, there is a critical flow rate of $Q_0 = 0.02$ below which there is no hematocrit entering branch A and a similar value for branch B. For this symmetric set of parameters, the hematocrit ratio in branch B is simply obtained by reflection about the $Q = 0.5$ axis. Note that there is a discontinuity in the derivative of either function at Q_0 and $1 - Q_0$. In Figure 3b we show the plasma skimming function in each branch for a nonsymmetric bifurcation. The parameter values are $H_F = 0.4$, $D_F = 20\mu m$, $D_A = 80\mu m$, and $D_B = 60\mu m$. The nonequal daughter branches mean that the parameter $r \neq 0$, which breaks the symmetry of the plasma skimming functions about $Q = 0.5$. Although the plasma skimming functions are no longer mirror images, red blood cell flow is still conserved; i.e.,

$$Q_F H_F = Q_A H_A + Q_B H_B. \tag{2.5}$$

2.3. Parametric models. The empirical models of Pries et al. [29, 30] provide a detailed description of the dependence of viscosity and plasma skimming on the various vessel and flow parameters. In order to gain insight, however, we propose using simple parametric models for the Fåhræus–Lindqvist effect and plasma skimming effect as follows. We assume the viscosity has an exponential dependence on hematocrit,

$$\mu(H) = e^{\delta H}, \tag{2.6}$$

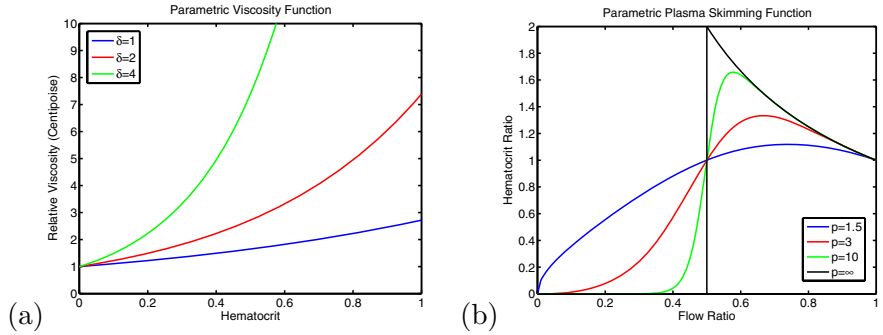


Figure 4. *Simple parametric models.* (a) *Exponential viscosity model.* (b) *Logit plasma skimming model.*

where δ is an adjustable parameter. In what follows we will assume that the same δ applies in any branch. In Figure 4a we plot the exponential function for three different values of δ. This parametric model captures the relevant feature of the Pries et al. viscosity model, namely, that viscosity is a monotonically increasing function of hematocrit with $\mu(0) = 1$. The choice of an exponential model is not unreasonable: the viscosity model of Pries, Neuhaus, and Gaehtgens [30] demonstrates a strong dependence on hematocrit; we have used other monotonically increasing models and find no substantial differences in the results that follow; and, perhaps most importantly, the exponential model makes the analysis tractable.

It is also impossible to capture all of the features of the plasma skimming function with only a single parameter. However, Carr, Geddes, and Wu [5] showed that the key ingredient is the existence of a maximum. We will use a semiempirical model proposed by Klitzman and Johnson [22],

$$\frac{H_A Q_A}{H_F Q_F} = \frac{Q^p}{Q^p + (1-Q)^p}, \tag{2.7}$$

where $p > 1$ is an adjustable parameter and $Q \in [0, 1]$. This model does not have a critical flow rate, nor does it give rise to nonsymmetric plasma skimming functions. It does, however, have a single maximum. Increasing the value of p leads to very low values of hematocrit for low values of Q while also leading to an increase in the maximum hematocrit. In Figure 4b we plot the hematocrit ratio as a function of flow ratio for several different values of p. Note that increasing p leads to a greater maximum value which slowly shifts toward $Q = 0.5$. In the limit as $p \to \infty$, the plasma skimming function is piecewise continuous and behaves as $1/Q$ for $Q \geq 0.5$.

3. The two-node network model. The model which we use was first proposed by Kiani et al. [21] and later reformulated by Carr and LeCoin [4]. It is a continuous model and assumes that the essence of microvascular blood flow can be captured with a position- and time-dependent hematocrit function. It is therefore a one-dimensional model and ignores the three-dimensional aspect of the problem and the discrete nature of red blood cells. In sections 3.1 and 3.2 we describe the partial differential equation model and relevant boundary conditions, while in section 3.3 we prove that there exists at least one equilibrium solution of the model, and we also find conditions under which there are multiple equilibria.

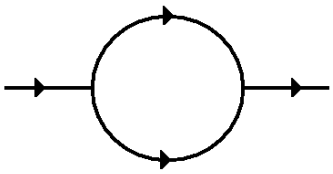


Figure 5. *The two-node network consists of a single inlet and a single outlet.*

3.1. The PDE model. Consider the simple two-node network shown in Figure 5. We assume that the hematocrit in each branch is governed by a first-order wave equation of the form

$$\frac{\partial H_A}{\partial t} + v_A \frac{\partial H_A}{\partial x_A} = 0, \qquad 0 \le x_A \le l_A, \quad t \ge 0,$$
$$\frac{\partial H_B}{\partial t} + v_B \frac{\partial H_B}{\partial x_B} = 0, \qquad 0 \le x_B \le l_B, \quad t \ge 0,$$

where $H_A(x_A, t)$ and $H_B(x_B, t)$ are the hematocrits in branches A and B, respectively. This is an appropriate description for the one-dimensional transport of red cells with small dispersion. The propagation velocity in each branch is proportional to the flow in each branch

$$v_A(t) = \frac{4Q_A(t)}{\pi d_A^2},$$
$$v_B(t) = \frac{4Q_B(t)}{\pi d_B^2},$$

where Q_A and Q_B are the flows in branches A and B, respectively, and d_A and d_B are the diameters of branches A and B, respectively. It is possible to express both velocities in terms of the fractional flow in branch A, $Q(t) = Q_A(t)/Q_F$, where Q_F is the steady volumetric flow rate in the feed branch. Conservation of volumetric flow at the branch implies that $Q_A(t) + Q_B(t) = Q_F$. It also implies that $0 \le Q(t) \le 1$.

A dimensionless form of the governing equations may be derived by scaling space and time according to

$$\hat{x}_A = \frac{x_A}{l_A},$$
$$\hat{x}_B = \frac{x_B}{l_B},$$
$$\hat{t} = t\frac{4Q_F}{\pi d_A^2 l_A},$$

which results in the following dimensionless propagation equations:

$$\frac{\partial H_A}{\partial t} + Q\frac{\partial H_A}{\partial x_A} = 0, \qquad 0 \le x_A \le 1, \quad t \ge 0, \tag{3.1}$$

$$\frac{\partial H_B}{\partial t} + \alpha(1-Q)\frac{\partial H_B}{\partial x_B} = 0, \qquad 0 \le x_B \le 1, \quad t \ge 0. \tag{3.2}$$

The hats have been dropped for convenience and the parameter α has been introduced. It is the ratio of the volume of branch A to that of branch B,

$$\alpha = \frac{l_A d_A^2}{l_B d_B^2}.$$

The propagation velocity in branch A is now Q, while in branch B the velocity is $\alpha(1-Q)$.

3.2. Constitutive relations and boundary conditions. The governing equations, (3.1)–(3.2), are accompanied by constitutive relations and boundary conditions. Conservation of flow and the two-node network topology determine the fractional flow rate in branch A in terms of the hydraulic resistance of each branch,

$$Q(t) = \frac{R_B(t)}{R_A(t) + R_B(t)}, \tag{3.3}$$

with a similar expression for branch B. In turn, the hydraulic resistance in each branch is determined by the Fåhræus–Lindqvist effect,

$$R_A(t) = \frac{128 l_A}{\pi d_A^4} \mu_A(\bar{H}_A(t), d_A), \tag{3.4}$$

$$R_B(t) = \frac{128 l_B}{\pi d_B^4} \mu_B(\bar{H}_B(t), d_B), \tag{3.5}$$

where μ_A and μ_B are the relative viscosities in branches A and B, respectively. The viscosity in each branch is a function of the axially averaged hematocrit in the branch and the diameter of the branch. Finally, the entrance hematocrit to each branch is governed by the plasma skimming effect,

$$H_A(0,t) = H_F f(Q(t)), \tag{3.6}$$

$$H_B(0,t) = H_F g(Q(t)), \tag{3.7}$$

where H_F is the hematocrit of the feed branch, and the functions f and g specify the plasma skimming effect for branches A and B, respectively. The governing equations (3.1)–(3.2), along with the constitutive relations (3.3)–(3.5) and boundary conditions (3.6)–(3.7), completely determine the problem.

3.3. Equilibrium solutions. In steady state, $\partial H_A/\partial t = 0$, the hematocrit along each branch is constant and equal to the entrance hematocrit. The entrance hematocrit depends on the steady state flow Q^*, which in turn implies that the steady state viscosity μ_A^* is a function of Q^*. The hydraulic resistance in each branch is therefore a function of Q^* in steady state, and (3.3) becomes

$$Q^* = \psi(Q^*), \tag{3.8}$$

where the nonlinear function on the right-hand side may be written as

$$\psi(Q^*) = \frac{1}{1 + \frac{d_B^4 l_A \mu_A^*}{d_A^4 l_B \mu_B^*}}. \tag{3.9}$$

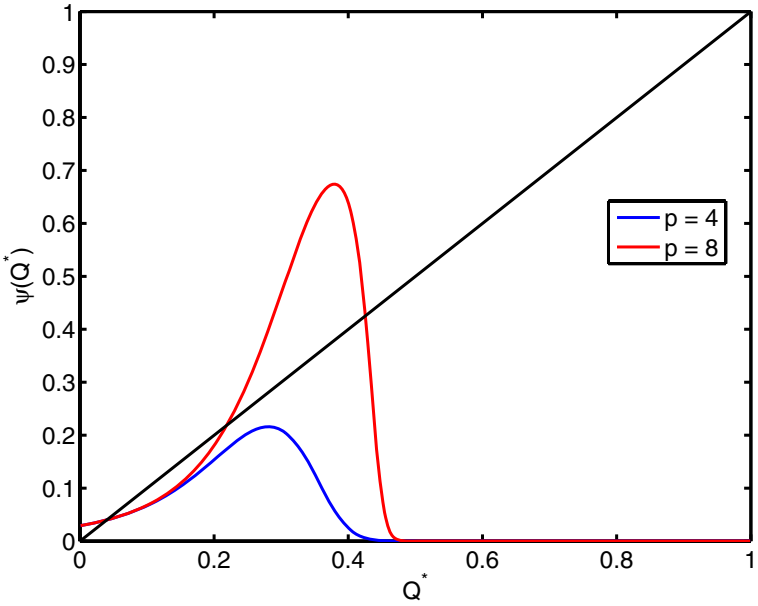


Figure 6. *Steady state solutions are given where the function ψ, defined in (3.9), crosses the line of unit slope. The parameter values are $\delta H_F = 8$, $l_A/l_B = 10$, and $d_B/d_A = 10$. For $p = 4$ there is one steady state solution, while for $p = 8$ there are three steady state solutions.*

The steady state viscosity is determined via the relations

$$\mu_A^* = \mu_A(\bar{H}_A^*, d_A), \tag{3.10}$$

$$\mu_B^* = \mu_B(\bar{H}_B^*, d_B), \tag{3.11}$$

$$\bar{H}_A^* = H_F f(Q^*), \tag{3.12}$$

$$\bar{H}_B^* = H_F g(Q^*). \tag{3.13}$$

We are now ready to state our first result concerning the existence of equilibrium solutions.

Theorem 3.1. *Assume that μ_A and μ_B are positive, continuous functions of $\bar{H}_A$ and $\bar{H}_B$. Also assume that f and g are continuous functions of Q^*. Then at least one steady state solution to (3.8)–(3.10) exists.*

Proof. We appeal to the Brouwer fixed-point theorem. If the viscosity functions are continuous functions of $\bar{H}_A$ and $\bar{H}_B$ and the plasma skimming functions are continuous functions of Q^*, then $\psi(Q^*)$ is continuous. In addition, the positivity of the viscosity functions implies that $\psi \in [0, 1]$ for all $Q^* \in [0, 1]$. By the fixed-point theorem this implies that at least one root exists with $Q^* \in [0, 1]$. ∎

A closed form solution for Q^* cannot be obtained except in the case of simple viscosity models. Solutions can be visualized, however, by plotting $\psi(Q^*)$ versus Q^* as demonstrated in Figure 6 for different parameter values. Equilibrium solutions are given where the function ψ crosses the line of unit slope. Here we use the parametric models with the following parameter values: $\delta H_F = 8$, $l_A/l_B = 10$, and $d_B/d_A = 10$. Notice that while Theorem 3.1 guarantees that a root exists, it does not guarantee that it is unique. Indeed, Figure 6 demonstrates

that, as we hold the value of δH_F fixed and change from $p = 4$ to $p = 8$, we transition from a system with a single steady state to one with multiple steady states. We state and prove below a condition which guarantees that the steady state is unique.

Theorem 3.2. *Define two dimensionless parameters, b and c, according to*

$$b = -H_F Q^*(1-Q^*) f'(Q^*) \frac{\mu_A'^*}{\mu_A^*},$$
$$c = H_F Q^*(1-Q^*) g'(Q^*) \frac{\mu_B'^*}{\mu_B^*}.$$

The steady state solution is unique if and only if

$$b + c < 1$$

for all $Q^* \in (0,1)$. The derivative of μ_A with respect to the average hematocrit is

$$\mu_A'^* = \left.\frac{\partial \mu_A}{\partial \bar{H}_A}\right|_{Q^*},$$

while the derivative of f with respect to the flow rate is

$$f'(Q^*) = \left.\frac{df}{dQ}\right|_{Q^*}.$$

Similar expressions apply to branch B.

Proof. Since Q^* is constrained to the closed interval $[0,1]$, we can rearrange (3.8) for the ratio of the lengths,

$$\frac{l_A}{l_B} = \left(\frac{d_A}{d_B}\right)^4 \left(\frac{1-Q^*}{Q^*}\right) \frac{\mu_B^*}{\mu_A^*}, \tag{3.14}$$

and use this formulation to determine whether the equilibrium solution is unique. The positive, continuous viscosity functions imply that $l_A/l_B \geq 0$ for all $Q^* \in [0,1]$ and that

$$\lim_{Q^* \to 0} \frac{l_A}{l_B} = +\infty,$$
$$\lim_{Q^* \to 1} \frac{l_A}{l_B} = 0.$$

The length ratio therefore cannot be a monotonically increasing function of Q^*. If $\frac{d(l_A/l_B)}{dQ^*} < 0$ for all $Q^* \in (0,1)$, then the length ratio is a monotonically decreasing function of Q^*, which implies that the steady state solution Q^* is unique for any set of network parameters. The derivative is

$$\frac{d}{dQ^*}\left(\frac{l_A}{l_B}\right) = \frac{1}{Q^{*2}} \frac{\mu_B^*}{\mu_A^*} \left(\frac{d_A}{d_B}\right)^4 (b + c - 1),$$

thus proving that the equilibrium solution is unique if $b + c - 1 < 0$ for all Q^*. If, on the other hand, there exists a $Q^* \in (0,1)$ for which $\frac{d(l_A/l_B)}{dQ^*} > 0$, then, by continuity of l_A/l_B, there exists multiple equilibrium solutions for some set of network parameters. ∎

4. Linearized equations. While the governing partial differential equations and accompanying constitutive relations are convenient for numerical simulation, they are not very useful from an analytical point of view. They can, however, be linearized and then transformed into a corresponding system of delay equations which can in turn be manipulated into a single linear integral equation. This proves to be most convenient in analyzing the linear stability of the steady state solution(s).

We begin by rewriting the governing partial differential equations (3.1)–(3.2) about the steady state solution. We assume that $Q(t) = Q^*$ for $t < 0$, $\bar{H}_A(t) = \bar{H}_A^*$, and $\bar{H}_B(t) = \bar{H}_B^*$ for $t \leq 0$. A perturbation will be introduced into the flow rate by specifying $Q(0) \neq Q^*$ and defining

$$\begin{aligned} Q(t) &= Q^* + \tilde{Q}(t), \\ H_A(x_A, t) &= \bar{H}_A^* + \tilde{H}_A(x_A, t), \\ H_B(x_B, t) &= \bar{H}_B^* + \tilde{H}_B(x_B, t). \end{aligned}$$

Replacing into (3.1)–(3.2) and grouping terms lead to the system of partial differential equations

$$\frac{\partial \tilde{H}_A}{\partial t} + Q^* \frac{\partial \tilde{H}_A}{\partial x_A} = -\tilde{Q} \frac{\partial \tilde{H}_A}{\partial x_A}, \tag{4.1}$$

$$\frac{\partial \tilde{H}_B}{\partial t} + \alpha(1 - Q^*) \frac{\partial \tilde{H}_B}{\partial x_B} = \alpha \tilde{Q} \frac{\partial \tilde{H}_B}{\partial x_B}. \tag{4.2}$$

If we drop the nonlinear terms on the right-hand side, then the linearized equations consist of two coupled, constant velocity propagation equations for perturbations to the equilibrium hematocrit. We can define the steady state propagation (or delay) times in branches A and B, respectively, as

$$\begin{aligned} \tau^* &= \frac{1}{Q^*}, \\ \theta^* &= \frac{1}{\alpha(1 - Q^*)}. \end{aligned}$$

A more convenient form for analysis can be found by integrating (4.1)–(4.2) over the spatial variable in each branch,

$$\begin{aligned} \tau^* \frac{d\hat{H}_A}{dt} &= \tilde{H}_A(0, t) - \tilde{H}_A(1, t), \\ \theta^* \frac{d\hat{H}_B}{dt} &= \tilde{H}_B(0, t) - \tilde{H}_B(1, t), \end{aligned}$$

where $\hat{H}_A$ is defined by

$$\hat{H}_A(t) = \int_0^1 \tilde{H}_A(x_A, t)\, dx_A,$$

and similarly for $\hat{H}_B$. This implies that

$$\bar{H}_A(t) = \bar{H}_A^* + \hat{H}_A(t),$$

with a similar expression for the other branch.

We now make use of the fact that our governing linearized equations are a pair of first-order wave equations with constant velocity. This implies that the hematocrit in each branch simply propagates along the appropriate characteristic with velocity Q^* in branch A and $\alpha(1-Q^*)$ in branch B. The hematocrit at the exit of vessel A is therefore the hematocrit at the entrance at an earlier time, namely, τ^*. With this in mind, the governing equations reduce to

$$\tau^* \frac{d\hat{H}_A}{dt} = \tilde{H}_A(0,t) - \tilde{H}_A(0,t-\tau^*),$$
$$\theta^* \frac{d\hat{H}_B}{dt} = \tilde{H}_B(0,t) - \tilde{H}_B(0,t-\theta^*).$$

Recall that the entrance hematocrit in branch A is specified by the plasma skimming rule (3.6) which linearizes to give

$$\begin{aligned} & H_A(0,t) = H_F f(Q), \\ \Rightarrow\ & H_A(0,t) = H_F(f(Q^*) + f'(Q^*)\tilde{Q}), \\ \Rightarrow\ & \tilde{H}_A(0,t) = H_F f'(Q^*)\tilde{Q}, \end{aligned}$$

with a similar expression for branch B. Our linearized governing equations become

$$\tau^* \frac{d\hat{H}_A}{dt} = H_F f'(Q^*)\left(\tilde{Q}(t) - \tilde{Q}(t-\tau^*)\right),$$
$$\theta^* \frac{d\hat{H}_B}{dt} = H_F g'(Q^*)\left(\tilde{Q}(t) - \tilde{Q}(t-\theta^*)\right),$$

which is a pair of delay differential equations with constant delay. If we integrate each of these from $t = 0$ to $t = T$ we find

$$\tau^* \hat{H}_A(T) = H_F f'(Q^*) \int_0^T \left(\tilde{Q}(s) - \tilde{Q}(s-\tau^*)\right) ds, \tag{4.3}$$

$$\theta^* \hat{H}_B(T) = H_F g'(Q^*) \int_0^T \left(\tilde{Q}(s) - \tilde{Q}(s-\theta^*)\right) ds, \tag{4.4}$$

where we have used the initial conditions $\hat{H}_A(0) = 0$ and $\hat{H}_B(0) = 0$. In addition, we can decompose the remaining integrals into two pieces,

$$\int_0^T \tilde{Q}(s-\tau^*)\, ds = \int_0^{\tau^*} \tilde{Q}(s-\tau^*)\, ds + \int_{\tau^*}^T \tilde{Q}(s-\tau^*)\, ds,$$

and use the initial condition on the flow perturbation to conclude that the first term is zero. A change of variables results in

$$\int_0^T \tilde{Q}(s-\tau^*)\, ds = \int_0^{T-\tau^*} \tilde{Q}(u)\, du.$$

Replacing and combining with the other integral give the following set of integral equations:

$$\tau^* \hat{H}_A(T) = H_F f'(Q^*) \int_{T-\tau^*}^{T} \tilde{Q}(u)\, du, \tag{4.5}$$

$$\theta^* \hat{H}_B(T) = H_F g'(Q^*) \int_{T-\theta^*}^{T} \tilde{Q}(u)\, du. \tag{4.6}$$

Equations (4.5)–(4.6) allow us to compute the axially averaged hematocrit perturbation. To do so we require the flow perturbation and its history. The problem is closed when we linearize the constitutive relation (3.3). Using the notation developed in section 3.3 we see that

$$\begin{aligned} Q(t) &= Q^* + \left.\frac{\partial \psi}{\partial \bar{H}_A}\right|_{Q^*} \left(\bar{H}_A(t) - \bar{H}_A^*\right) + \left.\frac{\partial \psi}{\partial \bar{H}_B}\right|_{Q^*} \left(\bar{H}_B(t) - \bar{H}_B^*\right), \\ \Rightarrow \tilde{Q}(t) &= \left.\frac{\partial \psi}{\partial \bar{H}_A}\right|_{Q^*} \hat{H}_A(t) + \left.\frac{\partial \psi}{\partial \bar{H}_B}\right|_{Q^*} \hat{H}_B(t). \end{aligned}$$

The partial derivatives are straightforward to compute and are given by

$$\begin{aligned} \left.\frac{\partial \psi}{\partial \bar{H}_A}\right|_{Q^*} &= -Q^*(1-Q^*)\frac{\mu_A'^*}{\mu_A^*}, \\ \left.\frac{\partial \psi}{\partial \bar{H}_B}\right|_{Q^*} &= Q^*(1-Q^*)\frac{\mu_B'^*}{\mu_B^*}. \end{aligned}$$

Combining (4.5)–(4.6) and the expression for $\tilde{Q}(t)$ results in the linear integral equation

$$\tilde{Q}(T) = \frac{b}{\tau^*} \int_{T-\tau^*}^{T} \tilde{Q}(u)\, du + \frac{c}{\theta^*} \int_{T-\theta^*}^{T} \tilde{Q}(u)\, du, \tag{4.7}$$

where the dimensionless parameters b and c are the same as those defined in Theorem 3.2,

$$b = -H_F Q^*(1-Q^*) f'(Q^*) \frac{\mu_A'^*}{\mu_A^*}, \tag{4.8}$$

$$c = +H_F Q^*(1-Q^*) g'(Q^*) \frac{\mu_B'^*}{\mu_B^*}. \tag{4.9}$$

Equations (4.7)–(4.8) completely determine the linear stability of the steady state flow rate Q^* to a perturbation $\tilde{Q}(T)$ such that $\tilde{Q}(T) = 0$ for all $T < 0$ and $\tilde{Q}(0) \neq 0$.

The characteristic equation can be obtained by seeking a solution of the linear integral equation in the form

$$\tilde{Q}(T) = \tilde{Q}(0) e^{\lambda T}, \tag{4.10}$$

where λ is complex. A nontrivial solution of (4.7) subject to the solution (4.10) exists if and only if λ satisfies the characteristic equation

$$1 = b\frac{(1-e^{-\lambda\tau})}{\lambda\tau} + c\frac{(1-e^{-\lambda\theta})}{\lambda\theta},$$

where we have dropped the $*$ for convenience. It will prove more fruitful if we make a change of variables and eliminate one of the delay times in favor of the ratio of the delay times. If we redefine $\hat{\lambda} = \lambda\tau$, then we obtain

$$1 = b\frac{(1-e^{-\lambda})}{\lambda} + c\frac{(1-e^{-\lambda\gamma})}{\lambda\gamma}, \tag{4.11}$$

where again we have dropped the hats for convenience. We have introduced a new dimensionless parameter, γ, which is the ratio of the steady state delay times, i.e.,

$$\gamma = \frac{\theta}{\tau} = \frac{Q^*}{\alpha(1-Q^*)}. \tag{4.12}$$

It is clear from l'Hôpital's rule that $\lambda = 0$ is not a root of (4.11) unless $b + c = 1$. This is a degenerate case and coincides with the creation of multiple equilibria, as demonstrated in section 3.3. We will not analyze this case here but focus instead on the modified characteristic equation,

$$\lambda = b(1-e^{-\lambda}) + \frac{c}{\gamma}(1-e^{-\lambda\gamma}), \tag{4.13}$$

where we ignore the zero root of (4.13). This characteristic equation depends on the three dimensionless parameters, b, c, and γ. The delay time ratio γ generally takes values on $(0,\infty)$ depending on the network parameters. However, we will consider only networks with values of $\gamma \in (0,1)$ since interchanging the vessel lengths and diameters results in $\alpha \to 1/\alpha$ and $Q^* \to 1-Q^*$. This implies that $\gamma \to 1/\gamma$, which recovers the original domain.

5. The characteristic equation. The characteristic equation (4.13) governs the stability of the two-node network. It is similar to a characteristic equation that has received some attention in the delay differential equation literature,

$$\lambda + a + be^{-\tau_1\lambda} + ce^{-\tau_2\lambda} = 0, \tag{5.1}$$

where τ_1 and τ_2 are the delays and a, b, and c are free parameters. Clearly our characteristic equation is related, but it is not simply a special case of (5.1)—it is important that the delay γ appear explicitly both inside and outside the exponential function in (4.13).

Several authors have considered the solutions of (5.1) using a variety of techniques [19, 9, 18, 2, 25]. The central question is whether there are any roots with $\mathrm{Re}(\lambda) > 0$. Hayes [19] studied the single delay problem ($c = 0$) and found conditions on the parameters for which all the roots lie to the left of the imaginary axis. The single delay equation is also discussed by El'sgol'ts and Norkin using the method of D-partitions [9]. In this method, the ab-parameter plane is divided into regions, each of which contains an integer number of roots with $\mathrm{Re}(\lambda) > 0$. The regions are separated by partition curves on which roots cross the imaginary axis. The general two delay problem was studied by Hale and Huang [18], who partitioned the $\tau_1\tau_2$-parameter plane into regions containing an integer number of roots with $\mathrm{Re}(\lambda) > 0$. Bélair and Campbell [2] considered the case with $a = 0$ and found the stability region in the $c\tau_2$-parameter plane after normalizing $b = 1$. Mahaffy, Zak, and Joiner [25] returned to the general

case, after normalizing $\tau_1 = 1$, and determined instability surfaces in the *abc*-parameter space as a function of delay τ_2. Our characteristic equation is easy to analyze using a similar approach; we divide the *cb*-plane into regions which contain a given number of roots, and we do this for different values of γ.

In section 5.1 we show that the roots of the characteristic equation with $\mathrm{Re}(\lambda) > 0$ (if they exist) are located within a bounded domain in the complex plane. In section 5.2 we show that if $b + c > 1$, there is exactly one positive real root, and that if $b + c < 1$, there are either two positive real roots or zero positive real roots. We also derive a parametric expression for the curve that separates the region of zero positive real roots from that of two positive real roots. In section 5.3 we derive parametric expressions for the Hopf bifurcation curves in the *cb*-plane on which the roots are purely imaginary, and we prove that $bc < 0$ is a necessary condition for the existence of purely imaginary roots. By combining the real root case and the imaginary root case we are therefore able to build a complete picture of the distribution of roots in the *cb*-plane as we vary the value of $\gamma \in (0, 1)$.

5.1. Complex roots. The most general solution of the characteristic equation (4.13) is complex and takes the form

$$\lambda = \sigma + i\omega, \tag{5.2}$$

where σ is the real part and ω is the imaginary part. Substituting into the characteristic equation and separating real and imaginary terms lead to the set of equations

$$\sigma - b\left(1 - e^{-\sigma}\cos(\omega)\right) - \frac{c}{\gamma}\left(1 - e^{-\sigma\gamma}\cos(\omega\gamma)\right) = 0, \tag{5.3}$$

$$\omega - be^{-\sigma}\sin(\omega) - \frac{c}{\gamma}e^{-\sigma\gamma}\sin(\omega\gamma) = 0, \tag{5.4}$$

which must be satisfied simultaneously. The following theorem is proved below.

Theorem 5.1.

1. *Assume that* $b < 0$ *and* $c < 0$. *Then there are no roots with* $\sigma > 0$.
2. *Assume that* $b > 0$ *and* $c < 0$. *If a root with* $\sigma > 0$ *exists, then it is located within a bounded region, namely,* $(\sigma, \omega) \in (0, 2b) \times [-(b + |c|/\gamma), b + |c|/\gamma]$.
3. *Assume that* $b < 0$ *and* $c > 0$. *If a root with* $\sigma > 0$ *exists, then it is located within a bounded region, namely,* $(\sigma, \omega) \in (0, 2c/\gamma) \times [-(|b| + c/\gamma), |b| + c/\gamma]$.
4. *Assume that* $b > 0$ *and* $c > 0$. *If a root with* $\sigma > 0$ *exists, then it is located within a bounded region, namely,* $(\sigma, \omega) \in (0, 2b + 2c/\gamma) \times [-(b + c/\gamma), b + c/\gamma]$.

Proof. Assume that there is a root with $\sigma > 0$. Then the following inequalities are true:

$$0 < 1 - e^{-\sigma}\cos(\omega) < 2, \tag{5.5}$$

$$0 < 1 - e^{-\sigma\gamma}\cos(\omega\gamma) < 2, \tag{5.6}$$

$$-1 < e^{-\sigma}\sin(\omega) < 1, \tag{5.7}$$

$$-1 < e^{-\sigma\gamma}\sin(\omega\gamma) < 1. \tag{5.8}$$

First consider the case of $b < 0$ and $c < 0$. The inequalities and equation (5.3) imply that $\sigma < 0$, which is a contradiction. Thus there are no roots with $\sigma > 0$ when $b < 0$ and $c < 0$.

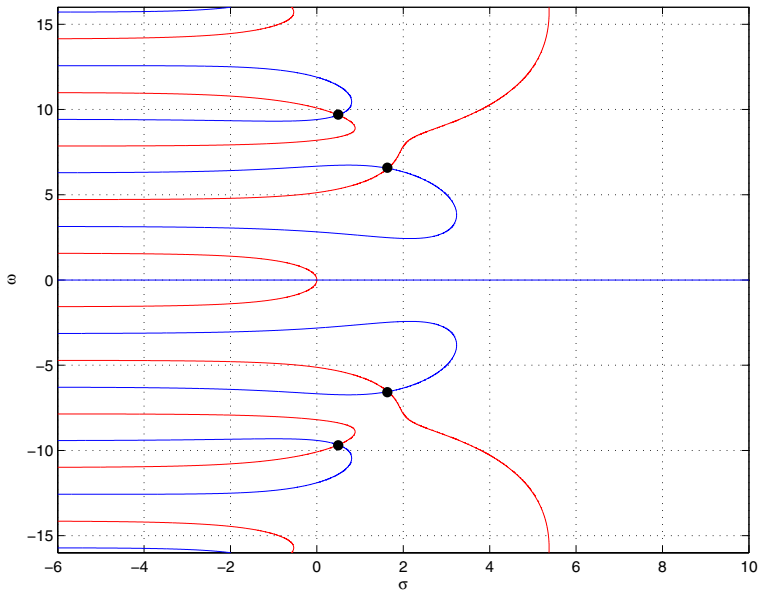


Figure 7. *The zero contours of $R(\sigma,\omega)$ and $I(\sigma,\omega)$ are shown in red and blue, respectively. The intersection points are solutions of (5.3)–(5.4). For $b=-8$, $c=2$, and $\gamma=0.2$ there are two complex conjugate roots with $\sigma>0$.*

Consider now the case of $b>0$ and $c<0$. The inequalities and equations (5.3) and (5.4) imply that

$$\sigma < 2b, \tag{5.9}$$

$$|\omega| \le b + \frac{|c|}{\gamma}, \tag{5.10}$$

which shows that if a root with $\sigma>0$ exists, then it is located within a bounded region in the complex plane. A similar argument can be made for the case of $b<0$ and $c>0$ and for the case of $b>0$ and $c>0$. ■

The solutions to the characteristic equation (4.13) may be visualized as follows. For a given set of parameters b, c, and γ, the left-hand sides of (5.3) and (5.4) are functions of σ and ω, and we will denote these functions as $R(\sigma,\omega)$ and $I(\sigma,\omega)$. The intersection of the zero contours of both functions will be the roots of the characteristic equation. An example is shown in Figure 7 for the values of $b=-8$, $c=2$, and $\gamma=0.2$. The roots of the characteristic equation are clearly marked, and in this case there are two pairs of complex conjugate roots with $\sigma>0$. The root at $(0,0)$ is not a solution of (4.11) since $b+c\neq 1$.

5.2. Real roots. We will search for real roots by looking for a solution of the form $\lambda=\sigma+i0$, which leads to the simplified characteristic equation

$$\sigma - b(1-e^{-\sigma}) - \frac{c}{\gamma}\left(1-e^{-\sigma\gamma}\right) = 0. \tag{5.11}$$

The following theorem is proved below.

Theorem 5.2.

1. *If $b + c > 1$ there is exactly one positive real root.*
2. *If $b + c < 1$ there are either two positive real roots or none at all.*
   (a) *The boundary that separates the regions of zero positive real roots from two positive real roots is defined parametrically by*

$$\det = -\frac{1}{\gamma}e^{-\sigma} + e^{-\sigma\gamma} + \frac{(1-\gamma)}{\gamma}e^{-\sigma(1+\gamma)}, \tag{5.12}$$

$$b(\sigma) = \frac{1}{\det}\left(\sigma e^{-\sigma\gamma} + \frac{e^{-\sigma\gamma}-1}{\gamma}\right), \tag{5.13}$$

$$c(\sigma) = \frac{1}{\det}\left(-\sigma e^{-\sigma} + (1-e^{-\sigma})\right), \tag{5.14}$$

   where $\sigma \in [0, \infty)$. This parametric curve requires $bc < 0$ and lies in the fourth quadrant (for $0 < \gamma < 1$).
   (b) *The boundary that separates the regions of two positive real roots from one positive real root is defined by the line segment $b + c = 1$, with $b < -\gamma/(1-\gamma)$ and $c > 1/(1-\gamma)$.*

Proof. The proof of these results is straightforward and is accomplished mostly by curve sketching. Roots of the characteristic equation (5.11) are given by roots of the function $F(\sigma)$,

$$F(\sigma) = \sigma - b(1-e^{-\sigma}) - \frac{c}{\gamma}\left(1-e^{-\sigma\gamma}\right). \tag{5.15}$$

The function $F(\sigma)$ is continuous and has the following properties:

$$\begin{aligned} F(0) &= 0, \\ \lim_{\sigma\to+\infty} F(\sigma) &= +\infty, \\ \lim_{\sigma\to-\infty} F(\sigma) &= \operatorname{sign}(b)\,\infty, \\ F'(0) &= 1-(b+c). \end{aligned}$$

In each region of the cb-plane we can sketch the function $F(\sigma)$ and determine the number of roots. This is simple in the first, second, and third quadrants and confirms the theorem in these regions. It is also unambiguous in the fourth quadrant with $b + c > 1$. However, in the fourth quadrant with $b + c < 1$ the curves are ambiguous and it is helpful to reconsider the original characteristic equation and write it in the form

$$-\sigma + b + \frac{c}{\gamma} = be^{-\sigma} + \frac{c}{\gamma}e^{-\sigma\gamma}. \tag{5.16}$$

The function on the left, which we denote as $L(\sigma)$, represents a straight line of slope -1 and intersects $b + c/\gamma$. The function on the right, which we denote as $R(\sigma)$, represents a linear combination of two different exponentials. We see that it has the same intersection at $\sigma = 0$, and that $R'(0) = -(b+c)$. As $\sigma \to \infty$ we see that $R \to 0$ from above if $c > 0$ and from below if $c < 0$ (remember that $\gamma < 1$). In addition, there exists one value of σ where $R = 0$ if and only if $-\frac{\gamma b}{c} > 1$. Finally, there exists one value of σ where $R' = 0$ if and only if $-\frac{b}{c} > 1$.

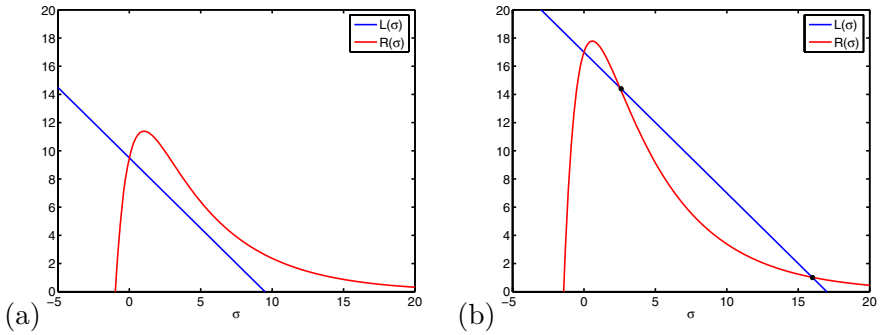


Figure 8. *The functions L and R are shown in blue and red, respectively, at two points in the fourth quadrant with $b + c \leq 1$ and $\gamma = 0.2$. (a) There are no nonzero intersections of L and R for $c = 3.5$ and $b = -8$. (b) There are two nonzero intersections of L and R for $c = 5$ and $b = -8$.*

In the fourth quadrant with $b+c < 1$ the function $R(\sigma)$ may never intersect the line $L(\sigma)$ or it may intersect it twice. Typical curves are shown in Figure 8. In Figure 8a we plot L and R as functions of σ in blue and red, respectively, for $b = -8$, $c = 3.5$, and $\gamma = 0.2$. Notice that there are no nonzero intersections in this case (the intersection at $\sigma = 0$ is not a root of (4.11) since $b + c \neq 1$). Increasing the value of c to $c = 5$ leads to two nonzero intersections, as shown in Figure 8b. The transition from zero to two roots takes place when the two curves are tangent; this occurs when both the functions and their derivatives are equal:

$$b(1 - e^{-\sigma}) + c\frac{(1 - e^{-\sigma\gamma})}{\gamma} = \sigma, \tag{5.17}$$

$$be^{-\sigma} + ce^{-\sigma\gamma} = 1. \tag{5.18}$$

This represents two linear equations in b and c which can be solved to give

$$\det = -\frac{1}{\gamma}e^{-\sigma} + e^{-\sigma\gamma} + \frac{(1-\gamma)}{\gamma}e^{-\sigma(1+\gamma)}, \tag{5.19}$$

$$b(\sigma) = \frac{1}{\det}\left(\sigma e^{-\sigma\gamma} + \frac{e^{-\sigma\gamma} - 1}{\gamma}\right), \tag{5.20}$$

$$c(\sigma) = \frac{1}{\det}\left(-\sigma e^{-\sigma} + (1 - e^{-\sigma})\right). \tag{5.21}$$

Since we are searching for nonnegative roots we can assume that $\sigma \in [0, \infty)$ and these equations therefore define a parametric curve in the cb-plane. In addition, applying l'Hôpital's rule demonstrates that

$$\lim_{\sigma \to 0^+} b(\sigma) = \frac{-\gamma}{1-\gamma}, \tag{5.22}$$

$$\lim_{\sigma \to 0^+} c(\sigma) = \frac{1}{1-\gamma}, \tag{5.23}$$

which is the point of intersection of the lines $b + c = 1$ and $b = -\gamma c$ and is in the fourth quadrant for $0 < \gamma < 1$. Moreover, it is true that $\lim_{\sigma\to\infty} b(\sigma) = -\infty$ and $\lim_{\sigma\to\infty} c(\sigma) = \infty$

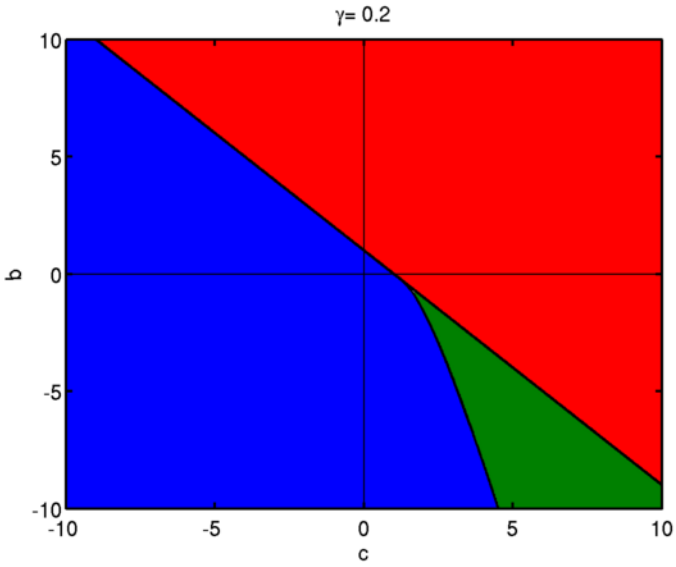


Figure 9. *The cb-plane is partitioned into three regions. In the blue region there are no real positive roots. In the red region there is one real positive root. In the green region there are two real positive roots.*

and that neither b nor c can ever become zero. The parametric curve is therefore located in the fourth quadrant. ■

In Figure 9 we plot the cb-plane for $\gamma = 0.2$ as an example. The regions containing zero, one, or two positive real roots are clearly marked. Let us define the $b+c=1$ line to be γ_1 and the parametric curve that separates the region of zero positive real roots from two positive real roots to be γ_2. For this value of γ, the point of intersection of γ_1 and γ_2 is $(5/4, -1/4)$.

5.3. Imaginary roots. We search for imaginary roots by looking for solutions of the characteristic equation of the form $\lambda = 2i\omega$. This particular choice leads to convenient algebra, and the characteristic equation decomposes into

$$\omega = b\cos(\omega)\sin(\omega) + \frac{c}{\gamma}\cos(\omega\gamma)\sin(\omega\gamma), \tag{5.24}$$

$$0 = -b\sin^2(\omega) - \frac{c}{\gamma}\sin^2(\omega\gamma). \tag{5.25}$$

We prove the following theorem below.

Theorem 5.3. *For any given value of $\gamma \in (0,1)$ there are an infinite number of Hopf bifurcation curves in the cb-plane on which there exists a purely imaginary root. These bifurcation curves are denumerable and we will label them as Γ_j for $j = 0, 1, 2, \ldots$. Each bifurcation curve is defined on a finite, closed interval of the form $[\omega_j, \omega_{j+1}]$. The curve Γ_j is defined parametrically by*

$$b(\omega) = \frac{-\omega\sin(\omega\gamma)}{\sin(\omega)\sin(\omega(1-\gamma))}, \tag{5.26}$$

$$c(\omega) = \frac{\omega\gamma\sin(\omega)}{\sin(\omega\gamma)\sin(\omega(1-\gamma))} \tag{5.27}$$

for $\omega \in [\omega_j, \omega_{j+1}]$ and is located in either the second or fourth quadrant. The endpoints ω_j and ω_{j+1} are found by sorting the roots of $\sin(\omega)$, $\sin(\omega\gamma)$, *and* $\sin(\omega(1-\gamma))$ *into ascending order and choosing the jth and $(j+1)$th roots, respectively. The Hopf bifurcation curve Γ_0 is defined on $[0, \pi]$, is semi-infinite, and originates from $(-\gamma/(1-\gamma), 1/(1-\gamma))$. All other bifurcation curves are infinite.*

Proof. Equations (5.24)–(5.25) define a linear system of equations in b and c which can be solved to give (5.26)–(5.27). These are parametric equations in ω with $bc < 0$. There is a new branch whenever one or more of the terms $\sin(\omega)$, $\sin(\omega\gamma)$, or $\sin(\omega(1-\gamma))$ is zero. The roots of these functions are $k\pi$, $l\pi/\gamma$, and $m\pi/(1-\gamma)$, respectively, for integers k, l, and m. Since we are considering $\gamma < 1$ the first two roots are $\omega_0 = 0$ and $\omega_1 = \pi$. The bifurcation curve Γ_0 is therefore defined for $\omega \in [0, \pi]$. It is semi-infinite since

$$\lim_{\omega \to 0^+} b(\omega) = \frac{-\gamma}{1-\gamma}, \tag{5.28}$$

$$\lim_{\omega \to 0^+} c(\omega) = \frac{1}{1-\gamma}, \tag{5.29}$$

$$\lim_{\omega \to \pi^-} b(\omega) = -\infty, \tag{5.30}$$

$$\lim_{\omega \to \pi^-} c(\omega) = 0^+ \tag{5.31}$$

and can be viewed as a continuation of the real root curve γ_2. All other bifurcation curves are infinite because of the linear dependence on ω in (5.26)–(5.27), even if there exist ω values where all three functions are simultaneously zero. Since $bc < 0$ every bifurcation curve is located in either the second or fourth quadrant. ∎

In Figure 10 we plot the first few Hopf bifurcation curves in the cb-plane for $\gamma = 0.2$ as an example. As expected from section 5.1, all of the Hopf bifurcation curves Γ_j lie in the second and fourth quadrants where $bc < 0$. The first Hopf bifurcation curve, Γ_0, always lies in the fourth quadrant since $\gamma < 1$. The Hopf bifurcation curves Γ_j and the real root curves γ_1 and γ_2 separate the cb-plane into distinct regions with different numbers of roots with positive real part. Changing the value of γ changes the location of the curves Γ_j and γ_2. For small γ the curves start relatively close to the origin, while as $\gamma \to 1$ the curves move off to infinity. It is clear graphically that on every curve there exists a unique point which is closest to the origin. In addition, it is also clear that the point closest to the origin lies on the first bifurcation curve Γ_0, but we have been unable to prove either of these results analytically. The next section will demonstrate that this conjecture has important implications for the onset of instability.

6. The onset of instability. The linear stability of the equilibrium solution to microvascular blood flow in a two-node network is governed by the characteristic equation (4.13). In the previous section we built a complete picture of the distribution of roots in the cb-plane for different values of γ. In particular, we found parametric equations for curves in the cb-plane that separated regions with different numbers of roots with positive real part—equations (5.26)–(5.27). The question of stability of a given equilibrium state of the microvascular blood flow model can now be answered as follows. Choose a set of network parameters. We can use (3.8) to compute the equilibrium flow ratio Q^*. If we recall from section 4 the definitions of

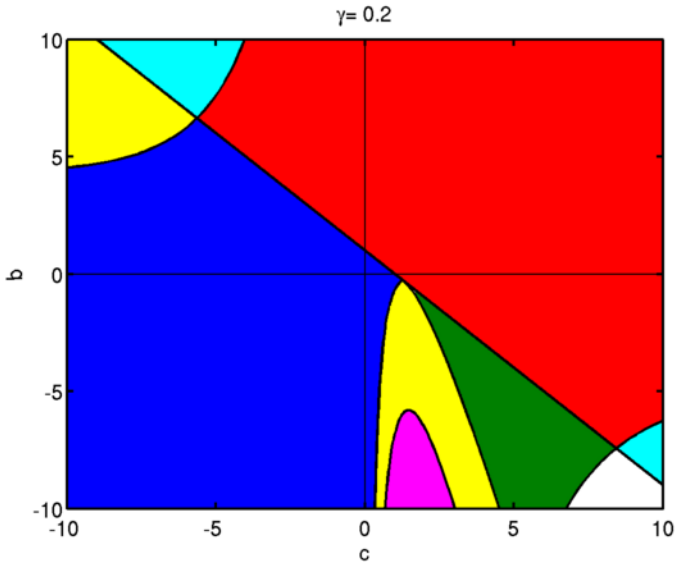


Figure 10. *The cb-plane is partitioned into distinct regions with different numbers of real and complex roots. In the blue region there are no real positive or complex roots with positive real part. In the red region there is one real positive root. In the green region there are two real positive roots. In the yellow region there is one pair of complex conjugate roots. In the magenta region there are two pairs of complex conjugate roots. In the cyan region there is one real positive root and a pair of complex conjugate roots. In the white region there are two real positive roots and a pair of complex conjugate roots.*

b, c, and γ,

$$b = -H_F Q^*(1-Q^*)f'(Q^*)\frac{\mu_A'^*}{\mu_A^*},$$
$$c = +H_F Q^*(1-Q^*)g'(Q^*)\frac{\mu_B'^*}{\mu_B^*},$$
$$\gamma = \frac{Q^*}{\alpha(1-Q^*)},$$

then each set of network parameters will map to a point in the cb-plane defined by γ. The system is stable or unstable depending on the location of the point.

The shortcomings of this approach are obvious; every point in network parameter space potentially maps to a different point in a different cb-plane defined by a different value of γ. While it is possible to use this approach to test the stability of a given network, this is all but useless as a tool for finding unstable parameter values. If, however, we focus our attention on the parametric models of (2.6) and (2.7), then b and c simplify to

$$b = -\delta H_F Q^*(1-Q^*)f'(Q^*), \tag{6.1}$$
$$c = +\delta H_F Q^*(1-Q^*)g'(Q^*). \tag{6.2}$$

As in the section on equilibrium solutions, only four network parameters are required in order to determine Q^*. These parameters are the length ratio l_A/l_B, the diameter ratio d_B/d_A, the combined Fåhræus–Lindqvist parameter δH_F, and the plasma skimming parameter p.

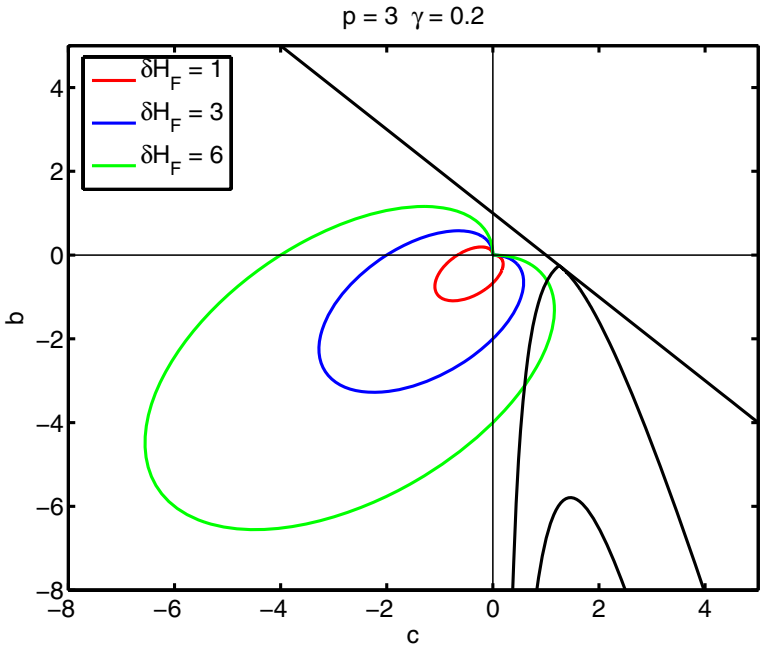


Figure 11. *The partitioned cb-plane in Figure 10 is shown with the colors removed and the scale changed for clarity. In addition, the parametric curve defined by (6.1)–(6.2) is shown for three different values of δH_F. In each case the curve is parameterized by $Q^* \in [0,1]$, launches from the origin, moves in a counter-clockwise direction, and ends at the origin. For $\delta H_F = 1$ (red curve) and $\delta H_F = 3$ (blue curve) the curve always remains in the stable region, while for $\delta H_F = 6$ (green curve) the curve crosses the first bifurcation curve Γ_1.*

In section 6.1 we show that it is possible to map the bifurcation curves to the $(Q^*, \delta H_F)$-plane for a fixed value of γ and p. In section 6.2 we consider the impact of changing γ, while in section 6.3 we show that the bifurcation curves can be mapped to a variety of network parameter planes. Finally, in section 6.4 we investigate the dependence on p, and we confirm our predictions by direct numerical simulation in section 6.5.

6.1. The $(Q^*, \delta H_F)$-plane. As a good example, let us choose $p = 3$ and $\gamma = 0.2$ and hold these values fixed. Then (6.1)–(6.2) define a parametric curve in the cb-plane parameterized by Q^*. The shape and size of this curve is influenced by the plasma skimming parameter p (via f' and g') and the combined parameter δH_F. For given values of δH_F and p, this curve is continuous and closed and has clockwise orientation for $Q^* \in [0,1]$.

In Figure 11 we show parametric curves in the cb-plane with $\gamma = 0.2$ and $p = 3$ for $\delta H_F = 1, 3, 6$. In each case, the curve originates from the origin at $Q^* = 0$, intersects the $c = 0$ axis where $g' = 0$, intersects the $b = 0$ axis where $f' = 0$, and returns to the origin as $Q^* \to 1$. Increasing δH_F results in larger parametric curves with the same basic shape since δH_F is simply a linear scale factor in (6.1)–(6.2). Notice, however, that for $\delta H_F = 1$ and $\delta H_F = 3$, the curve remains in the stable region while for $\delta H_F = 6$ the curve passes into the unstable region. There must therefore be a minimum value of δH_F below which the system is stable. For δH_F slightly larger than this minimum value, there are Hopf bifurcations at the two locations where the parametric curve crosses the first Hopf bifurcation curve. Increasing

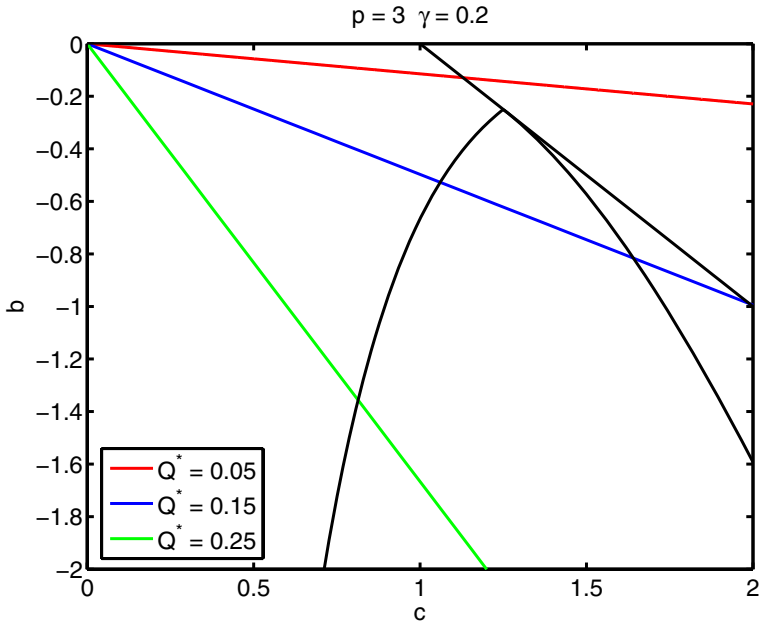


Figure 12. *The partitioned cb-plane in Figure 10 is shown with the colors removed and the scale changed for clarity. In addition, the parametric ray defined by (6.1)–(6.2) is shown for three different values of Q^*. In each case the ray is parameterized by $\delta H_F \in [0, \infty)$ and launches from the origin. For $Q^* = 0.05$ (red ray) the ray crosses into the single real root region for large enough δH_F but never enters a region with a complex root. For $Q^* = 0.15$ (blue ray) and $Q^* = 0.25$ (green ray) the ray crosses the first bifurcation curve Γ_1 for large enough δH_F.*

δH_F would result in the parametric curve crossing into the one real root and two real roots regions, respectively. For very large δH_F it is possible for the parametric curve to cross other Hopf bifurcation curves.

It is possible to map the real root curves γ_1 and γ_2 and the Hopf bifurcation curves into the $(Q^*, \delta H_F)$-plane using an alternative interpretation of (6.1)–(6.2). If we choose a value of Q^* and let δH_F vary from 0 to ∞, then the parametric curve is simply a parametric ray with slope

$$\frac{b}{c} = -\frac{f'(Q^*)}{g'(Q^*)}.$$

In Figure 12 we plot three different rays for $p = 3$ and $\gamma = 0.2$. For small Q^* there is an intersection with γ_1, but there may be no intersection with either γ_2 or Γ_0. For Q^* larger than some critical value, however, the parametric ray first intersects Γ_0, then intersects γ_2, and then intersects γ_1. This pattern continues as we increase Q^* until the parametric ray no longer intersects γ_1. For even larger Q^* the ray no longer intersects γ_2, but it continues to intersect Γ_0 as $Q^* \to Q_g$. Any ray may also intersect a higher order Hopf bifurcation curve, but this occurs for much larger values of δH_F since these curves are located farther away from the origin.

The value of δH_F at which any parametric ray crosses γ_1 may be computed as follows. Recall from section 5.2 that γ_1 is defined by $b + c = 1$. Using the definition of b and c in

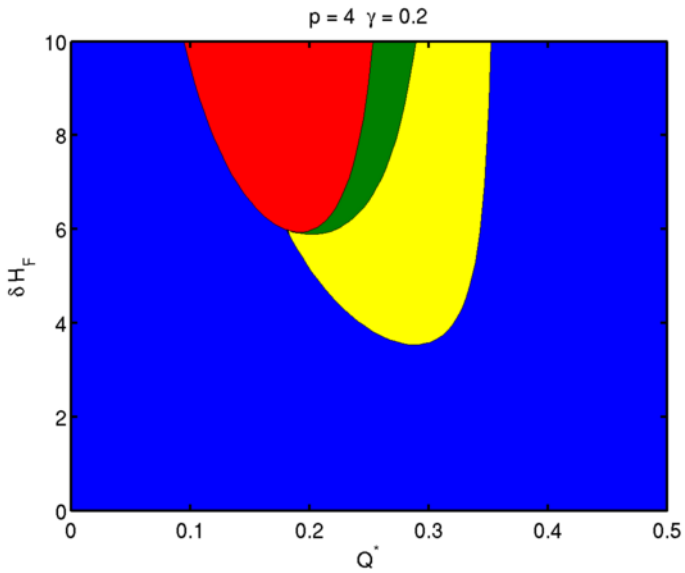


Figure 13. *The bifurcation curves and associated regions in the cb-plane are mapped to the $(Q^*, \delta H_F)$-plane. The boundaries between the regions are computed by determining, for a given value of Q^*, the value of δH_F for which the ray crosses the boundary curves of Figure 10. In the blue region there are no complex roots with positive real part. In the yellow region there is one pair of complex conjugate roots with positive real part. In the red region there is one positive real root, and in the green region there are two positive real roots.*

(6.1)–(6.2) we can solve for δH_F:

$$\delta H_F = \frac{1}{Q^*(1-Q^*)(g'(Q^*) - f'(Q^*))}.$$

In Figure 13 we plot γ_1 in the $(Q^*, \delta H_F)$-plane for $p = 4$ and $\gamma = 0.2$. For small Q^* the value of δH_F is very large. As Q^* increases, the value of δH_F decreases and then increases, reaching an asymptote at the value of Q^* where $g'(Q^*) - f'(Q^*) = 0$.

The value of δH_F at which any parametric ray crosses Γ_0 may be computed by using the definition of the Hopf bifurcation curves, equations (5.26)–(5.27). A ray from the origin to a point on any Hopf bifurcation curve has slope

$$\frac{b}{c} = -\frac{\sin^2(\omega^*\gamma)}{\gamma \sin^2(\omega^*)}.$$

Recall from section 5.3 that the first Hopf bifurcation curve, Γ_0, is defined for $\omega \in [0, \pi]$. For a given value of Q^* we can compute the value of $\omega^* \in [0, \pi]$ which results in identical slopes. The coordinates of a Hopf bifurcation point on Γ_0 can then be computed by solving

$$\frac{-\omega^* \sin(\omega^*\gamma)}{\sin(\omega^*)\sin(\omega^*(1-\gamma))} = -\delta H_F f'(Q^*)Q^*(1-Q^*)$$

for δH_F. A similar calculation can be made in order to map γ_2 onto the $(Q^*, \delta H_F)$-plane. In Figure 13 we also plot γ_2 and Γ_0 in the $(Q^*, \delta H_F)$-plane for $p = 4$ and $\gamma = 0.2$. Notice

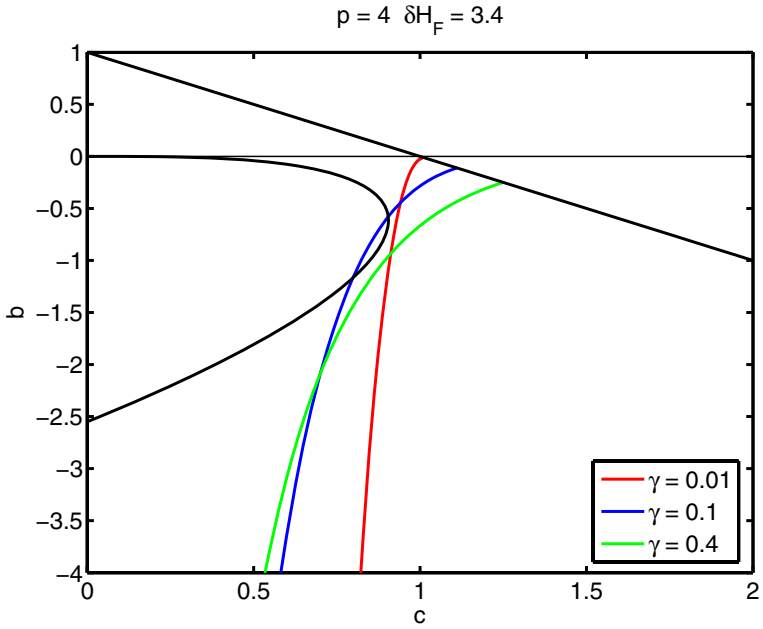


Figure 14. *The first Hopf bifurcation curve Γ_0 is shown for different values of γ. The line $b + c = 1$ is shown (in black) as well as the parametric curve for $p = 4$ and $\delta H_F = 3.4$ (also in black). For very small γ (red), the bifurcation curve Γ_0 starts close to $(1, 0)$ but never crosses the parametric curve. For an intermediate value of γ (blue), Γ_0 crosses the parametric curve twice. For a larger value of γ (green), Γ_0 starts too far from the origin to be able to cross the parametric curve.*

that there is a critical value of Q^* which results in a minimum value of δH_F below which the system is stable. As we increase δH_F above this minimum value, there exists an interval of Q^* for which the system is unstable.

6.2. Changing γ. We have so far focused our attention on an arbitrary value of γ for the sake of clarity. As indicated in section 5.3, changing γ results in the Hopf bifurcation curves moving in the cb-plane. In Figure 14 we show a fixed parametric curve with $p = 4$ and $\delta H_F = 3.4$ in the presence of Hopf bifurcation curves for several different values of γ. We plot only the first Hopf bifurcation curve Γ_0 for clarity. For $\gamma = 0.01$ the Hopf bifurcation curve begins close to $(1, 0)$, and b decreases so rapidly that it does not intersect the parametric curve. As γ increases, the Hopf bifurcation curve begins further from the origin, but b decreases less rapidly. Intersections are therefore possible, which is precisely the case when $\gamma = 0.1$. For larger γ, b decreases less rapidly still, but the Hopf bifurcation curve begins so far from the origin that intersections are impossible as shown for $\gamma = 0.4$. There are no further intersections for increasing γ.

As indicated earlier, there is a minimum value of δH_F at which we cross the first Hopf bifurcation curve Γ_0. This minimum value δH_F^{min} depends on the value of γ. By repeating our previous calculation, we can extract δH_F^{min} for each value of γ. In Figure 15 we show the results of this computation for $p = 4$. For very small γ the minimum value of δH_F is very large. As γ increases, the minimum value decreases to a global minimum and then increases

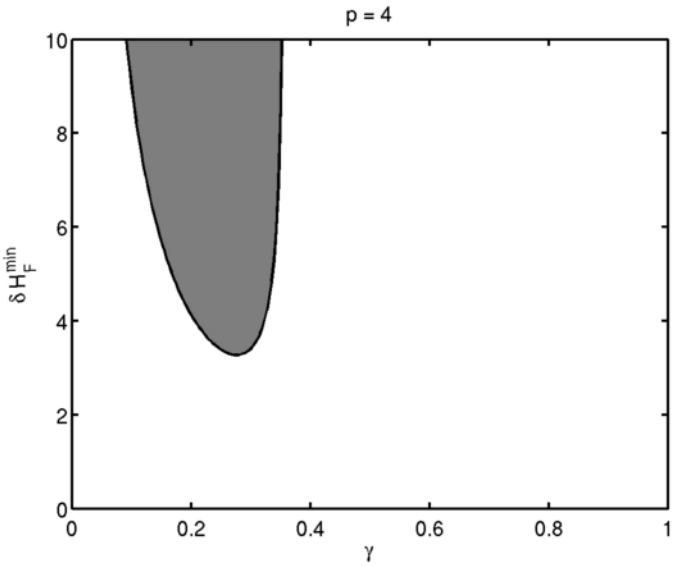


Figure 15. *The minimum value of δH_F is shown versus γ for $p = 4$. In the gray region the parametric curve defined by (6.1)–(6.2) is guaranteed to cross the first Hopf bifurcation curve Γ_0. In the white region there are no such crossings and the system is stable.*

again. We will denote this minimum value as δH_F^c, the critical value of δH_F below which the system is stable for any value of γ. For $p = 4$ the value of $\delta H_F^c = 3.2709$. The system is guaranteed to be stable if δH_F is less than this value. On the contrary, there exists an unstable set of parameters if δH_F is larger than this critical value.

6.3. Network parameter planes. Let us now choose a value of δH_F greater than the critical value, say, $\delta H_F = 3.4$. In Figure 16 we plot the region of instability in the (γ, Q^*)-plane by drawing a Hopf bifurcation contour at $\delta H_F = 3.4$ for $p = 4$. This contour is a simple closed curve for the following reason. Figure 15 shows that there is an interval of γ for which the system is unstable. In addition, Figure 13 shows that for a given value of γ there is an interval of Q^* for which the system is unstable. Combining both of these findings implies that the Hopf bifurcation curve Γ_0 is a simple closed curve in the (γ, Q^*)-plane. As we increase δH_F, the simple closed curve grows and evolves.

Every point on the Hopf bifurcation curve in Figure 16 has coordinates (γ, Q^*). Since $p = 4$ and $\delta H_F = 3.4$, there are only two network parameters undetermined—the length ratio l_A/l_B and the diameter ratio d_B/d_A. However, we have two equations—the definition of the steady state Q^*, (3.8), and the definition of γ, (4.12). Using these equations we can solve for unique values of the length and diameter ratios for every point on the Hopf bifurcation curve Γ_0. The result is a simple closed curve in the $(l_A/l_B, d_B/d_A)$-plane as shown in Figure 17. The Hopf bifurcation curve emerges at approximately $(16.5, 1.8)$ as δH_F exceeds the critical value of 3.2709. Notice that the instability region is quite narrow for δH_F close to the critical value; as δH_F increases, the size of the unstable region also increases.

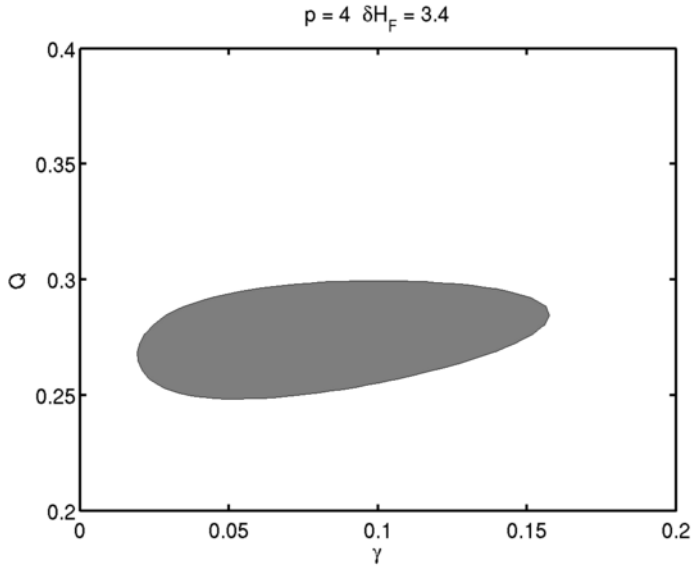


Figure 16. *The first Hopf bifurcation curve maps to a closed curve in the* (γ, Q^*)*-plane for* $p = 4$*,* $\delta H_F = 3.4$*. In the gray region the system is unstable to a Hopf bifurcation. In the white region it is stable. For larger values of* δH_F *the gray region moves and expands.*

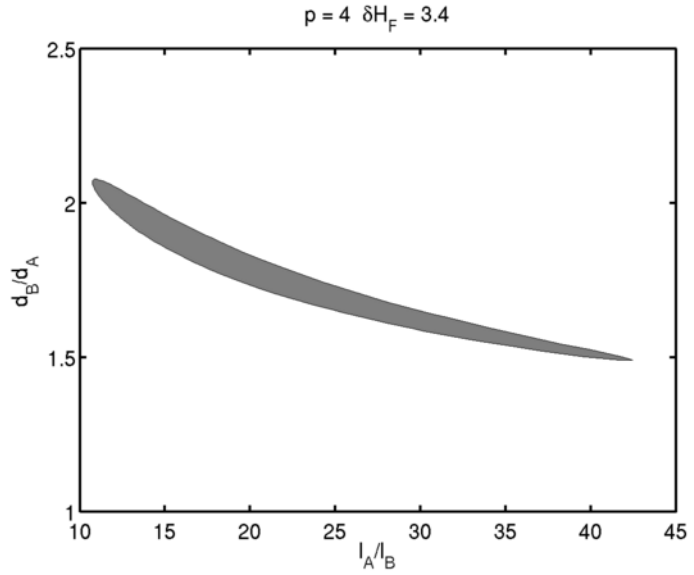


Figure 17. *The first Hopf bifurcation curve also maps to a closed curve in the* $(l_A/l_B, d_b/d_A)$*-plane for* $p = 4$*,* $\delta H_F = 3.4$*. The system is unstable to a Hopf bifurcation in the gray region and is stable otherwise. As* δH_F *increases the unstable region moves and expands.*

6.4. Changing p. The discussion so far has focused on determining unstable network parameters for a single value of p. If we change the value of p we change the shape of the parametric curves. In Figure 18 we show the parametric curves in the cb-plane ($\gamma = 0.2$) with

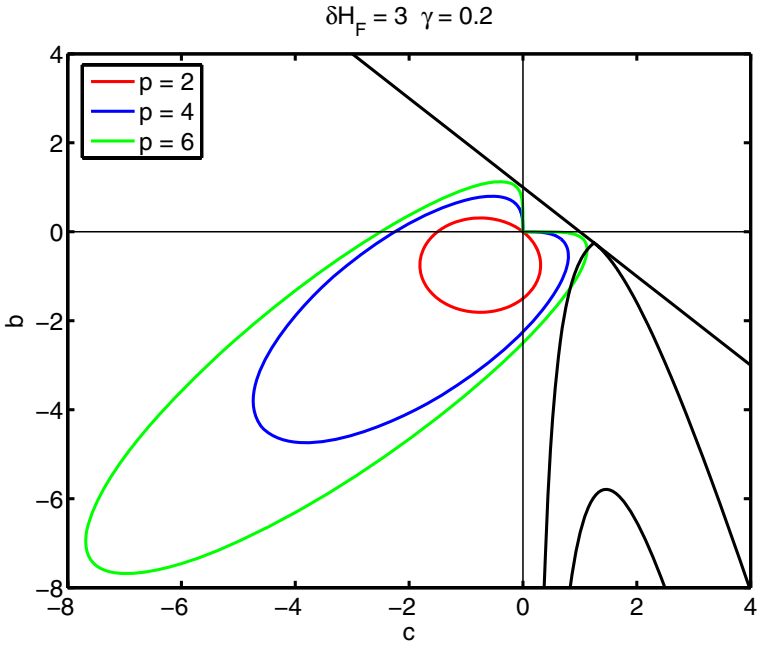


Figure 18. *The cb-plane with the bifurcation curves and three different parametric curves is shown for $\gamma = 0.2$, $\delta H_F = 3$. The parametric curve grows as p is increased until for a sufficiently large value it crosses into the unstable region (green).*

$\delta H_F = 3$ for $p = 2, 4, 6$. Notice again that as p increases the curve grows in size. For $p = 2$ and $p = 4$ it remains in the stable region, while for $p = 6$ the curve crosses the Hopf bifurcation curve Γ_0.

All of the previous computations can be repeated for any value of p. In particular, for every value of p there is a critical value of δH_F below which the system is stable. In Figure 19 we plot the critical value of δH_F as a function of p. This curve separates a region of stability from a region of instability. It is important to note that a point in the unstable region will have associated with it a region of instability in the (γ, Q^*)-plane and a corresponding region of instability in the $(l_A/l_B, d_B/d_A)$-plane. Close to the instability boundary the regions of instability are quite small, as shown previously in Figures 16 and 17. As we move further into the unstable domain, the instability regions grow and the other Hopf bifurcation curves and the real root curves γ_1 and γ_2 can come into play. We have not explored the regions far from the instability boundary for reasons that will become clear in section 7.

6.5. Numerical confirmation. We have confirmed these predictions by direct numerical simulation of the governing equations (3.1)–(3.2), along with the constitutive relations (3.3)–(3.5) and boundary conditions (3.6)–(3.7). We have used a variety of algorithms, including the upwinding convection scheme and the second-order Lax–Wendroff method [1]. While the accuracy of the schemes depends on the choice of space- and time-step, all of the schemes used have agreed with the theoretical predictions on instability. This, along with previous studies [21, 4, 5], makes it clear that the fluctuations are not a result of numerical instability.

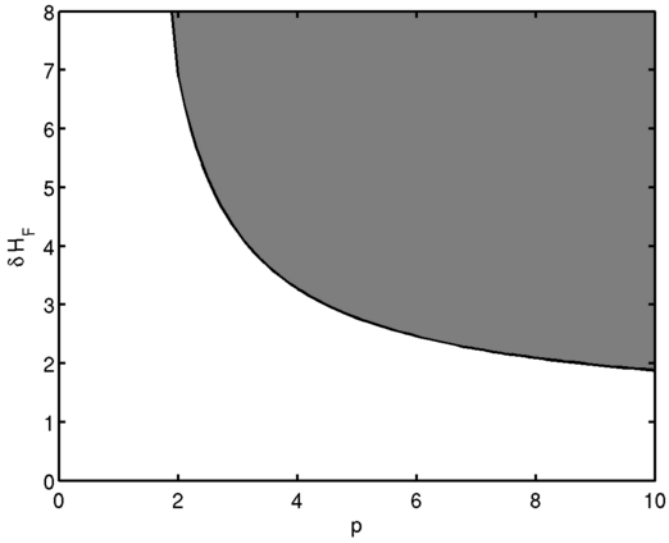


Figure 19. *For every value of p there exists a critical value of δH_F below which the system is stable. For values of δH_F greater than this critical value (gray region), there exists a domain in parameter space (similar to Figures 16 and 17) for which the system is unstable. This unstable domain grows in size as you move further into the gray region.*

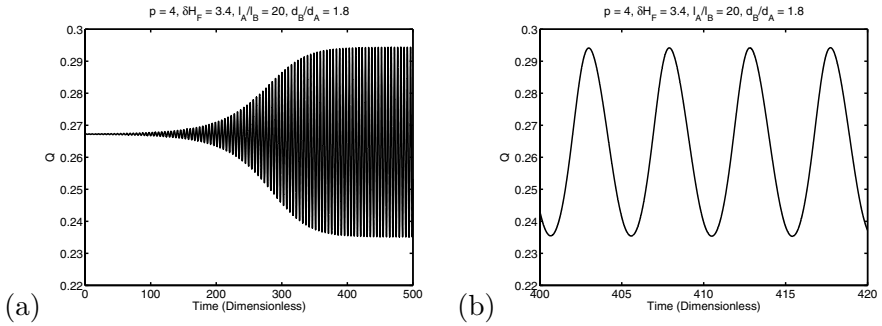


Figure 20. *Direct numerical simulation of (3.1)–(3.2). The system is initiated close to steady state and parameter values are chosen in the unstable domain.* (a) *The flow $Q(t)$ as a function of time demonstrating instability to a Hopf bifurcation with a growth rate close to that predicted by analysis.* (b) *A close-up of the oscillation showing that the period matches that predicted by analysis.*

In Figure 20a we plot the flow as a function of time. We used the following parameter values: $l_A/l_B = 20$, $d_B/d_A = 1.8$, $p = 4$, and $\delta H_F = 3.4$. At these parameter values the equilibrium solution is $Q \approx 0.2672$, which implies that $\gamma \approx 0.0591$. These parameter values place us in the unstable regions of Figures 16, 17, and 19. For these parameter values, there is a pair of complex conjugate roots of the characteristic equation with real and imaginary components $\sigma \approx 0.1728$ and $\omega \approx 4.883$. Recall that we scaled λ by $1/\tau$ in section 4. Undoing these scalings predicts a growth rate of approximately 0.046 and an oscillation period of approximately 4.82. Initially, we filled the vessels with hematocrit very close to the appropriate steady state values. We see from Figure 20b that the oscillation period is very close to that

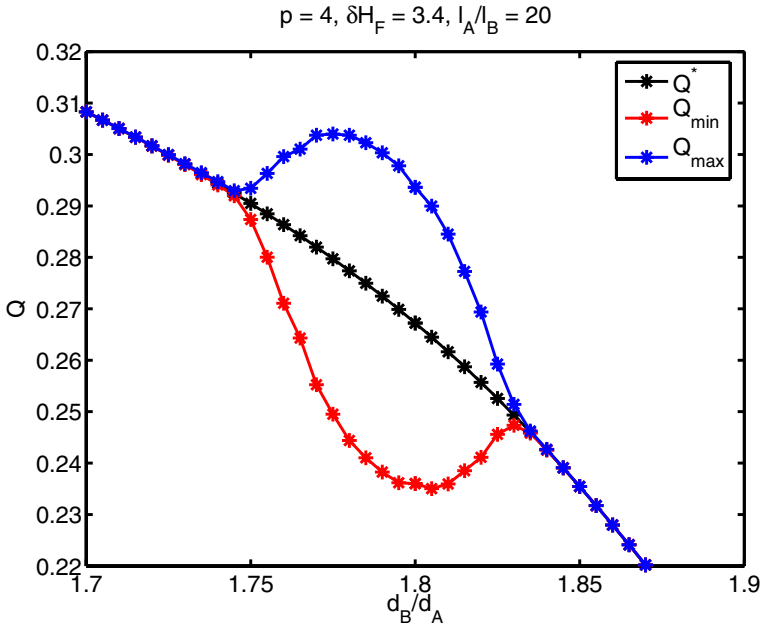


Figure 21. *A preliminary bifurcation diagram for changing diameter with $p = 4$, $\delta H_F = 3.4$, and $l_A/l_B = 20$. A periodic solution is initially obtained with $d_B/d_A = 1.8$, and the maximum and minimum values are shown in blue and red, respectively. The diameter ratio is then increased in increments of 0.005, the simulation is continued until the transients die off, and the minimum and maximum values are again recorded. This is repeated by decreasing the diameter ratio in increments of 0.005. The equilibrium solution is shown for reference in black.*

predicted by linear stability analysis.

We have also confirmed the existence of a finite region of instability in the $(l_A/l_B, d_B/d_A)$-plane by continuing the simulation for various values of d_B/d_A. In Figure 21 we plot the maximum and minimum values of the periodic solution as d_B/d_A is varied. The branch of periodic solutions is created by starting at $d_B/d_A = 1.8$ and then continuing the resulting periodic solution by increasing and decreasing d_B/d_A. The periodic solution emerges and vanishes at values of d_B/d_A which agree well with those predicted in section 6.2, and Figure 21 also suggests that the Hopf bifurcations are supercritical. We have not explored the nonlinear dynamics in detail but plan to in the future.

7. Conclusions. Numerous experimental studies have demonstrated that oscillations in microvascular blood flow are ubiquitous. While fluctuations in blood hematocrit and velocity are usually identified with biological control mechanisms such as vasomotion, there is evidence that oscillations may be possible in the absence of biological control. In a previous paper [5] we identified the simplest possible network, the two-node network, that could exhibit oscillatory dynamics, and we found and interpreted the three dimensionless parameters b, c, and γ which govern the stability of equilibrium.

In this paper we develop a complete description of the roots of the characteristic equation and their dependence on the three dimensionless parameters. We show that there exists an infinite number of Hopf bifurcation curves in the cb-plane for a given value of γ. Using simple

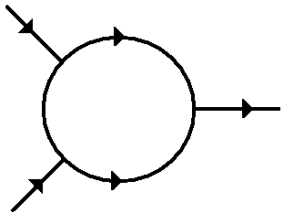


Figure 22. *The three-node network consists of two inlets and one outlet. One of the inlet nodes is converging, while the other is diverging. The flow in the middle branch could be in either direction.*

parametric models for the Fåhræus–Lindqvist effect and the plasma skimming effect, we map the bifurcation curves to the network parameter plane, $(l_A/l_B, d_B/d_A)$. For a given value of the plasma skimming parameter p we demonstrate that there is a critical value of the Fåhræus–Lindqvist parameter δ for instability. We also show that above this threshold value there exists a finite region of instability in the $(l_A/l_B, d_B/d_A)$-plane, and we confirm these predictions with direct numerical simulation.

It is worth noting that we do not pretend that the simple parametric models provide a realistic model for microvascular blood flow. Rather, we have used these models to provide insight by making the analysis tractable. There is still some question as to the possibility of instability if we use the more realistic models of Pries et al. The major difficulty is that there is no easy way to map from the five-dimensional network parameter space, $H_F, d_F, d_A, d_B, l_A/l_B$, to the three-dimensional dimensionless parameter space, b, c, γ. For a given set of network parameters it is simple to compute b, c, and γ and therefore determine whether the network is stable or unstable. This, however, is a rather unwieldy and unsatisfying method given the size of the five-dimensional parameter space. Our results to date indicate that instability is unlikely using the models of Pries et al., but we have no proof of this conjecture.

Even if oscillations are possible in the two-node network, the regions of instability in the network parameter space are likely to be very small. This makes the likelihood of confirming these predictions experimentally very remote. However, this parametric study does suggest several alternatives. As demonstrated by Kiani et al. [21] and Carr and LeCoin [4], oscillations are possible in larger networks. With a detailed understanding of the instability in hand, we can pursue the validation of this work in slightly larger networks. For example, we have preliminary evidence that the three-node network shown in Figure 22 has a much larger region of instability. A careful analysis of this network might pave the way for an in vitro experimental verification of these predictions.

Alternatively, we could explore other networks that are described by similar governing equations. While it is potentially difficult to find an analogy with both flow and hematocrit equivalents, we can reformulate our model into a single integral equation for the flow alone as follows. The resistance in any vessel depends on the average hematocrit $\bar{H}_i$ in vessel i. Since the hematocrit at any point in the vessel is determined by the inlet hematocrit at an earlier time, we can write the average hematocrit as

$$\bar{H}_i(t) = \frac{1}{l_i} \int_{t-\tau_i(t)}^{t} H_i(0,s) v_i(s)\, ds, \tag{7.1}$$

where l_i is the length of the vessel and $v_i(t)$ is the speed of flow in the vessel at time t. The time delay $\tau_i(t)$ is determined implicitly by the threshold condition,

$$\int_{t-\tau_i(t)}^{t} v_i(s)\,ds = l_i. \tag{7.2}$$

In the two-node network the flow $Q(t)$ is governed by (3.3),

$$Q(t) = \frac{R_B(\bar{H}_B(t))}{R_A(\bar{H}_A(t)) + R_B(\bar{H}_B(t))}, \tag{7.3}$$

which can now be viewed as an integral equation with two state-dependent delays since the entrance hematocrit, $H_i(0,t)$, is a function of the flow $Q(t)$ via the plasma skimming effect. Electric circuits provide an obvious analogy and may be worth investigating. In particular, the steady state IV-characteristic of a transistor can display *snapback* [34], which beautifully mimics the flow-pressure characteristic of blood flow in the two-node network. We hope to pursue both avenues of research in the near future.

Acknowledgments. We would like to thank the numerous undergraduate and graduate students at the University of New Hampshire and Olin College who have contributed in some way. These include David Cerra, Molly Goulet, Gabriel Withington, Yingyi Lao, David Gebhart, Kelly Butcher, and Keoni Mahelona. We would also like to thank Professor Mark Somerville of Olin College for his contribution regarding the voltage transfer characteristics of transistors.

REFERENCES

[1] W. F. Ames, *Numerical Methods for Partial Differential Equations*, 3rd ed., Academic Press, New York, 1992.
[2] J. Bélair and S. A. Campbell, *Stability and bifurcations of equilibria in a multiple-delayed differential equation*, SIAM J. Appl. Math., 54 (1994), pp. 1402–1424.
[3] G. Bugliarello and G. C. C. Hsiao, *Phase separation in suspensions flowing through bifurcations: A simplified hemodynamic model*, Science, 143 (1964), pp. 469–471.
[4] R. T. Carr and M. LeCoin, *Nonlinear dynamics in microvascular networks*, Annals of Biomedical Engineering, 28 (2000), pp. 641–652.
[5] R. T. Carr, J. B. Geddes, and F. Wu, *Oscillations in a simple microvascular network*, Annals of Biomedical Engineering, 33 (2005), pp. 764–771.
[6] S. Cavalcanti and M. Ursino, *Chaotic oscillations in microvessel arterial networks*, Annals of Biomedical Engineering, 24 (1996), pp. 37–47.
[7] S. Chien, C. D. Tvetenstrand, M. A. F. Epstein, and G. W. Schmid-Schönbein, *Model studies on distributions of blood cells at microvascular bifurcations*, American J. Physiology, 248 (1985), pp. H568–576.
[8] J. W. Dellimore, M. J. Dunlop, and P. B. Canham, *Ratio of cells and plasma in blood flowing past branches in small plastic channels*, American J. Physiology, 244 (1983), pp. H635–H643.
[9] L. E. El'sgol'ts and S. B. Norkin, *Introduction to the Theory and Application of Differential Equations with Deviating Arguments*, Academic Press, New York, 1973.
[10] G. Enden and A. S. Popel, *A numerical study of plasma skimming in small vascular bifurcations*, J. Biomechanical Engineering, 116 (1994), pp. 79–88.
[11] R. Fåhræus and T. Lindqvist, *The viscosity of blood in narrow capillary tubes*, American J. Physiology, 96 (1931), pp. 562–568.

[12] B. M. Fenton, R. T. Carr, and G. R. Cokelet, *Nonuniform red cell distribution in* 20 *to* 100 μm *bifurcations*, Microvascular Research, 29 (1985), pp. 103–126.
[13] Y. C. Fung, *Stochastic flow in capillary blood vessels*, Microvascular Research, 5 (1973), pp. 34–48.
[14] L. Glass, *Synchronization and rythmic processes in physiology*, Nature, 410 (2001), pp. 277–283.
[15] L. Glass and M. C. Mackey, *From Clocks to Chaos: The Rhythms of Life*, Princeton University Press, Princeton, NJ, 1988.
[16] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, *PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals*, Circulation, 101 (2000), pp. e215–e220.
[17] T. M. Griffith, *Temporal chaos in the microcirculation*, Cardiovascular Research, 31 (1996), pp. 342–358.
[18] J. K. Hale and W. Huang, *Global geometry of the stable regions for two delay differential equations*, J. Math. Anal. Appl., 178 (1993), pp. 344–362.
[19] N. D. Hayes, *Roots of the transcendental equations associated with a certain difference-differential equation*, J. London Math. Soc., 25 (1950), pp. 226–232.
[20] P. C. Johnson and H. Wayland, *Regulation of blood flow in single capillaries*, American J. Physiology, 212 (1967), pp. 1405–1415.
[21] M. F. Kiani, A. R. Pries, L. L. Hsu, I. H. Sarelius, and G. R. Cokelet, *Fluctuations in microvascular blood flow parameters caused by hemodynamic mechanisms*, American J. Physiology, 266 (1994), pp. H1822–H1828.
[22] B. Klitzman and P. C. Johnson, *Capillary network geometry and red cell distribution in hamster cremaster muscle*, American J. Physiology, 242 (1982), pp. H211–H219.
[23] A. Krogh, *The Anatomy and Physiology of Capillaries*, Yale University Press, New Haven, CT, 1922.
[24] A. Krogh, *Studies on the physiology of capillaries:* II. *The reactions to local stimuli of the blood-vessels in the skin and web of the frog*, J. Physiology, 55 (1921), pp. 414–422.
[25] J. M. Mahaffy, P. J. Zak, and K. M. Joiner, *A geometric analysis of stability regions for a linear differential equation with two delays*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 5 (1995), pp. 779–796.
[26] F. Mollica, R. K. Jain, and P. A. Netti, *A model for temporal heterogeneities of tumor blood flow*, Microvascular Research, 65 (2003), pp. 56–60.
[27] D. Parthimos, K. Osterloh, A. R. Pries, and T. M. Griffith, *Deterministic nonlinear characteristics of in vivo blood flow velocity and arteriolar diameter fluctuations*, Physics in Medicine and Biology, 49 (2004), pp. 1789–1802.
[28] A. S. Popel and P. C. Johnson, *Microcirculation and Hemorheology*, in Annual Review of Fluid Mechanics. Vol. 37, Annu. Rev. Fluid Mech. 37, Annual Reviews, Palo Alto, CA, 2005, pp. 43–69.
[29] A. R. Pries, K. Ley, and P. Gaehtgens, *Red cell distribution at microvascular bifurcations*, Microvascular Research, 38 (1989), pp. 81–101.
[30] A. R. Pries, D. Neuhaus, and P. Gaehtgens, *Blood viscosity in tube flow: Dependence on diameter and hematocrit*, American J. Physiology, 263 (1992), pp. H1770–H1778.
[31] A. R. Pries, T. W. Secomb, T. Gessner, M. B. Sperandio, J. F. Gross, and P. Gaehtgens, *Resistance to blood flow in microvessels in vivo*, Circulation Research, 75 (1994), pp. 904–915.
[32] G. P. Rodgers, A. N. Schechter, C. T. Noguchi, H. G. Klein, A. W. Niehuis, and R. F. Bonner, *Periodic microcirculatory flow in patients with sickle-cell disease*, New England J. Medicine, 311 (1984), pp. 1534–1538.
[33] S. P. Sutera and R. Skalak, *The history of Poiseuille's law*, in Annual Review of Fluid Mechanics. Vol. 25, Annu. Rev. Fluid Mech. 25, Annual Reviews, Palo Alto, CA, 1993, pp. 1–19.
[34] S. H. Voldman, *ESD: Circuits and Devices*, John Wiley and Sons, New York, 2006.
[35] C. Wiederhielm, J. W. Woodbury, S. Kirk, and R. F. Rushmer, *Pulsatile pressures in the microcirculation of frog mesentery*, American J. Physiology, 207 (1964), pp. 173–176.
[36] K. P. Yip, N. H. Nolstein-Rathlou, and D. J. Marsh, *Chaos in blood flow control in genetic and renovascular hypertensive rats*, American J. Physiology, 261 (1991), pp. F400–F408.

SIAM J. Applied Dynamical Systems
Vol. 6, No. 4, pp. 728–758

Dynamics on Networks of Cluster States for Globally Coupled Phase Oscillators*

Peter Ashwin†, Gábor Orosz†, John Wordsworth†, and Stuart Townley†

Abstract. Systems of globally coupled phase oscillators can have robust attractors that are heteroclinic networks. We investigate such a heteroclinic network between partially synchronized states where the phases cluster into three groups. For the coupling considered there exist 30 different three-cluster states in the case of five oscillators. We study the structure of the heteroclinic network and demonstrate that it is possible to navigate around the network by applying small impulsive inputs to the oscillator phases. This paper shows that such navigation may be done reliably even in the presence of noise and frequency detuning, as long as the input amplitude dominates the noise strength and the detuning magnitude, and the time between the applied pulses is in a suitable range. Furthermore, we show that, by exploiting the heteroclinic dynamics, frequency detuning can be encoded as a spatiotemporal code. By changing a coupling parameter we can stabilize the three-cluster states and replace the heteroclinic network by a network of excitable three-cluster states. The resulting "excitable network" has the same structure as the heteroclinic network and navigation around the excitable network is also possible by applying large impulsive inputs. We also discuss features that have implications for related models of neural activity.

Key words. globally coupled oscillators, three-cluster state, heteroclinic connection/network, winnerless competition, excitable dynamics, neural computation

AMS subject classifications. 34C15, 37C29, 37G40, 92B20

DOI. 10.1137/070683969

1. Introduction.

Coupled phase oscillators are a simple class of dynamical systems that can model phenomena such as synchronization between individual neurons [7, 16]. If neurons are assumed to produce periodic output (e.g., periodic spiking) when uncoupled, then they may be represented by a scalar phase variable. (In case of spiking a phase can be defined such that it crosses zero when the neuron spikes.) Coupled oscillators that are subject to external perturbations can be modeled in many cases by using only the phase variables [8].

Recent work [1] proposes that certain types of neural system are well modeled by what they call *winnerless competition*. Such dynamics consists of a number of saddle states in the phase space that are connected by their unstable manifolds to form a *heteroclinic network*. This type of dynamics can robustly produce a number of behaviors in the phase oscillator system that are of interest for modeling neural systems:

(i) it can generate a sequence of states in response to a sequence of inputs,
(ii) it can act as a very sensitive classifier for inputs, and

*Received by the editors February 28, 2007; accepted for publication (in revised form) by B. Ermentrout June 5, 2007; published electronically December 7, 2007. This research was supported by the EPSRC grant EP/C510771/1.
http://www.siam.org/journals/siads/6-4/68396.html

†Mathematics Research Institute, School of Engineering, Computer Science and Mathematics, University of Exeter, EX4 4QF, United Kingdom (p.ashwin@exeter.ac.uk, g.orosz@exeter.ac.uk, j.wordsworth@exeter.ac.uk, s.b.townley@exeter.ac.uk). The last author was supported by the Leverhulme Trust Research Fellowship RF/9/RFG/2005/0153.

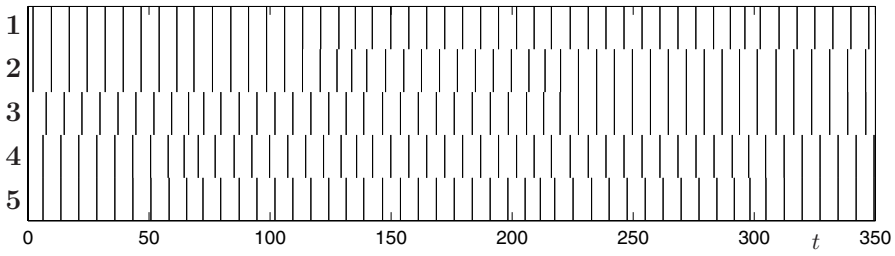


Figure 1. *Raster plot showing the spiking of five oscillators in time. The vertical lines indicate when the phases of oscillators* $1,\dots,5$ *pass through zero. The figure is obtained using the coupled phase oscillator model* (2.1), (2.2) *with parameters* (3.1) *in the case of low noise (of strength* $\eta = 10^{-5}$*), no input (*$\epsilon = 0$*), and no detuning (*$\delta = 0$*). Observe that the oscillators form into three-cluster states of two synchronized pairs and one singleton, and that the cluster formations spontaneously change as time progresses. The dynamics in this figure is shown also in Figure* 9(a2)*, where clusters are distinguished by color.*

(iii) it can convert inputs into spatiotemporal coding;

see [6, 23]. The last behavior may have importance for encoding odor information in antennal lobes of insects [11].

Attracting robust heteroclinic networks have been recognized in a variety of systems. For instance, they are found in the Lotka–Volterra dynamics of three or more interacting populations [18], in many symmetric systems [12], and in systems of four or more globally coupled phase oscillators [5] where they cause the phenomena of *slow oscillations/switching* among cluster states [13, 17, 21]. In case of slow oscillations/switching, the presence of noise causes a near-periodic switching motion such that the asymptotic period of switching becomes unbounded as the noise level reduces to zero.

To motivate our work we present here an example of slow switching. Figure 1 shows a time series of "spikes" for five oscillators which is a widely accepted representation of the dynamics in neural systems [23]. This figure is produced using the phase oscillator model studied in this article with added low noise. Vertical lines show where the phases of oscillators pass through zero. From this "raster plot" the underlying dynamics is not immediately obvious. However, one can recognize that three-cluster states are formed consisting of two synchronized pairs of oscillators and one singleton, and that there are regular switches between these cluster states as time progresses. This figure corresponds to Figure 9(a2), where the clusters are colored yellow and blue (and different time scales are used). Our aim is to show the existence of a heteroclinic network of cluster states which underlies and explains this dynamics. By exploiting the features of this network we can predict possible transitions between cluster states and estimate the period of switching.

Even for quite simple coupling between fairly small numbers of globally coupled phase oscillators one can find attracting robust (structurally stable) heteroclinic networks of high complexity. For example, a system of five oscillators may possess an attractor with 20 different two-cluster states linked together as a heteroclinic network [2]. This network can be used to perform simple computations by applying small inputs to the individual oscillators [3], but these computations are very sensitive to noise. For this reason, we introduce a modified model in section 2 which presents a new type of heteroclinic network between three-cluster states and where the performed computations are considerably less sensitive to noise. The goal of

this paper is to describe the structure and dynamics of this network and also to investigate its computational properties.

In section 3 we examine the structure of the proposed heteroclinic network in detail. For five oscillators we find an open set of parameters such that the system has a robust heteroclinic attractor comprising 30 symmetrically related three-cluster states of saddle type connected by their unstable manifolds. We also investigate the behavior of the network under perturbations. Namely, we consider the effect of noise, the effect of detuning the natural frequencies of the oscillators, and the effect of (periodic) impulsive inputs.

In section 4 we demonstrate that the heteroclinic network can robustly perform finite-state computation in the presence of background noise and frequency detuning. By applying periodic impulsive inputs to the phase oscillators it is possible to induce switches between cluster states and so predictably navigate around the network. We characterize the bit error rate of these switches as a function of the input period, the input amplitude, and the noise strength. It is shown that as long as the input amplitude is large enough (compared to the noise strength) there exists a range of periods for which the computations are reliable. The left boundary of this range is proportional to the logarithm of the input amplitude, while its right boundary is proportional to the logarithm of the noise strength. (Frequency detuning also brings errors into the finite-state computation similarly to noise.) At the end of section 4 we also discuss that frequency detuning can be encoded as spatiotemporal codes.

Section 5 shows that on varying a single parameter in the coupling function one can achieve that the above attracting heteroclinic network becomes an attracting network of linearly stable three-cluster states. These states are excitable: there is a minimum threshold for the input amplitude to switch from one state to the other. This "excitable network" can perform finite-state computation similarly to the heteroclinic network, except that the input amplitudes must be sufficiently large.

We conclude our results in section 6 and discuss future research in section 7.

2. The phase oscillator model. In this section we introduce the globally coupled phase oscillator model we study in this paper. The first model of this kind was developed by Kuramoto [19], but models with several different coupling functions have been investigated in the last three decades; see, e.g., [2, 7, 13]. We consider N oscillators complying with the equations

$$\dot{\theta_n} = \omega_n + \frac{1}{N}\sum_{m=1}^{N} g(\theta_n - \theta_m) + \epsilon I_n(t) + \eta\, w_n, \quad n = 1, \ldots, N, \tag{2.1}$$

where $\theta_n(t) = \theta_n(t + 2\pi)$ is the phase of the nth oscillator, ω_n is the *natural frequency* of the nth oscillator, $I_n(t)$ is an *impulsive input* with unit magnitude, and $w_n(t)$ is *uncorrelated white noise* such that the associated random walk has unit growth of variance per unit time. The quantity ϵ represents the *input amplitude*, while η is the *noise strength*. The globally coupled system (2.1), (2.2) is represented by the graphs in Figure 2 for $N = 3, 5, 7$ oscillators (black dots) with $N(N-1)/2 = 3, 10, 21$ bidirectional connections (two-headed arrows). We remark that we are interested only in the case when N is finite. In particular, $N = 5$ is studied in detail, and implications for larger N are discussed in section 7. In some cases we will use the

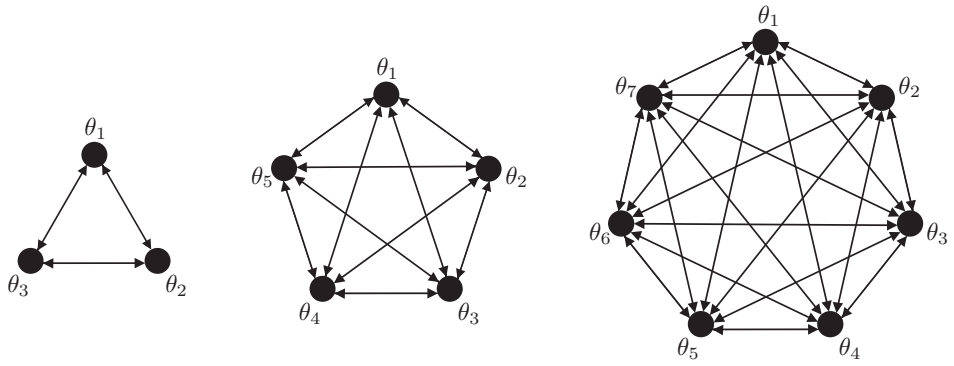


Figure 2. *Global coupling between $N = 3, 5, 7$ oscillators with $N(N-1)/2 = 3, 10, 21$ bidirectional connections, respectively. Black dots represent the oscillators, and two-headed arrows represent the mutual coupling between them.*

vector notation $\theta(t) = \mathrm{col}[\theta_1(t), \ldots, \theta_N(t)]$. Furthermore, we consider a *coupling function*

$$g(\varphi) = -\sin(\varphi + \alpha) + r\sin(2\varphi + \beta), \tag{2.2}$$

where α, β, and r are constants. This function is 2π-periodic containing Fourier modes with frequencies 1 and 2. Note that in order to find nontrivial clustering behavior in the system (2.1), (2.2) it is necessary to include Fourier modes with frequencies higher than 1 [13]. Furthermore, the phase reductions of well-known neural models naturally lead to periodic functions where several harmonics are included [14]. Function (2.2) is the simplest generalization of the functions used in [2, 13, 17].

One may define the *average natural frequency* by

$$\omega = \frac{1}{N}\sum_{n=1}^{N}\omega_n, \tag{2.3}$$

and a *uniform detuning of natural frequencies* is when

$$\omega_{n+1} - \omega_n = \delta, \quad n = 1, \ldots, N-1. \tag{2.4}$$

Here $\delta \ll 1$ represents the *detuning magnitude*. Observe that for uniform detuning given ω and δ we have

$$\omega_n = \omega + \big(n - \tfrac{N+1}{2}\big)\delta, \quad n = 1, \ldots, N. \tag{2.5}$$

Permuting the oscillators in (2.4), and consequently in (2.5), gives detunings which we also call uniform. In the case of nonuniform detunings the frequencies ω_n are not uniformly spaced and the detuning magnitude may be defined by

$$\delta = \max_n\big\{\min_{m\neq n}\{|\omega_n - \omega_m|\}\big\}. \tag{2.6}$$

In this article we investigate only the effects of uniform detunings. Note that $\delta = 0$ if and only if $\omega_n = \omega$ for $n = 1, \ldots, N$.

One may consider the detuning as a steady external input applied to the "ideal" system of identical oscillators ($\omega_n = \omega$ for all n) in addition to impulsive inputs and noise in (2.1). The effects of the external perturbations in the form of impulsive inputs (when $\epsilon \neq 0$), noise (when $\eta \neq 0$), and detuning (when $\delta \neq 0$) will be discussed in section 3.3.

In the "unperturbed" case

$$\epsilon = \eta = \delta = 0, \tag{2.7}$$

the systems of coupled oscillators (2.1), (2.2) can be thought of as ODEs on the N-torus $\mathbb{T}^N$ with symmetry S_N of all permutations of the oscillators. The codimension one invariant subspaces

$$\theta_n(t) \equiv \theta_m(t) \qquad \text{for} \qquad n \neq m \tag{2.8}$$

form barriers to the flow. These divide the torus up into $(N-1)!$ invariant regions that are symmetric images of the canonical invariant region

$$\{\theta_1 \leq \theta_2 \leq \cdots \leq \theta_N \leq \theta_1 + 2\pi\}; \tag{2.9}$$

see [5] for more details. Subspaces (2.8) are essential for the formation of the robust heteroclinic network, as will be explained in section 3.1.

We refer to an *ℓ-cluster state of type* $(N_1, \ldots, N_\ell)$ as a state where each oscillator has one of only ℓ different phases $\phi_1, \ldots, \phi_\ell$: for each r there are N_r oscillators that have the same phase ϕ_r such that $\sum_{r=1}^{\ell} N_r = N$. In this way we can think of full synchrony as a one-cluster state of type (N) and full asynchrony as an N-cluster state of type $(1, \ldots, 1)$ with N entries. Note that choosing a clustered distribution of frequencies ω_n can give rise to cluster states of similar type. However, this is not necessary as cluster states may appear spontaneously as stable behavior, even for identical oscillators ($\omega_n = \omega$ for all n), as will be discussed in detail in section 5.

We wish to emphasize the importance of the parameter β in the coupling function (2.2). In [2, 13, 17] $\beta = 0$ was considered; i.e., only the odd part of the second harmonic was included. Due to this "degeneracy" of the coupling function it appears that any attracting heteroclinic network contains only two-cluster states. In this paper we unfold the above "degeneracy" by considering $\beta \neq 0$ and find heteroclinic networks between three-cluster states.

Because the right-hand side of (2.1) depends only on phase differences, there is an extra S^1 continuous rotational symmetry given by performing an identical phase shift to all oscillators. If there is a solution $\tilde{\theta}(t) = \mathrm{col}[\tilde{\theta}_1(t), \ldots, \tilde{\theta}_N(t)]$ of (2.1) then $\tilde{\theta}(t) + \varphi\,\mathrm{col}[1, \ldots, 1] = \mathrm{col}[\tilde{\theta}_1(t)+\varphi, \ldots, \tilde{\theta}_N(t)+\varphi]$ is also a solution for any $\varphi \in \mathbb{R}$. Consequently, it is sufficient to examine phase differences to determine the long-term behavior of the system. We will introduce the phase differences $\gamma_n = \theta_n - \theta_N$ for $n = 1, \ldots, N$. Note that similar continuous symmetries appear in many other applications, e.g., in car-following models of highway traffic [22].

3. Attracting network of three-cluster states.

In the unperturbed case $\epsilon = \eta = \delta = 0$ system (2.1), (2.2) displays a wide range of robust dynamics. This includes synchrony (r, α small regardless of β), antisynchrony (r small α near π regardless of β), and cluster states (typically in the region where α is close to $\pi/2$). In the last case one can find robust heteroclinic

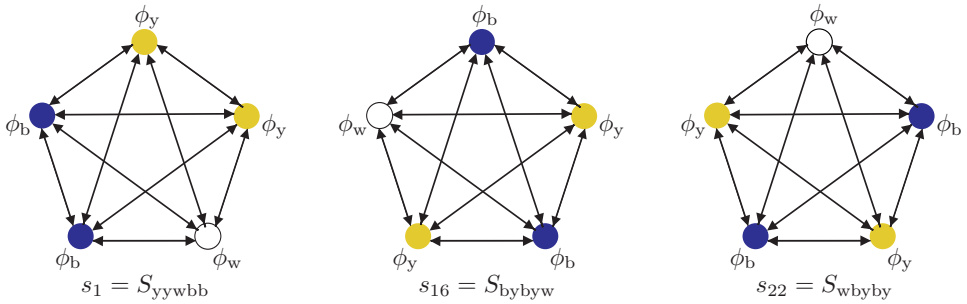


Figure 3. *Three of the* 30 *three-cluster states of Table* 1 *are represented by color. The pairs marked* "y" *are shown as yellow, the pairs marked* "b" *are shown as blue, and the single oscillators are displayed as white.*

connections. For $\beta = 0$ these connections are between two-cluster states [2, 17], but the system with $\beta \neq 0$ can have connections between three-cluster states.

Considering $N = 5$ the system (2.1), (2.2) has an attracting heteroclinic network for an open set of parameters near

$$r = 0.2, \quad \alpha = 1.8, \quad \beta = -2.0, \quad \omega = 1.0, \tag{3.1}$$

as explained below. Considering the invariant subspace

$$\begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ \theta_3(t) \\ \theta_4(t) \\ \theta_5(t) \end{bmatrix} = \begin{bmatrix} \theta_1(t) \\ \theta_1(t) \\ \theta_3(t) \\ \theta_4(t) \\ \theta_4(t) \end{bmatrix}, \tag{3.2}$$

where the first two oscillators form a cluster ($\theta_1(t) \equiv \theta_2(t)$) and the last two oscillators form another cluster ($\theta_4(t) \equiv \theta_5(t)$), one can find a three-cluster state of type $(2, 1, 2)$

$$\begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ \theta_3(t) \\ \theta_4(t) \\ \theta_5(t) \end{bmatrix} = \begin{bmatrix} \Omega t + \phi_y \\ \Omega t + \phi_y \\ \Omega t + \phi_w \\ \Omega t + \phi_b \\ \Omega t + \phi_b \end{bmatrix} := S_{yywbb} \tag{3.3}$$

for some constants $\phi_y < \phi_w < \phi_b < \phi_y + 2\pi$. Note that the subscripts of S_{yywbb} correspond to the subscripts of the coordinates where "y," "w," and "b" stand for yellow, white, and blue, respectively; see the coloring applied in the leftmost panel of Figure 3. In this article we refer to this coloring by using terms such as *yellow oscillator/cluster*, *white oscillator*, and *blue oscillator/cluster*.

In addition there are a number of states given by permutation of the subscripts of S_{yywbb}, that is, by permutation of the coordinates in (3.3). This symmetry gives a total of $5!/(2!1!2!) = 30$ possible states listed in Table 1. Some of these states are also represented by the graphs in Figure 3 with corresponding coloring.

Table 1

List of three-cluster states. Each column may be generated from another by cyclic permutation of the subscripts of $S_{i_1,\ldots,i_5}$.

$s_1 = S_{\mathrm{yywbb}}$	$s_7 = S_{\mathrm{byywb}}$	$s_{13} = S_{\mathrm{bbyyw}}$	$s_{19} = S_{\mathrm{wbbyy}}$	$s_{25} = S_{\mathrm{ywbby}}$
$s_2 = S_{\mathrm{bbwyy}}$	$s_8 = S_{\mathrm{ybbwy}}$	$s_{14} = S_{\mathrm{yybbw}}$	$s_{20} = S_{\mathrm{wyybb}}$	$s_{26} = S_{\mathrm{bwyyb}}$
$s_3 = S_{\mathrm{ybwyb}}$	$s_9 = S_{\mathrm{bybwy}}$	$s_{15} = S_{\mathrm{ybybw}}$	$s_{21} = S_{\mathrm{wybyb}}$	$s_{27} = S_{\mathrm{bwyby}}$
$s_4 = S_{\mathrm{bywby}}$	$s_{10} = S_{\mathrm{ybywb}}$	$s_{16} = S_{\mathrm{bybyw}}$	$s_{22} = S_{\mathrm{wbyby}}$	$s_{28} = S_{\mathrm{ywbyb}}$
$s_5 = S_{\mathrm{ybwby}}$	$s_{11} = S_{\mathrm{yybwb}}$	$s_{17} = S_{\mathrm{byybw}}$	$s_{23} = S_{\mathrm{wbyyb}}$	$s_{29} = S_{\mathrm{bwbyy}}$
$s_6 = S_{\mathrm{bywyb}}$	$s_{12} = S_{\mathrm{bbywy}}$	$s_{18} = S_{\mathrm{ybbyw}}$	$s_{24} = S_{\mathrm{wybby}}$	$s_{30} = S_{\mathrm{ywybb}}$

Notice that the three-cluster state (3.3) (and each symmetric copy in Table 1) has full frequency synchrony with frequency Ω and clustering into three-cluster states with phases ϕ_{y}, ϕ_{w}, and ϕ_{b}. One of these phases can be chosen arbitrarily corresponding to the S^1 continuous rotational symmetry of the system. This means that the three-cluster state (3.3) (and each symmetric copy in Table 1) is determined by the phase differences

$$\begin{aligned} \chi &:= \phi_{\mathrm{y}} - \phi_{\mathrm{w}}, \qquad \text{where } -\pi < \chi < 0, \\ \psi &:= \phi_{\mathrm{b}} - \phi_{\mathrm{w}}, \qquad \text{where } 0 < \psi < \pi, \end{aligned} \tag{3.4}$$

up to application of the continuous rotational symmetry. Substituting the solution (3.3) into (2.1), (2.2) when $\epsilon = \eta = \delta = 0$ and using (3.4), one obtains

$$\begin{aligned} \Omega &= \omega + \tfrac{1}{5}\Big(g(0) + 2g(-\chi) + 2g(-\psi)\Big), \\ \Omega &= \omega + \tfrac{1}{5}\Big(2g(0) + g(\chi) + 2g(\chi - \psi)\Big), \\ \Omega &= \omega + \tfrac{1}{5}\Big(2g(0) + g(\psi) + 2g(\psi - \chi)\Big), \end{aligned} \tag{3.5}$$

which determine χ, ψ, and Ω. Note that when solving (3.5) the frequency ω does not appear in the phase differences χ and ψ. For parameters (3.1) there is a unique solution:

$$\chi = -1.8212, \quad \psi = 1.1041, \quad \Omega = 0.8468. \tag{3.6}$$

Linearizing system (2.1), (2.2) about the three-cluster state $s_1 = S_{\mathrm{yywbb}}$ in (3.3) (or about any symmetric copy in Table 1) and using (3.4), one can investigate the linear stability of three-cluster states, that is, determine the eigenvalues

$$\begin{aligned} \lambda_1 &= 0, \\ \lambda_2 &= \tfrac{1}{5}\Big(2g'(0) + g'(\chi) + 2g'(\chi - \psi)\Big), \\ \lambda_3 &= \tfrac{1}{5}\Big(2g'(0) + g'(\psi) + 2g'(\psi - \chi)\Big), \\ \lambda_4 &= \lambda_5^* = \mu + \mathrm{i}\sqrt{\nu} \end{aligned} \tag{3.7}$$

of the linearized system. The rather complicated expressions of $\mu, \nu \in \mathbb{R}^+$ in terms of $g'(\chi)$, $g'(-\chi)$, $g'(\psi)$, $g'(-\psi)$, $g'(\chi - \psi)$, and $g'(\psi - \chi)$ are given by (A.1), (A.2) in Appendix A. For parameters (3.1) (that is, for the phase differences (3.6)) we have the eigenvalues

$$\lambda_2 = -0.2834, \quad \lambda_3 = 0.1703, \quad \lambda_{4,5} = -0.1012 \pm \mathrm{i}\,0.2848, \tag{3.8}$$

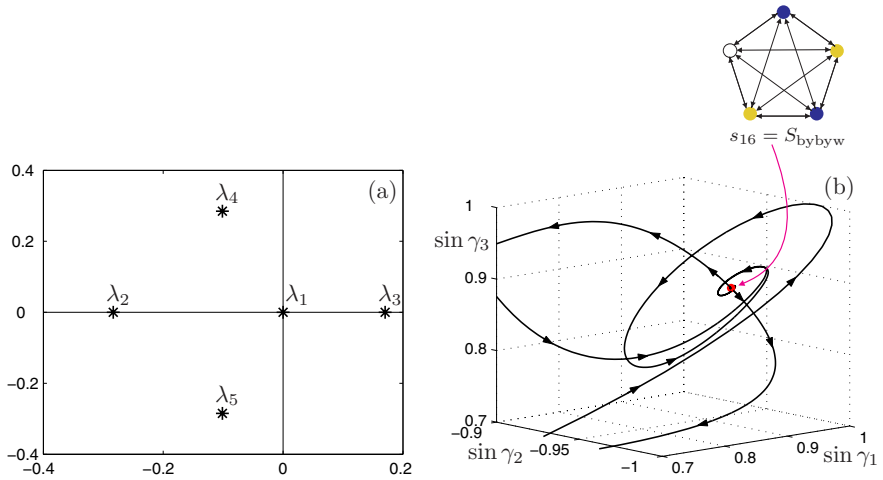


Figure 4. *Dynamics around a saddle three-cluster state for parameters* (3.1) *in the unperturbed case* $\epsilon = \eta = \delta = 0$. *In panel* (a) *the eigenvalues* (3.7), (3.8) *are displayed in the complex plane. Panel* (b) *shows the local dynamics in phase space in the vicinity of the three-cluster state* $s_{16} = S_{\text{bybyw}}$ *(red dot): The unstable eigendirection* v_3 *and the stable eigendirections* $\text{Re}(v_4) = \text{Re}(v_5)$, $\text{Im}(v_4) = -\text{Im}(v_5)$ *given in* (3.9) *can be visualized by the outgoing and spiralling-in trajectories, respectively.*

shown in Figure 4(a). For the three-cluster state $s_1 = S_{\text{yywbb}}$ the corresponding eigenvectors are

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \quad v_4 = v_5^* = \begin{bmatrix} r_{\text{y}} + \text{i}\, p_{\text{y}} \\ r_{\text{y}} + \text{i}\, p_{\text{y}} \\ 1 \\ r_{\text{b}} + \text{i}\, p_{\text{b}} \\ r_{\text{b}} + \text{i}\, p_{\text{b}} \end{bmatrix}, \tag{3.9}$$

where the expressions of $r_{\text{y}}, p_{\text{y}}, r_{\text{b}}, p_{\text{b}} \in \mathbb{R}$ contain $g'(\chi)$, $g'(-\chi)$, $g'(\psi)$, $g'(-\psi)$, $g'(\chi - \psi)$, and $g'(\psi - \chi)$ as given by (A.3)–(A.6) in Appendix A. For the other 29 states in Table 1, the eigenvectors can be obtained by permuting the components in (3.9) (in the same manner as the subscripts of $s_1 = S_{\text{yywbb}}$ to obtain these symmetric copies).

In order to eliminate the S^1 continuous rotational symmetry we use the phase differences

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{bmatrix} = \begin{bmatrix} \theta_1 - \theta_5 \\ \theta_2 - \theta_5 \\ \theta_3 - \theta_5 \\ \theta_4 - \theta_5 \\ 0 \end{bmatrix} \tag{3.10}$$

to plot the dynamics of the system. In the phase space of these coordinates the three-cluster state $s_1 = S_{\text{yywbb}}$ in (3.3) and its symmetrical copies in Table 1 are equilibria. Furthermore, in many of the subsequent figures we project the phase space $\mathbb{T}^4$ of the phase differences γ_n into $[-1, 1]^4$ by taking $\sin \gamma_n$.

The eigenvector v_1 with zero eigenvalue corresponds to translation along the S^1 group orbit of the system. The eigenvectors v_2 and v_3 correspond to splitting the yellow and the blue clusters, respectively. (These clusters are marked by subscripts "y" and "b" in (3.3) and Table 1, and are colored yellow and blue in Figures 3 and 4.) The eigenvectors v_4 and v_5 correspond to motion in the subspace (3.2). In Figure 4(b) the eigendirection v_3 and the eigendirections $\mathrm{Re}(v_4) = \mathrm{Re}(v_5)$, $\mathrm{Im}(v_4) = -\mathrm{Im}(v_5)$ can be visualized by the outgoing and spiralling-in trajectories, respectively, for the three-cluster state $s_{16} = S_{\mathrm{bybyw}}$ (red dot).

For parameters (3.1) we have $\lambda_1 = 0$, $\lambda_2 < 0$, $\lambda_3 > 0$, and $\mathrm{Re}(\lambda_4) = \mathrm{Re}(\lambda_5) < 0$ as given by (3.8) and plotted in Figure 4(a). Consequently, the eigendirection v_2 is attracting; that is, perturbations that involve splitting the yellow cluster will decay in time. Hence we say that *the yellow cluster is stable*. On the other hand, the eigendirection v_3 is repelling as the outgoing trajectories demonstrate in Figure 4(b). That is, perturbations that involve splitting the blue cluster grow in time. Hence we say that *the blue cluster is unstable*. The eigendirections $\mathrm{Re}(v_4) = \mathrm{Re}(v_5)$ and $\mathrm{Im}(v_4) = -\mathrm{Im}(v_5)$ are attracting, as demonstrated by the inward spiralling trajectories in Figure 4(b). That is, the three-cluster state is attracting in the subspace (3.2). In summary, the only perturbations giving rise to motion away from the saddle three-cluster state are those in the direction v_3 that involve splitting the blue cluster.

3.1. Heteroclinic connections between three-cluster states. One can verify numerically that the branches of the one-dimensional unstable manifold of the three-cluster state $s_1 = S_{\mathrm{yywbb}}$ are contained wholly within the stable manifolds of the three-cluster states $s_{12} = S_{\mathrm{bbywy}}$ and $s_{13} = S_{\mathrm{bbyyw}}$. The former branch corresponds to the direction where the 4th oscillator is advanced relative to the 5th one, i.e., "starts in direction v_3," while the latter branch corresponds to the opposite direction, i.e., "starts in direction $-v_3$"; see (3.9). These connections occur within the invariant subspace

$$\begin{bmatrix}\theta_1(t)\\ \theta_2(t)\\ \theta_3(t)\\ \theta_4(t)\\ \theta_5(t)\end{bmatrix} = \begin{bmatrix}\theta_1(t)\\ \theta_1(t)\\ \theta_3(t)\\ \theta_4(t)\\ \theta_5(t)\end{bmatrix}, \tag{3.11}$$

and the connections are source-to-sink and hence robust (persistent) to sufficiently small symmetry-preserving perturbations. Note that on the connecting orbit the first two oscillators remain synchronized ($\theta_1(t) \equiv \theta_2(t)$), but this cluster becomes unstable (changes from yellow to blue).

Similarly, one branch of the one-dimensional unstable manifold of each of the three-cluster states $s_{19} = S_{\mathrm{wbbyy}}$ and $s_{29} = S_{\mathrm{bwbyy}}$ connects to $s_1 = S_{\mathrm{yywbb}}$. These branches "arrive from directions $\mp v_2$," and they are embedded in the invariant subspace

$$\begin{bmatrix}\theta_1(t)\\ \theta_2(t)\\ \theta_3(t)\\ \theta_4(t)\\ \theta_5(t)\end{bmatrix} = \begin{bmatrix}\theta_1(t)\\ \theta_2(t)\\ \theta_3(t)\\ \theta_4(t)\\ \theta_4(t)\end{bmatrix}, \tag{3.12}$$

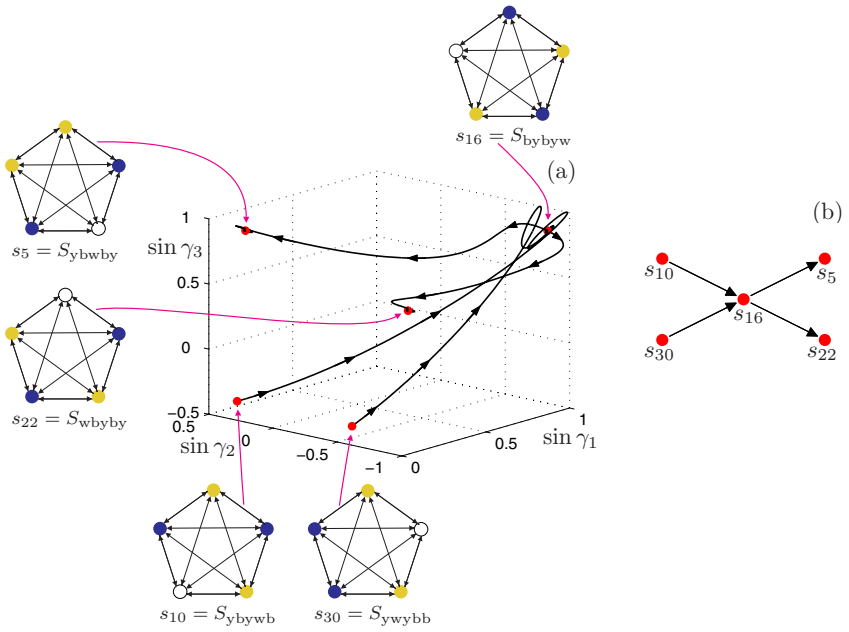


Figure 5. *In panel* (a) *the heteroclinic connections originated in and terminated at the three-cluster state* $s_{16} = S_{\text{bybyw}}$ *are shown in phase space for parameters* (3.1) *in the unperturbed case* $\epsilon = \eta = \delta = 0$. *The three-cluster states are represented by red dots. In panel* (b) *a simple graph representation is displayed such that the nodes and directed edges correspond to three-cluster states and heteroclinic connections, respectively.*

where the last two oscillators are synchronized ($\theta_4(t) \equiv \theta_5(t)$).

Indeed, the above argument can be interpreted for any three-cluster state in Table 1. For example, considering the three-cluster state $s_{16} = S_{\text{bybyw}}$, one may verify that the branches of its unstable manifold connect to the three-cluster states $s_5 = S_{\text{ybwby}}$ and $s_{22} = S_{\text{wbyby}}$, while the unstable manifolds of the three-cluster states $s_{10} = S_{\text{ybywb}}$ and $s_{30} = S_{\text{ywybb}}$ connect to it, as shown in Figure 5(a). The connected three-cluster states can be represented as nodes, while the connections can be represented as directed edges between them, as shown by the graph in Figure 5(b).

3.2. Graph structure of the heteroclinic network. Let us now examine all connections between three-cluster states in Table 1. We find that there is a heteroclinic network consisting of the 30 three-cluster states $s_i = S_{i_1 \ldots i_5}$ and their 60 connecting heteroclinic orbits; see Figure 6(a), where all states and connections are shown in phase space. The network can be represented as the directed graph shown in Figure 6(b). The nodes s_i of the graph correspond to the three-cluster states $s_i = S_{i_1 \ldots i_5}$, and there is a directed edge from node s_i to s_j if and only if there is connecting orbit from the three-cluster state $s_i = S_{i_1 \ldots i_5}$ to $s_j = S_{j_1 \ldots j_5}$. Notice the five-fold cyclic symmetry of the graph that reflects a five-fold cyclic permutation symmetry of the oscillators; see the middle panel in Figure 2.

Starting at a certain three-cluster state of the heteroclinic network in Figure 6(a) (or equivalently at the corresponding node of the graph in Figure 6(b)), we can choose which connection (or corresponding edge) we wish to follow. More precisely, an arbitrarily small perturbation with a component in the direction v_3 or $-v_3$ initiates a transition along one or

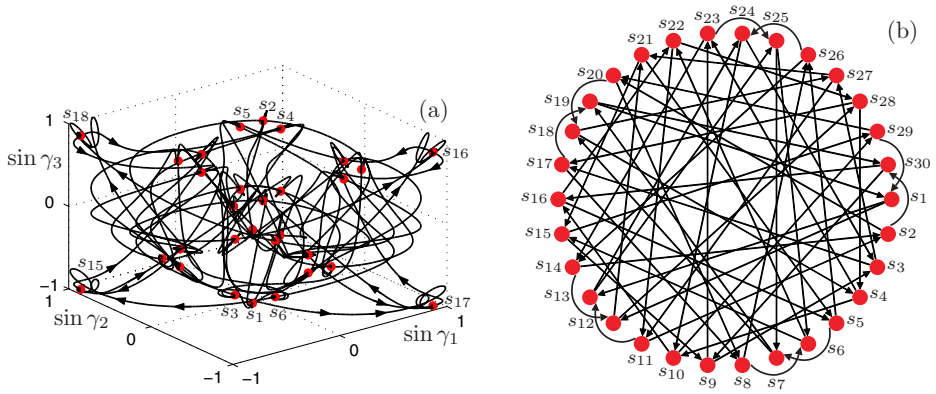


Figure 6. *Panel* (a) *shows the entire heteroclinic network in phase space for parameters* (3.1) *in the unperturbed case* $\epsilon = \eta = \delta = 0$. *Some of the three-cluster states (red dots) are labeled according to Table* 1. *In panel* (b) *the graph representation is depicted where each node represents a three-cluster state and each directed edge represents a heteroclinic connection from one state to the other.*

the other heteroclinic connection. Indeed, these perturbations correspond to advancing one oscillator of the unstable blue cluster (marked by "b" in the subscript of $s_i = S_{i_1 \ldots i_5}$ in (3.3) and Table 1 and shown as blue in Figures 3–5). This means that a small advance of phase of the pth oscillator gives rise to a transition from s_i to s_j if $i_p = \mathrm{b}$ (we perturb the unstable blue cluster) and there exists a heteroclinic connection between the three-cluster states $s_i = S_{i_1,\ldots,i_5}$ and $s_j = S_{j_1,\ldots,j_5}$. To formalize this we define operator O_p such that

$$O_p(s_i) = s_j \quad \Leftrightarrow \quad O_p(S_{i_1 \ldots i_5}) = S_{j_1 \ldots j_5} \tag{3.13}$$

if and only if there is a transition from s_i to s_j obtained by a small advance of the phase of the pth oscillator. The action of O_p is characterized by the subscript changes:

$$\text{if} \quad i_p = \mathrm{b}, \qquad \text{then} \quad \begin{cases} i_n = \mathrm{y} & \to \quad j_n = \mathrm{b}, \\ i_n = \mathrm{w} & \to \quad j_n = \mathrm{y}, \\ i_n = \mathrm{b} & \to \quad j_n = \mathrm{y} \quad \text{for} \quad n \neq p, \\ i_n = \mathrm{b} & \to \quad j_n = \mathrm{w} \quad \text{for} \quad n = p, \end{cases} \tag{3.14}$$

$$\text{if} \quad i_p \in \{\mathrm{y}, \mathrm{w}\}, \quad \text{then} \qquad j_n = i_n \quad \text{for all} \quad n.$$

According to (3.14) there are two "qualitatively" different choices of $p \in \{1, \ldots, 5\}$. Choosing an oscillator from the unstable blue cluster ($i_p = \mathrm{b}$) gives $O_p(s_i) = s_j$ with $j \neq i$, while choosing an oscillator from the stable yellow cluster ($i_p = \mathrm{y}$) or choosing the single white oscillator ($i_p = \mathrm{w}$) gives $O_p(s_i) = s_i$. Formula (3.14) evaluated for all s_i is presented as Table 2 in Appendix B. According to this, one can follow arbitrary paths on the directed graph in Figure 6(b) by making a sequence of arbitrarily small perturbations to the appropriate oscillators. We will exploit this feature of the network in sections 3.3 and 4.

We can define an adjacency matrix A with elements $A_{ij} \in \{0, 1\}$, where $A_{ij} = 1$ if and only if there is a directed edge from node s_i to s_j in the graph Figure 6(b). The number

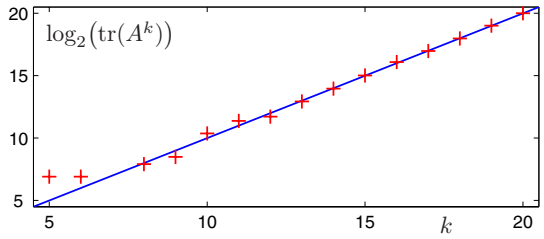


Figure 7. *The logarithm of the number of distinct cyclic paths* $\log_2(\mathrm{tr}(A^k))$ *as a function of the path length* k *(red crosses) for the graph shown in Figure* 6(b). *Note that the shortest cycle length is* 5 *and there are no cycles of length* 7. *The blue line corresponds to the growth* 2^k.

of cyclic paths of any given length can be calculated from the adjacency matrix as follows. The elements of the matrix A^k give the number of distinct paths of length k from node s_i to node s_j; see [15]. Thus, the diagonal elements of A^k give the number of distinct cyclic paths of length k from a node back to itself, and the trace $\mathrm{tr}(A^k)$ represents the total number of cyclic paths of length k in the network. In Figure 7 the logarithm of $\mathrm{tr}(A^k)$ is plotted as a function of k (red crosses). One can observe that there are no cycles of length $1, 2, 3, 4, 7$ and that $\mathrm{tr}(A^k) \approx 2^k$ for large k (the blue line represents the exact relation $\mathrm{tr}(A^k) = 2^k$). We remark that cycles with path length $k = 6$ become relevant when the natural frequencies of the oscillators are detuned in (2.1) according to (2.5) with $\delta \neq 0$, as will be discussed in sections 3.3 and 4.2.

3.3. Dynamics of the attracting heteroclinic network. For parameters (3.1) in the unperturbed case $\epsilon = \eta = \delta = 0$ the heteroclinic network described above is asymptotically stable/attracting (even though each three-cluster state is saddle). (Recall that the contracting eigenvalue is stronger than the expanding one, i.e., $|\lambda_2| > \lambda_3$ as given by (3.8) and plotted in Figure 4(a); so by [18] the network is asymptotically stable.) This attractivity is essential if one wishes to perform reliable computations by using the heteroclinic network: small perturbations do not drive the system away from the network as time progresses, allowing navigation along a chosen path on the network.

In Figure 8 the attractivity is demonstrated for an arbitrary initial condition. Figure 8(a) shows the dynamics in phase space. The trajectory approaches the network such that it approaches a sequence of three-cluster states (labeled red dots) and "travels close" to the heteroclinic connections from one state to the other while spending longer and longer periods of time close to the three-cluster states. This well-known slowing down of cycling [18] is seen in time profiles in Figure 8(b), where the plateaux correspond to "being close to" the three-cluster states and the quick transitions between them correspond to "traveling close to" the heteroclinic connections. In this panel a color code is also applied: yellow oscillators form the stable cluster, blue oscillators form the unstable cluster, and labels on the top identify the three-cluster states according to Table 1. Regions of quick transitions are colored light green. The length of time intervals spent close to the three-cluster states, i.e., the length of the plateaux, increases as

$$\lim_{k\to\infty} \frac{T_k}{T_{k-1}} = \frac{|\lambda_2|}{\lambda_3}, \tag{3.15}$$

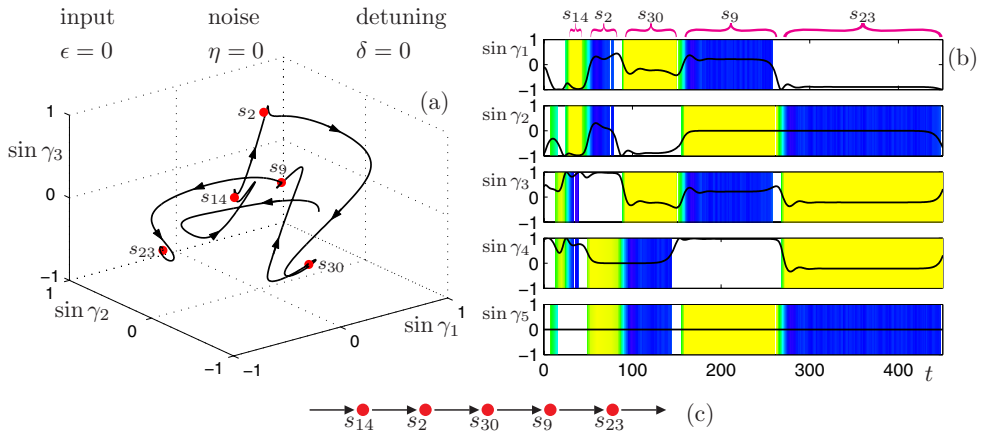


Figure 8. *Approaching the heteroclinic network from an arbitrary initial condition in the unperturbed case (as labeled on the top left) for parameters* (3.1). *Panel* (a) *shows the dynamics in phase space: The labeled red dots show the three-cluster states approached by the trajectory that "travels close to" the heteroclinic connections. In panel* (b) *the time profiles are displayed: Plateaux of* $\sin\gamma_n$ *correspond to three-cluster states and transitions between them correspond to heteroclinic connections. Lining up the colored regions vertically reveals which oscillators form the stable yellow and unstable blue clusters, and the labels on the top identify the three-cluster states according to Table* 1. *Transition regions are colored light green. Panel* (c) *shows the path followed on the graph in Figure* 6(b). *See also the accompanying animation* (68396_01.avi [3.9MB]).

which corresponds to the strength of attraction of the heteroclinic connections [18]. Notice in Figure 8(b) that the limit (3.15) is almost achieved even for small values $k = 3, 4, 5$. Figure 8(c) shows the path followed on the graph in Figure 6(b) by the above dynamics.

The heteroclinic behavior changes when external perturbations such as noise ($\eta \neq 0$), detuning ($\delta \neq 0$), or impulsive inputs ($\epsilon \neq 0$) are applied; see (2.1)–(2.5). Figure 9 shows the response of the system to such perturbations; i.e., the effects of noise, detuning, and periodic impulsive inputs are shown in panels (a), (b), and (c), respectively, where exactly the same notation is used as in Figure 8. The precise dynamics followed depends on the initial condition, although after an initial transient the dynamics enters a neighborhood of the heteroclinic attractor. To avoid transient effects the initial condition is set close to the three-cluster state $s_1 = S_{\mathrm{yywbb}}$. We assume that ϵ, η, and δ are small enough so that the only attractors of the system are in a neighborhood of the heteroclinic attractor for the unperturbed case $\epsilon = \eta = \delta = 0$. When the system is close to a three-cluster state it is sensitive to perturbations; that is, the perturbations determine which oscillator in the unstable blue cluster is advanced with respect to the other and, consequently, which of the two outgoing heteroclinic connections should be followed. The asymptotic stability of the network ensures that, after a transient, trajectories remain close to the heteroclinic network. This means that we can describe the dynamics by the itinerary of the heteroclinic network in Figure 6(a) (or the corresponding graph in Figure 6(b)). According to this we use the expressions "at a three-cluster state" and "at a node" to mean that the system is close to a three-cluster state. Furthermore, for the perturbed system the formula (3.15) does not hold. Instead, a characteristic time of transition between states can be assigned as shown by the length of plateaux in Figure 9(a2), (b2), (c2). This time is inversely proportional to the logarithm of

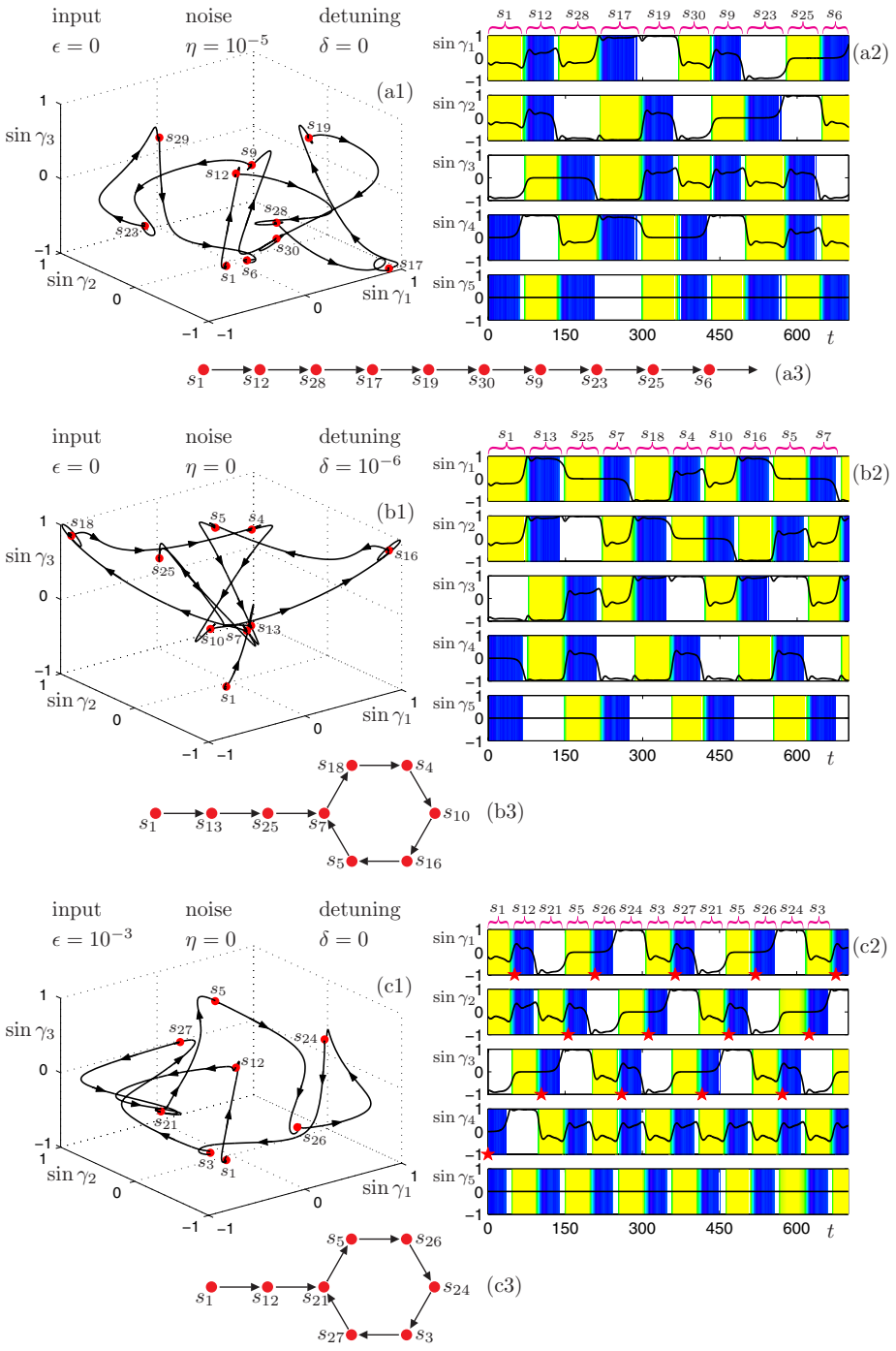


Figure 9. *Dynamics of the system in the presence of noise* (a), *detuning* (b), *and periodic impulsive inputs* (c) *for parameters* (3.1). *Panels* (a1), (b1), *and* (c1) *show the dynamics in phase space; panels* (a2), (b2), *and* (c2) *depict the time profiles; and panels* (a3), (b3), *and* (c3) *show the paths followed in Figure* 6(b). *The same notation is used as in Figure* 8. *Panel* (a2) *corresponds to the raster plot in Figure* 1. *In panel* (c2) *red stars in every* $t = kT = k\,52.0$ *indicate which oscillator receives the impulsive input. See also the accompanying animations* (68396_02.avi [6.4MB]), (68396_03.avi [6.4MB]), *and* (68396_04.avi [6.7MB]).

the magnitude of perturbations; see section 4.

In Figure 9(a) noise of strength $\eta = 10^{-5}$ forces the system to follow a randomly chosen path on the heteroclinic network. At each three-cluster state/node the noise randomly chooses which of the two oscillators in the unstable blue cluster is advanced with respect to the other and, consequently, which of the two outgoing heteroclinic connections/directed edges is followed. Note that the whole network is explored with probability 1 if the simulation is run for long enough. The time domain $t \in [0, 350]$ in Figure 9(a2) corresponds to the raster plot in Figure 1, where vertical lines indicate when the individual phases $\theta_n(t)$ go through zero.

Figure 9(b) shows the dynamics when oscillators are detuned according to the uniform detuning (2.5), which for $N = 5$ means

$$\omega_1 = \omega - 2\delta, \quad \omega_2 = \omega - \delta, \quad \omega_3 = \omega, \quad \omega_4 = \omega + \delta, \quad \omega_5 = \omega + 2\delta. \tag{3.16}$$

Here the detuning magnitude $\delta = 10^{-6}$ is considered. Since the heteroclinic attractor is only robust to perturbations that preserve the symmetries of the system, it is generically destroyed by this detuning. More precisely, at a three-cluster state/node the detuning (3.16) always advances the blue oscillator with larger index (the one which changes its color to white in Figure 9(b2)). Consequently, one particular outgoing heteroclinic connection/directed edge is followed. This results (after a transient path) in the system reaching a cyclic path of length 6 (which is followed as $t \to \infty$). This path corresponds to limit cycle oscillations of the phase differences in the detuned system. Note that this limit cycle is not unique; i.e., for different initial states we can see different attractors, as discussed in section 4.2.

In Figure 9(c) the system is forced by periodic impulsive inputs of the form

$$I_i(t) = \sum_{k=0}^{\infty} \delta(t - k\,T)\,\delta_{ip_k}, \tag{3.17}$$

where $\delta(t - k\,T)$ is a Dirac delta function while δ_{ip_k} is a Kronecker delta. The kth term of this sum represents a unit impulse applied to the p_kth oscillator at time $k\,T$, where $p_k \in \{0, 1, \ldots, 5\}$ ($p_k = 0$ corresponds to no input). That is, the phase of the oscillator receiving the input is simply advanced by the input amplitude ϵ. Here the input amplitude $\epsilon = 10^{-3}$ and period $T = 52.0$ are considered, and a red star is plotted in Figure 9(c2) when an oscillator receives an input. At a three-cluster state/node $s_i = S_{i_1,\ldots,i_5}$ at time $t = k\,T$ we can choose which of the two oscillators in the unstable blue cluster should receive the input (that is, $i_{p_k} = \mathrm{b}$ can be chosen) and, consequently, which of the two outgoing heteroclinic connections/directed edges should be followed. In Figure 9(c2) we always choose the blue oscillator with the smaller index to have its phase advanced (the blue oscillator with the red star changes its color to white). This results (after following the transient path) in a cyclic path of length 6 (which is followed as $t \to \infty$). Note that different input configurations lead to different paths and that repeated input sequences correspond to cyclic paths. This shows that by exploiting the natural dynamics of the system we can drive it along a chosen path on the heteroclinic network: we can think of the dynamics as an input-output system which is capable of performing finite-state computation [3]. The reliability of such computations is investigated in detail in the next section.

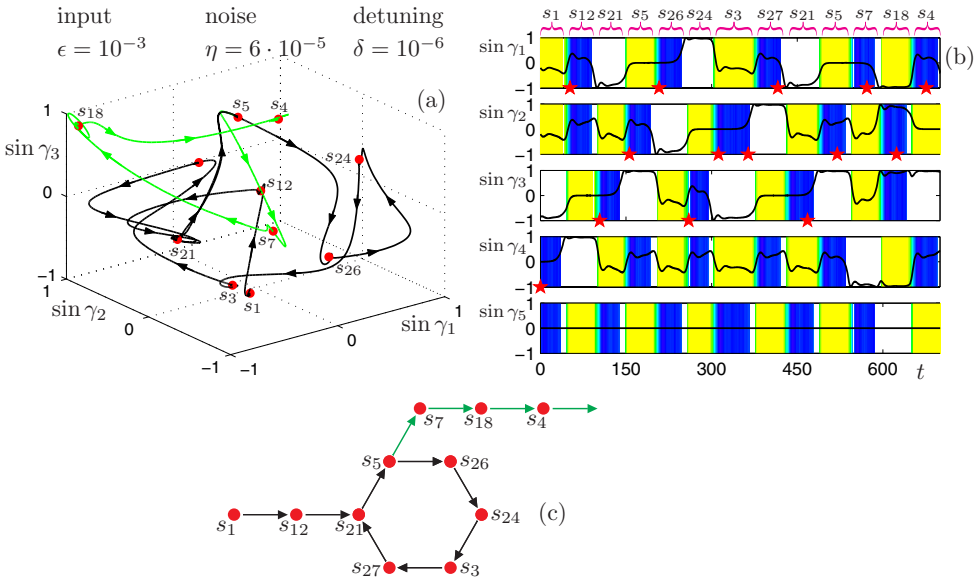


Figure 10. *Dynamics of the system in the presence of periodic impulsive inputs, noise, and detuning (as labeled on the top left) for parameters* (3.1). *The same notation is used as in Figure* 8. *In panel* (b) *red stars in every* $t = kT = k\,52.0$ *indicate which oscillator receives the impulsive input. In panels* (a) *and* (c) *the light green sections of the trajectory and the graph show the dynamics after an incorrect switch at the three-cluster state* $s_5 = S_{\mathrm{ybwby}}$.

4. Reliability of switching along the heteroclinic network. In this section we illustrate that it is possible to reliably drive the system around the heteroclinic network of three-cluster states using small impulsive inputs. This can be managed even when noise and detuning are present if the input amplitude is large enough relative to the noise strength and detuning magnitude. From now on we will use the term *switch* to mean a transition along a heteroclinic orbit between three-cluster states.

For example, one may check that by applying the same noise as in Figure 9(a) (with strength $\eta = 10^{-5}$), the same detuning as in Figure 9(b) (with magnitude $\delta = 10^{-6}$), and the same input as in Figure 9(c) (with amplitude $\epsilon = 10^{-3}$ and period $T = 52.0$), the obtained trajectory, time profiles, and, consequently, the path followed are indistinguishable from those in Figure 9(c), where neither noise nor detuning are present. This means that the dynamics is still determined by the sequence of inputs.

Figure 10 shows the response of the system to perturbations with noise increased to the extent that the switches are not all determined by the inputs. (The same notation is used as in Figure 8 and the initial condition is at the three-cluster state $s_1 = S_{\mathrm{yywbb}}$.) Again the same detuning is applied as in Figure 9(b) (with magnitude $\delta = 10^{-6}$), the same input is applied as in Figure 9(c) (with amplitude $\epsilon = 10^{-3}$ and period $T = 52.0$), but the noise (of strength $\eta = 6 \cdot 10^{-5}$) is larger than in Figure 9(a). The trajectory and the time profiles are initially close to those in Figure 9(a) and (b).

Figure 10(b) shows that the first "incorrect" switch happens at time $t = 6\,T = 312.0$, where the first blue oscillator receives an impulsive input but the system stays at the three-cluster state $s_3 = S_{\mathrm{ybwyb}}$ instead of switching to $s_{27} = S_{\mathrm{bwyby}}$. Another impulsive input has

to be applied to the same oscillator at time $t = 7T = 364.0$ in order to force the system to switch. The next "incorrect" switch takes place at time $t = 10T = 520.0$, where an impulsive input is applied to the first blue oscillator of the three-cluster state $s_5 = S_{\text{ybwby}}$; that is, a switch to $s_{26} = S_{\text{ybwby}}$ is initiated. Instead, a switch to $s_7 = S_{\text{byywb}}$ is observed (the blue oscillator without a red star changes its color to white; i.e., the "wrong" heteroclinic connection/directed edge is followed). This "incorrect" switch leads the system away from the cyclic path of length 6, as shown by the green part of the trajectory in Figure 10(a) and by the green part of the graph in Figure 10(c).

In order to measure the accuracy of switching, at each three-cluster state we randomly choose which blue oscillator receives the impulsive input (i.e., choose which heteroclinic connection/directed edge should be followed), and then detect whether the expected switch happens. To be more precise, we classify the system as being at a three-cluster state s_i if $\theta(t) \in B_\xi(s_i)$, where $B_\xi(s_i)$ denotes the ξ-neighborhood of s_i. The *detection tolerance* $\xi > 0$ is chosen such that the ξ-neighborhoods of different three-cluster states do not intersect each other, that is, $B_\xi(s_i) \cap B_\xi(s_j) = \emptyset$ if $i \neq j$. In this article we fix $\xi = 0.5$, but no qualitative change is observed when using different ξ.

Suppose that we are at the three-cluster state $s_i = S_{i_1,\ldots,i_5}$ at time $t = kT$ and randomly choose an oscillator p_k in the unstable blue cluster for input (that is, $i_{p_k} = \text{b}$). Then we predict $s_j = O_{p_k}(s_i)$ to be the state we expect after a time T; see definitions (3.13), (3.14) and Table 2. We declare a switch to be *correct* if $\theta(t + T)$ is at the expected three-cluster state s_j and otherwise *incorrect*. An incorrect switch can occur for one of the three reasons:

(i) we cannot classify the state of $\theta(t)$ (i.e., there is no i such that $\theta(t) \in B_\xi(s_i)$),
(ii) we cannot classify the state of $\theta(t + T)$, or
(iii) we can classify both states but the switch is not as expected (i.e., $O_{p_k}(s_i) \neq s_j$), as can be observed in Figure 10(b).

Figure 11 shows 14 randomly chosen switches for perturbations $\epsilon = 10^{-3}$, $\eta = 10^{-7}$, and $\delta = 0$ with $T = 85.0$. (Again the same notation is used as in Figure 8 and the initial condition is the three-cluster state $s_1 = S_{\text{yywbb}}$.) The system follows the randomly chosen path expected from the provided sequence of impulsive inputs; that is, all switches are correct (the blue oscillators with red stars change their color to white in Figure 11(a), and the sequence of three-cluster states can be predicted by $s_j = O_{p_k}(s_i)$). In the next section we compute the rate of incorrect switches as a function of input period T, input amplitude ϵ, noise strength η, and detuning magnitude δ.

4.1. Error rates for switching. In order to check how reliable the switches are when driving the system with impulsive inputs in the presence of noise and detuning, we perform long simulations and count the proportion of correct and incorrect switches. The *error rate* $0 \leq E \leq 1$ is calculated as the proportion of incorrect switches for a large number of attempts (2000 switches) for a random choice of the possible switches at each state. The random choice is uniformly chosen between the two possible inputs that initiate switching at each step. More precisely, if we are at the three-cluster state $s_i = S_{i_1,\ldots,i_5}$, we choose with equal probability one of the two p_k such that $i_{p_k} = \text{b}$; that is, an oscillator in the unstable blue cluster is perturbed. If no three cluster-state s_i is identified we choose $p_k = 0$; i.e., no input is provided. The error rate clearly depends on the exact values of the parameters in (2.2), which here are fixed as in (3.1). It also depends on the input period T, the input amplitude ϵ, the noise strength η,

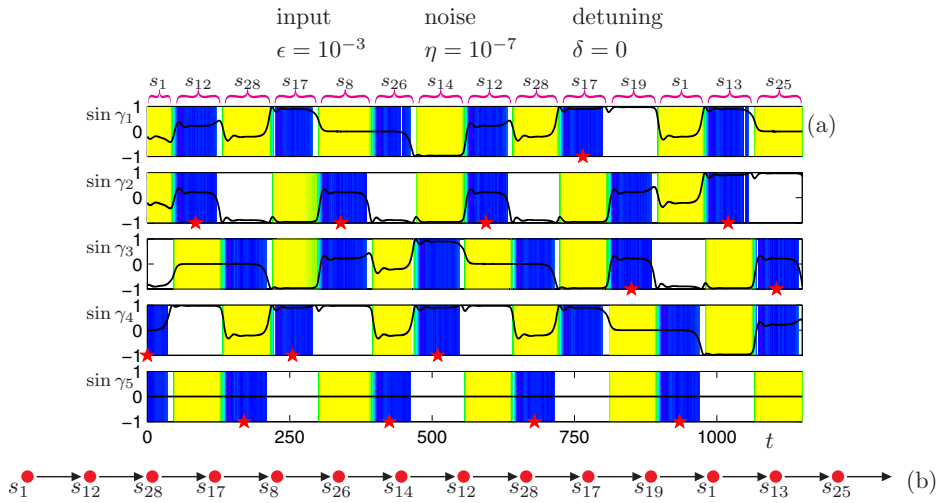


Figure 11. *Driving the system along a randomly chosen path in the heteroclinic network for parameters* (3.1) *with perturbations labeled on the top. The same notation is used as in Figure* 8. *In panels* (a) *and* (b) *the time profiles and the followed path are shown for the* 14 *switches. In panel* (a) *red stars in every* $t = kT = k\,85.0$ *indicate which oscillator receives the impulsive input. See also the accompanying animation* (68396_05.avi [10.2MB]).

and the detuning magnitude δ; that is, $E = E(T, \epsilon, \eta, \delta)$. We will assume that the error rate converges for our assumptions on random choice of $\{p_k\}$.

One can maintain a periodic orbit near the heteroclinic network with period $T > T_\epsilon$ between successive states by application of impulsive inputs of amplitude ϵ, where

$$T_\epsilon = -\frac{1}{\lambda_e}\ln\epsilon + \mathcal{O}(1) = -\frac{\ln(10)}{\lambda_e}\log_{10}\epsilon + \mathcal{O}(1), \tag{4.1}$$

as $\epsilon \to 0$. Here λ_e is the expanding eigenvalue of the three-cluster states (for parameters (3.1) we have $\lambda_e = \lambda_3 = 0.1703$, as given by (3.8) and shown in Figure 4(a)). Scaling (4.1) can be seen by the following argument. Consider a transition from the three-cluster state s_i to the three-cluster state s_j with symmetrically related surfaces of section Σ_i and Σ_j, intersecting the stable manifolds of s_i and s_j, respectively. If we provide an impulse of amplitude ϵ in the unstable direction at Σ_i, this will grow proportionally to $\epsilon\, e^{\lambda_e t}$, meaning that it saturates and causes a switch when $\epsilon\, e^{\lambda_e T} = \mathcal{O}(1)$. This in turn implies (4.1). If we perturb a three-cluster state s_i in its unstable direction, then the unstable direction for s_j is not affected, meaning that by choosing to perturb on a section Σ_i closer to s_i, we can provide impulses of amplitude ϵ to obtain periodic orbits with any period $T > T_\epsilon$.

Similarly, it can be shown [24] that the addition of noise of strength η will transform the heteroclinic attractor into an approximately periodic attractor with period $T \approx T_\eta$, where

$$T_\eta = -\frac{1}{\lambda_e}\ln\eta + \mathcal{O}(1) = -\frac{\ln(10)}{\lambda_e}\log_{10}\eta + \mathcal{O}(1), \tag{4.2}$$

as $\eta \to 0$. In consequence, we conclude that one can navigate around the network with error

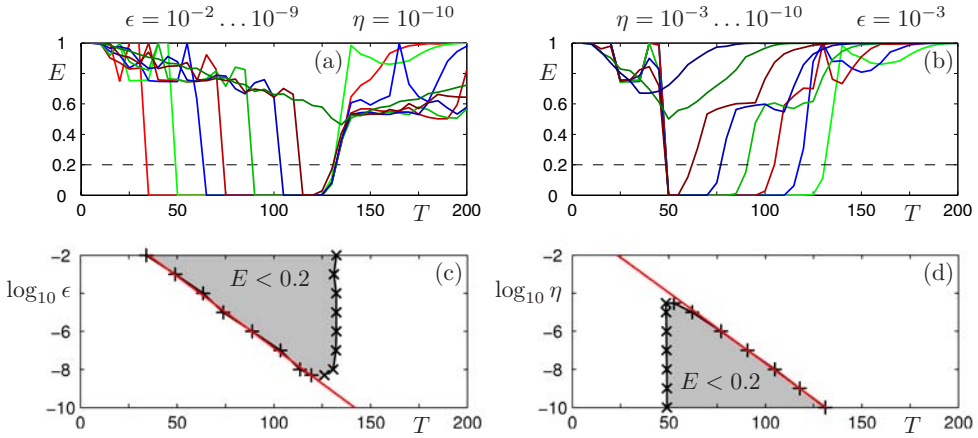


Figure 12. *In panels* (a) *and* (b) *the error rate* E *is shown as a function of the input period* T *for parameters* (3.1). *In panel* (a) *several different values of the input amplitude* ϵ *are considered for fixed noise strength* η, *while in panel* (b) *several different values of* η *are taken for fixed* ϵ. *No detuning is added* ($\delta = 0$). *The leftmost curves in panels* (a) *and* (b) *correspond to the largest values of* ϵ *and* η, *respectively. In panels* (c) *and* (d) *the shaded domains show where the error rate is below the threshold* 0.2 *in the* $(T, \log_{10} \epsilon)$*-plane and in the* $(T, \log_{10} \eta)$*-plane, respectively. Notice that the fitted red lines have approximately the same gradient in both panels* (c) *and* (d).

rate E close to zero as long as T is in the range

$$T_\epsilon < T < T_\eta. \tag{4.3}$$

If $T < T_\epsilon$, then errors take place because the time is not long enough to permit the connections to be followed. If $T > T_\eta$, then the noise will be the dominant influence in determining when transitions take place, and hence the error rate will be large. Note that a range of possible T in (4.3) is available as long as

$$\epsilon \gg \eta. \tag{4.4}$$

If both noise $\eta > 0$ and detuning $\delta > 0$ are present in the system, then the same argument holds except one must write $\max(\delta, \eta)$ in place of η in formulae (4.2)–(4.4).

In Figure 12(a) and (b) the error rates (obtained from 2000 switches) are shown as a function of the input period T for several different values of input amplitude ϵ and noise strength η, respectively (with no detuning, $\delta = 0$). The leftmost curves in Figure 12(a) and (b) correspond to the largest values of ϵ and η, respectively. These panels clearly show that when (4.4) is satisfied there exists an interval of T (4.3) where $E \approx 0$. Also notice that left and right boundaries of this interval move proportionally to $\log_{10} \epsilon$ in Figure 12(a) and proportionally to $\log_{10} \eta$ in Figure 12(b). In order to measure these boundaries more precisely we detect where the curves intersect the error rate $E = 0.2$ (denoted by horizontal dashed lines in Figure 12(a) and (b)). We plot the corresponding data in Figure 12(c) and (d), where the low error rate domains ($E < 0.2$) are shaded. The least square fitted red lines in Figure 12(c) and (d) with gradients

$$g_\epsilon = -0.0749 \pm 0.0014 \qquad \text{and} \qquad g_\eta = -0.0740 \pm 0.0005, \tag{4.5}$$

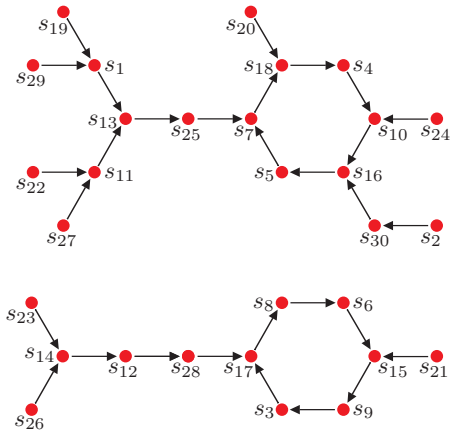


Figure 13. *The network in Figure* 6(b) *collapses to this network with two connected components on applying the uniform detuning* (3.16)*. Observe that there are two eventually periodic paths of length* 6*, indicating that detuning results (after a transient path) in one of two "periodic spatiotemporal codes" depending on initial state.*

respectively, agree well with the predicted gradients in (4.1) and (4.2), namely,

$$g = -\frac{\ln(10)}{\lambda_e} = -\frac{\ln(10)}{\lambda_3} = -0.0740. \tag{4.6}$$

This shows that the accuracy of navigation around the heteroclinic network can be predicted by the stability properties of the three-cluster states.

4.2. Detuning-driven spatiotemporal coding. In addition to giving a network that can perform finite-state computations, the system (2.1), (2.2) can give spatiotemporal coding of steady external inputs that are in the form of detuning. Assuming the uniform detuning (3.16) and setting the initial condition at $s_1 = S_{\text{yywbb}}$ the system reaches a cyclic path of length 6 (after some transient path) as shown in Figure 9(b). In fact, this path encodes information about the form of the detuning, as discussed below.

Here we assume the uniform detuning (3.16) and consider the system starting at different three-cluster states. Whenever the system is at a three-cluster state $s_i = S_{i_1,\dots,i_5}$ there will be a "preferred" direction along the graph in Figure 6(b). Namely, if $i_p = i_q = \text{b}$ are the unstable blue oscillators, then if $\omega_{i_p} > \omega_{i_q}$ the detuning will tend to initiate a switch to $O_{i_p}(s_i)$, while if $\omega_{i_p} < \omega_{i_q}$ the initiated transition will be to $O_{i_q}(s_i)$; see definitions (3.13), (3.14) and Table 2. Hence, uniform detuning causes a bias to the switchings, meaning that one of the outgoing edges is removed from each node and the network with the graph in Figure 13 remains for the original network. This "detuned" graph consists of two connected components, each of which has an absorbing cyclic path of length 6 and transient paths that are asymptotic to these circuits after a finite number of switches. Which of the cyclic paths is reached depends on the initial state. The sequences of three-cluster states along the circuits can therefore be considered as spatiotemporal encodings of the applied detuning.

Notice that the magnitude of δ does not change the graph in Figure 13, but it does affect the average period between switches T_δ in a similar way to (4.1). Different detuning configurations obtained by permutation of the indices in (3.16) result in different cyclic paths

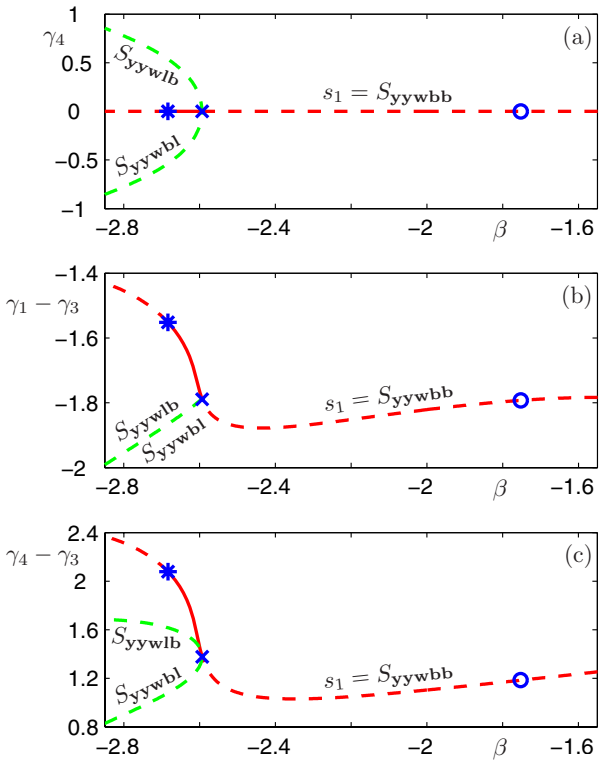


Figure 14. *Bifurcation diagrams showing the phase differences* $\gamma_4 = \theta_5 - \theta_4$ (a), $\gamma_1 - \gamma_3 = \theta_1 - \theta_3$ (b), *and* $\gamma_4 - \gamma_3 = \theta_4 - \theta_3$ (c) *as the parameter* β *is varied for* $r = 0.2$, $\alpha = 1.8$, $\omega = 1$ *with* $\epsilon = \eta = \delta = 0$. *The three-cluster state* $s_1 = S_{\mathrm{yywbb}}$ *and the bifurcated cluster states* S_{yywlb}, S_{yywbl} *are labeled on the corresponding red and light green branches. Solid curves denote stable states and dashed curves denote saddle states. The blue star at* $\beta = -2.6840$, *blue cross at* $\beta = -2.5933$, *and blue circle at* $\beta = -1.7515$ *denote the Hopf, pitchfork, and resonance bifurcations, respectively. The network shown in Figure 6 is attracting between the pitchfork and the resonance bifurcations.*

of length 6, that is, different spatiotemporal codes. One may check that 120 cyclic paths of length 6 exist (using the adjacency matrix defined in section 3.2 we obtain $\mathrm{tr}(A^6) = 120$). However, only 20 of these paths are possible absorbing cycles when one considers all possible $5! = 120$ permutations of (3.16).

5. Excitable dynamics. The dynamics described in sections 3 and 4 is robust (structurally stable); i.e., there is an open set of parameters around (3.1) where such an attracting heteroclinic network exists. We observe the network on varying β in the range $(-2.5933, -1.7515)$ and now analyze the bifurcations at the boundaries of this regime. At the lower boundary the saddles within the network are stabilized at a pitchfork bifurcation to become stable three-cluster states that remain arranged in a network as in Figure 6. At the upper boundary the heteroclinic network persists but loses stability at a resonance bifurcation [9].

As verification of this, Figure 14 shows bifurcation diagrams for the three-cluster state $s_1 = S_{\mathrm{yywbb}}$. The parameters r, α, and ω are given by (3.1) except that β is varied, and we set $\epsilon = \eta = \delta = 0$. The figure was generated by the continuation package AUTO [10].

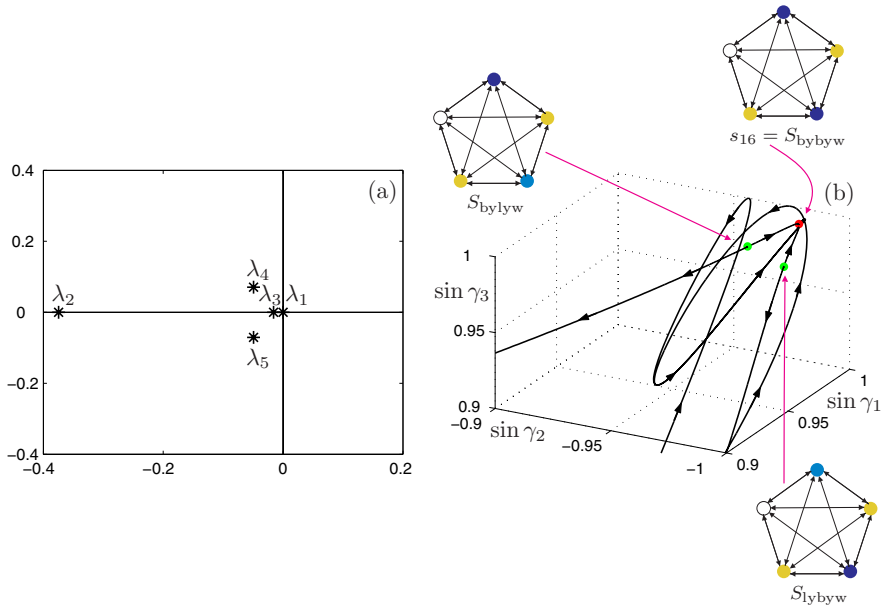


Figure 15. *Dynamics around a stable three-cluster state for parameters* (5.1) *in the unperturbed case* $\epsilon = \eta = \delta = 0$. *In panel* (a) *the eigenvalues* (3.7), (5.3) *are displayed in the complex plane. Panel* (b) *shows the local dynamics in phase space in the vicinity of the three-cluster state* $s_{16} = S_{\text{bybyw}}$ *(red dot). Light green dots show the bifurcated saddle cluster states* S_{lybyw} *and* S_{bylyw}. *One branch of the unstable manifold of each of these states connects to* $s_{16} = S_{\text{bybyw}}$.

In Figure 14 the phase differences are plotted as a function of the parameter β. Solid lines correspond to stable states, and dashed lines correspond to saddle states. Blue star, blue cross, and blue circle represent Hopf, pitchfork, and resonance bifurcations, respectively, from the three-cluster state. Between the pitchfork and the resonance bifurcation the heteroclinic network is attracting because $|\lambda_2| > \lambda_3$ (as in Figure 4(a)); that is, the contracting eigenvalue is stronger than the expanding one for the robust connection. The pitchfork bifurcation occurs when $\lambda_3 = 0$ ($\beta = -2.5933$), while the resonance occurs when $|\lambda_2| = \lambda_3$ ($\beta = -1.7515$).

Figure 14(a) shows the phase difference $\gamma_4 = \theta_5 - \theta_4$ as a function of β. For the saddle three-cluster state $s_1 = S_{\text{yywbb}}$ (horizontal red line) the phase difference $\gamma_4 = \theta_4 - \theta_5$ is zero, indicating that the 4th and the 5th oscillators are synchronized. At the pitchfork bifurcation (blue cross at $\beta = -2.5933$) the three-cluster state $s_1 = S_{\text{yywbb}}$ becomes stable with $\text{Re}(\lambda_n) < 0$ for $n = 2, 3, 4, 5$ (see Figure 15(a)). The bifurcation creates two branches of saddle cluster states (light green curves) where $\gamma_4 \neq 0$, i.e., the 4th and the 5th oscillators are not synchronized. If $\gamma_4 > 0$, for this cluster state we write S_{yywlb} (and S_{yywbl} if $\gamma_4 < 0$). The subscript "l" stands for the light blue color used in Figures 15–16 and indicates the oscillator in the unstable blue cluster whose phase has been advanced. These new cluster states are located in the subspace (3.11), and their unstable manifolds are connected to the three-cluster states $s_1 = S_{\text{yywbb}}$, $s_{12} = S_{\text{bbywy}}$, and $s_{13} = S_{\text{bbyyw}}$, as detailed further below.

Figure 14(b) and (c) show the phase differences $\gamma_1 - \gamma_3 = \theta_1 - \theta_3$ and $\gamma_4 - \gamma_3 = \theta_4 - \theta_3$ as a function of β for the relevant cluster states. For $s_1 = S_{\text{yywbb}}$ (red curve) we have $\gamma_1 - \gamma_3 = \chi$ and $\gamma_4 - \gamma_3 = \psi$; see definitions (3.3), (3.4). Notice that in Figure 14(b) the branches of

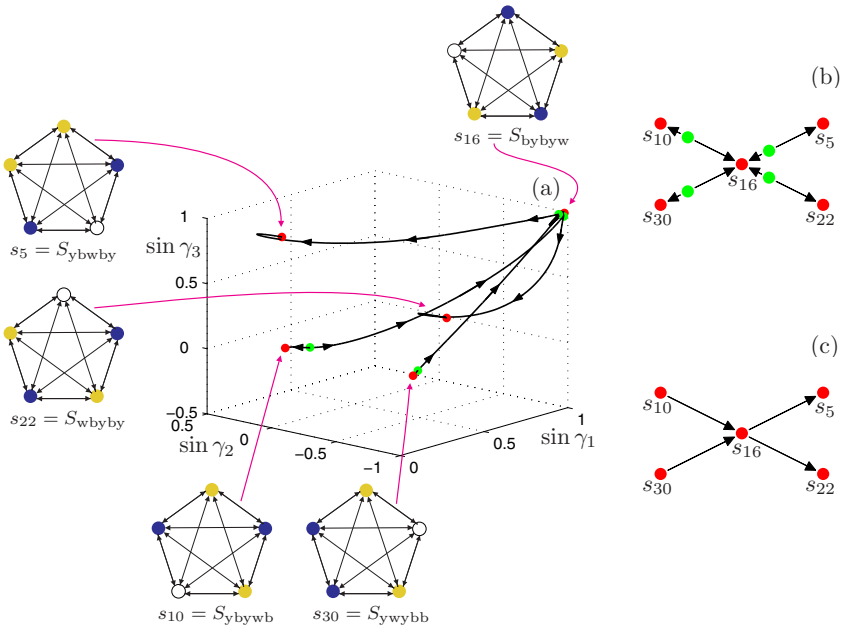


Figure 16. *In panel* (a) *the heteroclinic connections originated in and terminated at the three-cluster state* $s_{16} = S_{\text{bybyw}}$ *are shown in phase space for parameters* (5.1) *in the unperturbed case* $\epsilon = \eta = \delta = 0$. *The cluster states are represented by red and light green dots. In panel* (b) *a simple graph representation is displayed such that the nodes and directed edges correspond to cluster states and heteroclinic connections, respectively. Panel* (c) *shows a further simplified graph representation where only the three-cluster states are displayed and the directed edges represent possible large-input-triggered switches between them.*

cluster states S_{yywlb} and S_{yywbl} overlap (light green curve), which indeed shows that these cluster states are embedded in the subspace (3.11). The symmetric copies of $s_1 = S_{\text{yywbb}}$, as in Table 1, behave in a corresponding way at the pitchfork bifurcation.

On reducing β further, the complex conjugate pair of eigenvalues moves to the right-hand side of the imaginary axis at the Hopf bifurcation (blue star at $\beta = -2.6840$), and the three-cluster state $s_1 = S_{\text{yywbb}}$ becomes a saddle again since $\lambda_2, \lambda_3 < 0$ and $\operatorname{Re}\lambda_4 = \operatorname{Re}\lambda_5 > 0$. However, the resulting dynamics is different from that on the right side of the pitchfork bifurcation (the oscillatory branches emerging at the Hopf bifurcation are not studied further in this paper).

At the resonance bifurcation (blue circle at $\beta = -1.7515$) the contracting eigenvalue becomes weaker than the expanding eigenvalue, i.e., $|\lambda_2| < \lambda_3$. Consequently, the heteroclinic network becomes repelling; see [9]. Note that the three-cluster states persist as saddles and the network still persists on the right side of the resonance bifurcation. However, since it is not attracting, general initial conditions may approach long-period stable limit cycles or even chaotic attractors (see section 7).

5.1. Network of excitable three-cluster states. We now investigate in detail the network of stable excitable states that replaces the heteroclinic network to the left side of the pitchfork bifurcation, by considering the parameters

$$r = 0.2, \quad \alpha = 1.8, \quad \beta = -2.6, \quad \omega = 1.0. \tag{5.1}$$

Note that only β has been changed compared to parameters (3.1). Solving (3.5) for parameters (5.1), one can find a three-cluster state (3.3), (3.4) with

$$\chi = -1.7638, \quad \psi = 1.4594, \quad \Omega = 0.9343, \tag{5.2}$$

and symmetric copies as in Table 1. Substituting (5.1), (5.2) into (3.7), the eigenvalues are

$$\lambda_2 = -0.3746, \quad \lambda_3 = -0.0158, \quad \lambda_{4,5} = -0.0493 \pm \mathrm{i}\,0.0709, \tag{5.3}$$

as displayed in Figure 15(a). Notice that all eigenvalues are to the left of the imaginary axis indicating that the three-cluster states are stable, but λ_3 is close to zero as the system is close to the bifurcation point.

In Figure 15(b) the stable three-cluster state $s_{16} = S_{\mathrm{bybyw}}$ (red dot) and the bifurcated saddle cluster states S_{lybyw} and S_{bylyw} (light green dots) are displayed (with the oscillators corresponding to the subscript "l" colored light blue). The latter cluster states are located in the symmetric copy of the subspace (3.11). Observe that now $s_{16} = S_{\mathrm{bybyw}}$ is attracting from all directions. In particular, $\mathrm{Re}(v_4) = \mathrm{Re}(v_5)$ and $\mathrm{Im}(v_4) = -\mathrm{Im}(v_5)$ correspond to the spiralling directions, while v_3 and $-v_3$ are the directions toward the three-cluster states S_{lybyw} and S_{bylyw}. (The latter two directions are indeed the opposite of each other, but the curvature of the projection $(\sin\gamma_1, \sin\gamma_2, \sin\gamma_3)$ is large in the vicinity of the three-cluster state $s_{16} = S_{\mathrm{bybyw}}$ as it is close to $(1, -1, 1)$.)

It can be verified numerically that one branch of the one-dimensional unstable manifold of the cluster state S_{lybyw} returns to $s_{16} = S_{\mathrm{bybyw}}$ while the other branch connects to the three-cluster state $s_{22} = S_{\mathrm{wbyby}}$. Similarly, the unstable manifold of S_{bylyw} connects to $s_{16} = S_{\mathrm{bybyw}}$ and $s_5 = S_{\mathrm{ybwby}}$; this is illustrated in Figure 16(a). There are saddle cluster states S_{ybywl} and S_{ywybl} near $s_{10} = S_{\mathrm{ybywb}}$ and $s_{30} = S_{\mathrm{ywybb}}$, respectively. These are located in the symmetric copy of the subspace (3.12) and their unstable manifolds connect to $s_{16} = S_{\mathrm{bybyw}}$. This means that there are connections as shown in the graph Figure 16(b), where again the dots represent cluster states and directed edges represent heteroclinic connections between them.

Hence, by providing a large enough perturbation when the system is near $s_{16} = S_{\mathrm{bybyw}}$ in the direction of S_{lybyw}, it is possible to "jump over" the stable manifold of S_{lybyw} and then approach $s_{22} = S_{\mathrm{wbyby}}$. We say that the stable *three-cluster states are excitable* to large perturbations that separate the blue cluster. Taking these large perturbations into account we can obtain the simplified graph of connections in Figure 16(c). Notice that this graph is close to the one in Figure 5(b) for the original heteroclinic network with the difference that in Figure 5(b) infinitesimally small perturbations can trigger transitions while in Figure 16(c) the perturbations have to be of finite size.

Considering all three-cluster states the graph in Figure 6(b) is obtained again as a representation of the system when finite size perturbations are applied. We stress that in the resulting network of excitable three-cluster states, the excitability is a property of the three-cluster states and not a property of the individual oscillators. We will use the term *excitable network* to describe this dynamics.

5.2. Reliable switching in the excitable network. We now demonstrate that it is possible to reliably navigate around the graph of three-cluster states by applying finite size impulsive inputs in precisely the same way as described in section 4. Figure 17(a) shows an example

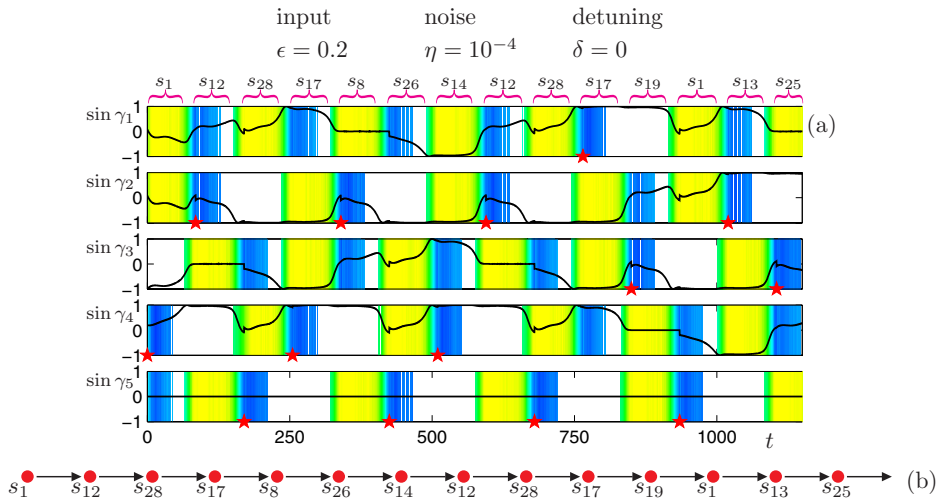


Figure 17. *Driving the system along a randomly chosen path of the network of excitable three-cluster states for parameters* (5.1) *with perturbations labeled on the top. The same notation is used as in Figure* 8. *In panels* (a) *and* (b) *the time profiles and the followed path are shown for the first* 14 *switches. In panel* (a) *red stars in every* $t = kT = k\,85.0$ *indicate which oscillator receives the input. Notice that on providing the same sequence of inputs as in Figure* 11 *the same path is followed. See also the accompanying animation* (68396_06.avi [11.0MB]).

where 14 correct switches are performed for parameters (5.1) and perturbations $\epsilon = 0.2$, $\eta = 10^{-4}$, $\delta = 0$ with $T = 85.0$. (Again the same notation is used as in Figure 8 and the initial condition is at the three-cluster state $s_1 = S_{\mathrm{yywbb}}$.) Comparing Figure 17 to Figure 11 the same sequence of impulsive inputs are provided (red stars in panel (a)) and the system follows the same randomly chosen path (panel (b)). Even though relatively high noise is applied in Figure 17, the switching pattern is reliably followed.

We test the reliability of switching by following random paths of the network when impulses of amplitude ϵ are provided with period T as in section 4.1. The error rate E is obtained (from 2000 switches again) when varying the input period T for several different values of noise strength η and fixed input amplitude ϵ. When the input amplitude ϵ is above some threshold, i.e., $\epsilon > \epsilon_c$ ($\epsilon_c \approx 0.18$ for parameters (5.1)), then there exist ranges of input period where reliable switching is possible as shown in Figure 18 by regions where $E \approx 0$. If the input amplitude is below the threshold, i.e., $\epsilon < \epsilon_c$, then no switches can be triggered by the input so the error rate $E(T, \epsilon, \eta, \delta) \equiv 1$.

Figure 18 shows that for noise strength $\eta \ll \epsilon$ the error rate $E \approx 0$ is achievable when T exceeds some critical period. That is, reliable switching is possible when a sufficiently long time is provided to travel along the connections described above. However, there also exists a window of input period in Figure 18 which corresponds to the case when the system gets close to a symmetrical copy of the subspace (3.11) but it is not in close vicinity of the three-cluster state embedded in the subspace. There exist regions in the subspace (e.g., the region around $(-0.95, 0.95, -0.9)$ in Figure 15(b)) where the provided inputs effectively place the system "on the other side" of the stable manifold of the symmetric copy of the cluster state S_{yywlb} or S_{yywbl} and thus lead to correct switches.

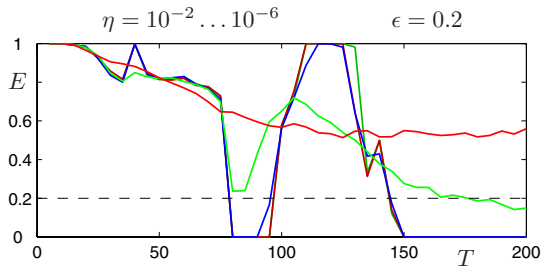


Figure 18. *The error rate E is shown as a function of the input period T for the excitable network with parameters (5.1) when several different values of the noise strength η are considered for fixed input amplitude ϵ and no detuning $\delta = 0$. Observe that the boundaries where the error rate rapidly drops do not move as the noise strength is varied (except for strong noise).*

The boundaries of these low error rate regions do not move much when the noise strength is varied except for very strong noise which increases the error rate for any T. This behavior can be explained by the stability of the three-cluster states and by the fact that ϵ is close to ϵ_c, that is, $\epsilon \gtrsim \epsilon_c$. Small noise $\eta \ll \epsilon_c$ does not provide enough perturbation to excite a switch between three-cluster states, while large enough noise $\eta > \epsilon_c$ violates the condition $\eta \ll \epsilon$; that is, reliable switching is not expected for any T.

6. Conclusions. We investigated a globally coupled phase oscillator system as a simple model for a highly connected neural system. We generalized the coupling considered in [2, 13, 17] and found an attracting robust heteroclinic network between partially synchronized three-cluster states of saddle type. A graph representation of the network was given that can be used to explain much of its dynamical behavior, and we studied its response to perturbations such as noise, frequency detuning, and impulsive inputs, each of which can drive the system around the heteroclinic network. One may interpret this dynamics as a form of *winnerless competition* [1, 6, 23] between three-cluster states since each of these states may be visited when tracking around the heteroclinic network but no state is an attractor for the system.

We found that the characteristic switching time between the three-cluster states is inversely proportional to the logarithm of the magnitude of perturbations. Broadly speaking, the bigger the perturbation, the faster the system "reacts" to it. We demonstrated that it is possible to reliably navigate the system along any chosen path on the heteroclinic network by providing small impulsive inputs to the appropriate oscillators, i.e., to perform finite-state computation as in [3]. Moreover, this can be done accurately in the presence of background noise and frequency detuning if the input amplitude dominates the noise strength and the detuning magnitude for a suitable range of input periods determined by these quantities. The boundaries of this range may be computed by considering the eigenvalues of the three-cluster states. In contrast, a steady input to the system (i.e., frequency detuning) resulted in a reduced network with attracting cyclic paths corresponding to a spatiotemporal output of the system.

We illustrated that on changing parameters in the coupling one can stabilize the three-cluster states at a pitchfork bifurcation. This results in a robust network of excitable states with the same graph structure as the heteroclinic network. Hence, it is still possible to navigate around this excitable network by applying finite size impulsive inputs to the appropriate oscillators.

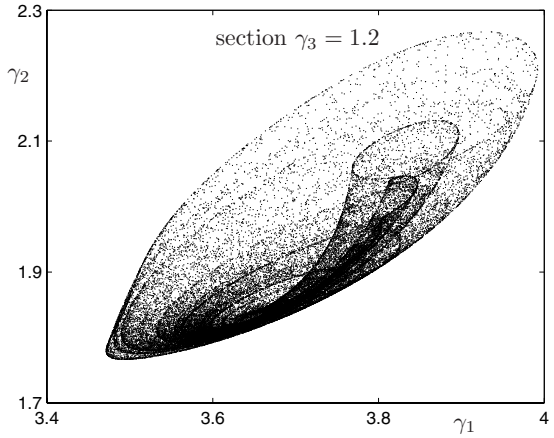


Figure 19. *Chaotic attractor through a Poincaré section taken at* $\gamma_3 = \theta_3 - \theta_5 = 1.2$ *by projecting into the plane of* $\gamma_1 = \theta_1 - \theta_5$ *and* $\gamma_2 = \theta_2 - \theta_5$. *The parameters* (7.1) *are considered and no external perturbation is added* $\epsilon = \eta = \delta = 0$.

The studied dynamics shows that complex phase clustering can occur by spontaneous symmetry breaking in globally coupled systems, even in the absence of any inhomogeneity. Both the heteroclinic network (which is sensitive to arbitrarily small inputs) and the excitable network (which is only sensitive to finite amplitude inputs) can be used to represent the dynamics of systems that are sensitive to inputs but robust against background noise and imperfections. In these systems encoding via phase differences can be used for reliable finite-state computation as well as for producing spatiotemporal codes from steady inputs [23]. These features mimic the information processing capabilities of relating neural systems, such as the antennal lobes of insects [11]. In these neural systems the encoding is not in average frequencies of the oscillators but in their phases; i.e., effectively it is a form of *spike-time coding*.

7. Discussion of future directions. In this paper a region of parameter space was studied for the model (2.1), (2.2) of $N = 5$ oscillators. The significance of $N = 5$ is that this is the smallest number of oscillators that permits nontrivial cluster formations and switching between them [4]. Although the dynamics discussed for $N = 5$ is robust, it does raise further questions:

(i) What other robust dynamics are possible for this system?

(ii) How does the dynamics change when the number of oscillators N is increased?

Concerning (i), we illustrate that in the vicinity of the explored parameter regime there is chaotic behavior. Recall that to the right of the resonance bifurcation in Figure 14 the heteroclinic network is unstable. This unstable network may coexist with a chaotic attractor at the same parameter values. Figure 19 depicts a chaotic attractor for

$$r = 0.2, \quad \alpha = 1.6, \quad \beta = -1.58, \quad \omega = 1.0 \tag{7.1}$$

near to (3.1) in the unperturbed case $\epsilon = \eta = \delta = 0$. The chaotic attractor is shown in a Poincaré section projected into the (γ_1, γ_2)-plane when $\gamma_3 = 1.2$; many symmetrical copies of

the same attractor coexist in phase space. We believe this is the first example of a chaotic attractor in a system of globally coupled identical phase oscillators with no detuning ($\delta = 0$). (For example, the chaos observed in [20] for $N \geq 4$ oscillators of the Kuramoto system requires large detuning in relation to the coupling strength.)

Concerning (ii), preliminary investigations show that systems of $N > 5$ phase oscillators also possess very rich dynamics including heteroclinic and excitable networks and chaotic attractors. In particular, for parameters (3.1) and moderate odd numbers of oscillators $N = 7, 9$ the only apparent attractor is a robust heteroclinic network of three-cluster states of type $((N-1)/2, 1, (N-1)/2)$. These are connected analogously to section 3 for $N = 5$: one of the clusters of $(N-1)/2$ oscillators is unstable to perturbations while the other is stable. These cluster states are well separated in phase space (the detection tolerance ξ may be kept the same when N is increased). We expect that most of the considerations in this paper will therefore follow through for these cases, except that the number of available three-cluster states/nodes and heteroclinic connections/directed edges are much larger (e.g., 140 states and 420 connections are available for $N = 7$, while 630 states and 2520 connections are available for $N = 9$). This combinatorial growth makes a thorough analysis much more difficult in the cases of $N > 5$. Note that for other parameters we have found heteroclinic networks between cluster states of other types. However, in order to find a certain ℓ-cluster state, more Fourier harmonics may need to be included in the coupling function. These preliminary results predict that the number of available cluster states can be enormously large for real neural ensembles.

Appendix A. Constants for eigenvalues and eigenvectors. The real part of (3.7) can be expressed as

$$\mu = \tfrac{1}{10}\Big(g'(\chi) + 2g'(-\chi) + g'(\psi) + 2g'(-\psi) + 2g'(\chi-\psi) + 2g'(\psi-\chi)\Big), \tag{A.1}$$

while its imaginary part contains

$$\begin{aligned}\nu = -\mu^2 + \tfrac{4}{100}\Big(& g'(\chi)g'(\psi) + 2g'(\chi)g'(-\psi) + 2g'(-\chi)g'(\psi) \\ & + 2g'(\chi-\psi)\Big(g'(\psi) + 2g'(-\chi) + 2g'(-\psi)\Big) \\ & + 2g'(\psi-\chi)\Big(g'(\chi) + 2g'(-\chi) + 2g'(-\psi)\Big)\Big).\end{aligned} \tag{A.2}$$

The real parts of the components of (3.9) are

$$\begin{aligned} r_{\mathrm{y}} = -\frac{1}{4Q}\Big(& N_1\Big(g'(\chi) - g'(\psi) - 2g'(-\chi) - 2g'(-\psi) + 2g'(\chi-\psi) + 2g'(\psi-\chi)\Big) \\ & - 4g'(\chi)g'(-\psi)g'(\psi-\chi) + 4g'(-\chi)g'(\psi)g'(\chi-\psi)\Big), \\ r_{\mathrm{b}} = \frac{1}{4Q}\Big(& N_3\Big(-g'(\chi) + g'(\psi) - 2g'(-\chi) - 2g'(-\psi) + 2g'(\chi-\psi) + 2g'(\psi-\chi)\Big) \\ & + 4g'(\chi)g'(-\psi)g'(\psi-\chi) - 4g'(-\chi)g'(\psi)g'(\chi-\psi)\Big), \end{aligned} \tag{A.3}$$

while the imaginary parts are

$$\begin{aligned} p_{\mathrm{y}} &= -\frac{1}{4Q} N_1 \sqrt{P}, \\ p_{\mathrm{b}} &= \frac{1}{4Q} N_3 \sqrt{P}, \end{aligned} \tag{A.4}$$

where

$$\begin{aligned} N_1 &= g'(\chi) g'(-\psi) + 2g'(\chi - \psi)\Big(g'(-\chi) + g'(-\psi)\Big), \\ N_3 &= g'(-\chi) g'(\psi) + 2g'(\psi - \chi)\Big(g'(-\chi) + g'(-\psi)\Big), \end{aligned} \tag{A.5}$$

and

(A.6)

$$\begin{aligned} P = &-\Big(g'(\chi) - g'(\psi)\Big)^2 - 4\Big(g'(-\chi) + g'(-\psi)\Big)^2 - 4\Big(g'(\chi - \psi) + g'(\psi - \chi)\Big)^2 \\ &- 4\Big(g'(\chi) - g'(\psi)\Big)\Big(g'(-\chi) - g'(-\psi)\Big) - 4\Big(g'(\chi) - g'(\psi)\Big)\Big(g'(\chi - \psi) - g'(\psi - \chi)\Big) \\ &+ 8\Big(g'(-\chi) + g'(-\psi)\Big)\Big(g'(\chi - \psi) + g'(\psi - \chi)\Big), \\ Q = &\, g'(-\chi) g'(-\psi)\Big(g'(\chi) - g'(\psi)\Big) + \Big(g'(-\chi) + g'(-\psi)\Big)\Big(2g'(-\chi) g'(\chi - \psi) \\ &- 2g'(-\psi) g'(\psi - \chi)\Big). \end{aligned}$$

Appendix B. The operator O_p.

Table 2

The effects of operator O_p, $p = 1, \ldots, 5$, in (3.13) on the three-cluster states s_i, $i = 1, \ldots, 30$. The resulting states are shown only when $i \neq j$ for $O_p(s_i) = s_j$; in all other cases $O_p(s_i) = (s_i)$.

	$O_1(s_i)$	$O_2(s_i)$	$O_3(s_i)$	$O_4(s_i)$	$O_5(s_i)$
s_1				s_{12}	s_{13}
s_2	s_{20}	s_{30}			
s_3		s_{27}			s_{17}
s_4	s_{23}			s_{10}	
s_5		s_{26}		s_7	
s_6	s_{22}				s_{15}
s_7	s_{19}				s_{18}
s_8		s_{26}	s_6		
s_9	s_{23}		s_3		
s_{10}		s_{29}			s_{16}
s_{11}			s_2		s_{13}
s_{12}	s_{21}	s_{28}			
s_{13}	s_{24}	s_{25}			
s_{14}			s_2	s_{12}	
s_{15}		s_{29}		s_9	
s_{16}	s_{22}		s_5		
s_{17}	s_{19}			s_8	
s_{18}		s_{27}	s_4		
s_{19}		s_{30}	s_1		
s_{20}				s_8	s_{18}
s_{21}			s_5		s_{15}
s_{22}		s_{28}		s_{11}	
s_{23}		s_{25}			s_{14}
s_{24}			s_3	s_{10}	
s_{25}			s_6	s_7	
s_{26}	s_{24}				s_{14}
s_{27}	s_{21}			s_{11}	
s_{28}			s_4		s_{17}
s_{29}	s_{20}		s_1		
s_{30}				s_9	s_{16}

Acknowledgments. The authors thank Roman Borisyuk for discussions on neurodynamics and Stephen Coombes and Sebastian Wieczorek for asking perceptive questions.

REFERENCES

[1] V. S. Afraimovich, M. I. Rabinovich, and P. Varona, *Heteroclinic contours in neural ensembles and the winnerless competition principle*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 14 (2004), pp. 1195–1208.

[2] P. Ashwin and J. Borresen, *Encoding via conjugate symmetries of slow oscillations for globally coupled oscillators*, Phys. Rev. E (3), 70 (2004), 026203.

[3] P. Ashwin and J. Borresen, *Discrete computation using a perturbed heteroclinic network*, Phys. Lett. A, 347 (2005), pp. 208–214.

[4] P. Ashwin, O. Burylko, Yu. Maistrenko, and O. Popovych, *Extreme sensitivity to detuning for globally coupled phase oscillators*, Phys. Rev. Lett., 96 (2006), 054102.
[5] P. Ashwin and J. W. Swift, *The dynamics of n weakly coupled identical oscillators*, J. Nonlinear Sci., 2 (1992), pp. 69–108.
[6] P. Ashwin and M. Timme, *When instability makes sense*, Nature, 436 (2005), pp. 36–37.
[7] E. Brown, P. Holmes, and J. Moehlis, *Globally coupled oscillator networks*, in Perspectives and Problems in Nonlinear Science: A Celebratory Volume in Honor of Larry Sirovich, K. Sreenivasan, E. Kaplan, and J. Marsden, eds., Springer, New York, 2003, pp. 183–215.
[8] E. Brown, J. Moehlis, and P. Holmes, *On the phase reduction and response dynamics of neural oscillator populations*, Neural Computation, 16 (2004), pp. 673–715.
[9] S.-N. Chow, B. Deng, and B. Fiedler, *Homoclinic bifurcation at resonant eigenvalues*, J. Dynam. Differential Equations, 2 (1990), pp. 177–244.
[10] E. J. Doedel, A. R. Champneys, T. F. Fairgrieve, Yu. A. Kuznetsov, B. Sandstede, and X. Wang, *AUTO97: Continuation and Bifurcation Software for Ordinary Differential Equations*, Tech. report, Department of Computer Science, Concordia University, Montreal, Canada, 1997, http://indy.cs.concordia.ca/auto/.
[11] C. G. Galizia, S. Sachse, A. Rappert, and R. Menzel, *The glomerular code for odor representation is species specific in the honeybee apis mellifera*, Nature Neuroscience, 2 (1999), pp. 473–478.
[12] M. Golubitsky and I. N. Stewart, *The Symmetry Perspective: From Equilibrium to Chaos in Phase Space and Physical Space*, Progr. Math. 200, Birkhäuser Boston, Boston, 2003.
[13] D. Hansel, G. Mato, and C. Meunier, *Clustering and slow switching in globally coupled phase oscillators*, Phys. Rev. E, 48 (1993), pp. 3470–3477.
[14] D. Hansel, G. Mato, and C. Meunier, *Phase dynamics for weakly coupled Hodkin-Huxley neurons*, Europhys. Lett., 23 (1993), pp. 367–372.
[15] B. Hasselblatt and A. Katok, *The development of dynamics in the 20th century and the contribution of Jürgen Moser*, Ergodic Theory Dynam. Systems, 22 (2002), pp. 1343–1364.
[16] E. M. Izhikevich, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*, MIT Press, Cambridge, MA, 2007.
[17] H. Kori and Y. Kuramoto, *Slow switching in globally coupled oscillators: Robustness and occurrence through delayed coupling*, Phys. Rev. E, 63 (2001), 046214.
[18] M. Krupa, *Robust heteroclinic cycles*, J. Nonlinear Sci., 7 (1997), pp. 129–176.
[19] Y. Kuramoto, *Self-entrainment of a population of coupled non-linear oscillators*, in International Symposium on Mathematical Problems in Theoretical Physics, Lecture Notes in Phys. 39, H. Araki, ed., Springer-Verlag, New York, 1975, pp. 420–422.
[20] Yu. Maistrenko, O. Popovych, and M. Hasler, *On strong and weak chaotic partial synchronization*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 10 (2000), pp. 179–203.
[21] K. Okuda, *Variety and generality of clustering in globally coupled oscillators*, Phys. D, 63 (1993), pp. 424–436.
[22] G. Orosz and G. Stépán, *Subcritical Hopf bifurcations in a car-following model with reaction-time delay*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 462 (2006), pp. 2643–2670.
[23] M. I. Rabinovich, P. Varona, A. I. Selverston, and H. D. I. Abarbanel, *Dynamical principles in neuroscience*, Rev. Modern Phys., 78 (2006), pp. 1213–1265.
[24] E. Stone and P. Holmes, *Random perturbations of heteroclinic attractors*, SIAM J. Appl. Math., 50 (1990), pp. 726–743.