J. Fluid Mech. (2004), *vol.* 499, *pp.* 277–314.
DOI: 10.1017/S0022112003007109 Printed in the United Kingdom

On finite-amplitude subcritical instability in narrow-gap spherical Couette flow

By ANDREW P. BASSOM[1,2] AND ANDREW M. SOWARD[1]

[1]Department of Mathematical Sciences, University of Exeter, Exeter EX4 4QE, Devon, UK

[2]School of Mathematics, University of New South Wales, Sydney 2052, Australia

(Received 21 January 2003 and in revised form 17 September 2003)

We consider the finite-amplitude instability of incompressible spherical Couette flow between two concentric spheres of radii R_1 and R_2 ($>R_1$) in the narrow-gap limit, $\varepsilon \equiv (R_2 - R_1)/R_1 \ll 1$, caused by rotating them both about a common axis with distinct angular velocities Ω_1 and Ω_2 respectively. In this limit it is well-known that the onset of (global) linear instability is manifested by Taylor vortices of roughly square cross-section close to the equator. According to linear theory this occurs at a critical Taylor number $T_{\rm crit}$ which, remarkably, exceeds the local value T_c obtained by approximating the spheres as cylinders in the vicinity of the equator even as $\varepsilon \downarrow 0$. Previous theoretical work on this problem has concentrated on the case of almost co-rotation with $\delta \approx (\Omega_1 - \Omega_2)/\Omega_1 = O(\varepsilon^{1/2})$ for which $T_{\rm crit} = T_c + O(\delta^2) + O(\varepsilon)$. In this limit the amplitude equation that governs the spatio-temporal modulation of the vortices on the latitudinal extent $O(\varepsilon^{1/2}R_1)$ gives rise to an interesting bifurcation sequence. In particular, the appearance of global bifurcations heralds the onset of complicated subcritical time-dependent finite-amplitude solutions.

Here we switch attention to the case when $\varepsilon^{1/2} \ll \delta \leqslant 1$. We show that for Taylor numbers $T = T_c + O((\delta\varepsilon)^{2/3})$ there exists a locally unstable region of width $O((\delta\varepsilon)^{1/3}R_1)$ within which the amplitude equation admits solutions in the form of pulse-trains. Each pulse oscillates at a frequency proportional to its distance from the equatorial plane and consists of a wave propagating towards the equator under an envelope. The pulse drifts at a slow speed (relative to the wave velocity) proportional to its distance (and away) from the equator. Both the wavelength and the envelope width possess the same relatively short length scale $O((\varepsilon^2/\delta)^{1/3}R_1)$. The appropriate theory of spatially periodic pulse-trains is developed and numerical solutions found. Significantly, these solutions are strongly subcritical and have the property that $T \to T_c$ as $\varepsilon \downarrow 0$.

Two particular limits of our theory are examined. In the first, $\varepsilon^{1/2} \ll \delta \ll 1$, the spheres almost co-rotate and the pulse drift velocity is negligible. A comparison is made of the pulse-train predictions with previously obtained numerical results pertaining to large (but finite) values of $\delta/\varepsilon^{1/2}$. The agreement is excellent, despite the complicated long-time behaviour caused by inhomogeneity across the relatively wide unstable region.

Our second special case $\delta = 1$ relates to the situation when the outer sphere is at rest. Now the poleward drift of the pulses leads to a slow but exponential increase of their separation with time. This systematic pulse movement, over and above the spatial inhomogeneity just mentioned, necessarily leads to complicated and presumably chaotic spatio-temporal behaviour across the wide unstable region of width $O(\varepsilon^{1/3}R_1)$ on its associated time scale, which is $O(\varepsilon^{-1/3})$ longer than the

wave period. In view of the several length and time scales involved only qualitative comparison with experimental results is feasible. Nevertheless, the pulse-train structure is robust and likely to provide the building block of the ensuing complex dynamics.

1. Introduction

Our study concerns the motion of incompressible viscous fluid confined between concentric spheres of radii R_1 and R_2 ($>R_1$) in the narrow gap limit,

$$\varepsilon \equiv (R_2 - R_1)/R_1 \ll 1, \tag{1.1}$$

caused by rotating them both about a common axis with distinct angular velocities Ω_1 and Ω_2 respectively. Unlike classical Couette flow between concentric cylinders, which consists of pure azimuthal motion, spherical Couette flow exhibits a secondary axisymmetric meridional circulation induced as a consequence of the spherical geometry. In most experiments the outer sphere is at rest ($\Omega_2 = 0$) and so many of the numerical investigations have focused on that case. Studies have shown that the bifurcation sequence that ensues depends on the gap aspect ratio ε and, at least for the case of the stationary outer sphere, the experiments highlight three parameter regimes – these can be loosely classified as narrow ($\varepsilon < 0.12$), medium ($0.12 < \varepsilon < 0.24$) and wide ($\varepsilon > 0.24$) gap geometries. For narrow to medium-sized gaps the undisturbed flow first becomes susceptible to axisymmetric vortices; these are analogous to the familiar Taylor vortices which occur in the cylindrical geometry. The vortices form in the vicinity of the equator of the spherical system – in some cases the cells have been observed to be symmetric with respect to the equatorial plane while in other circumstances they are asymmetric (see, for example, Bühler 1990). The numerical results of Marcus & Tuckerman (1987*a*, *b*), who largely focus their attention on the medium-gap case, indicate that even the transitions between the basic state and alternative steady states may be very complicated. Further bifurcations may lead to time-dependent structures, either axisymmetric (see, for example, Bartels 1988; Mamum & Tuckerman 1995) or non-axisymmetric (Nakagayashi 1983). Experimental studies (Nakagayashi 1983; Nakagayashi & Tsuchida 1988*a*, *b*) for $\varepsilon = 0.14$ demonstrate that the non-axisymmetric flow consists of spiral vortices. These have also been identified in numerical results for the same gap width by Sha & Nakagayashi (2001) and in the narrow gap ($\varepsilon = 0.06$) by Dumas & Leonard (1994).

Wimmer (1976, 1981) describes experimental results for a variety of gap widths in the range $0.0063 \leqslant \varepsilon \leqslant 0.6$. His findings for the narrow-gap limit show that, at the onset of instability, axisymmetric Taylor vortices with their roughly square cross-section are localized in the vicinity of the equator. Such a configuration may be analysed by multiple-scale asymptotic methods, based on the idea that the structure of each vortex on its short $O(\varepsilon R_1)$ length scale is determined by conditions locally whereas its amplitude varies over a longer latitudinal length scale fixed by a higher-order theory. The early studies of Walton (1978), Hocking & Skiepko (1981) and Hocking (1981) assumed that the critical Taylor number $T_{\rm crit}$ is simply a perturbation of the value T_c obtained by approximating the equatorial shellular region by infinite cylinders. Thus 'local instability' occurs when the Taylor number T exceeds T_c, whereas for 'global instability' it is necessary that $T > T_{\rm crit}$. However, once the effects of the spatial modulation are correctly accounted for, Soward & Jones (1983) have shown that $T_{\rm crit} > T_c$. This distinction between 'local' and 'global instability' is well-known in

the context of spatially evolving shear flows (see, for example, Huerre & Monkewitz 1990).

There are two important physical ingredients involved in the latitudinal modulation of the vortices. One is boundary curvature, while the other is the secondary meridional circulation present in the basic state (the primary flow being azimuthal). Both lead to the physical mechanism of phase mixing (a name adopted in astrophysical contexts; see, for instance, Heyvaerts & Priest 1983), which tends to shorten length scales and so enhance dissipation. This stabilizing process is so potent that it causes the true critical (or global) Taylor number $T_{\rm crit}$ to exceed the local cylinder value T_c in the limit $\varepsilon \downarrow 0$ (Soward & Jones 1983) as mentioned above.

Harris, Bassom & Soward (2000) extended the Soward & Jones (1983) linear analysis for axisymmetric Taylor vortices into the weakly nonlinear regime by including the Stuart–Landau term as derived by Davey (1962). To appreciate the nature of the asymptotic method we recap briefly the notation employed previously by both Soward & Jones (1983) and Harris *et al.* (2000). The relative sizes of the angular velocities and angular momenta of the two boundaries are measured by

$$\widetilde{\mu} \equiv \Omega_2/\Omega_1, \qquad \widetilde{\delta} \equiv (1-\widetilde{\mu})/(1+\widetilde{\mu}), \tag{1.2a, b}$$

and

$$\mu \equiv R_2^2\Omega_2\big/R_1^2\Omega_1, \qquad \delta \equiv (1-\mu)/(1+\mu), \tag{1.2c, d}$$

respectively. Scaled Reynolds and Taylor numbers are introduced according to

$$R_M \equiv \varepsilon^2(1-\mu)R_1^2\Omega_1\big/\nu, \qquad T \equiv \varepsilon^{-1}\left(1+\tfrac{1}{2}\varepsilon\right)^{-3}\delta^{-1}R_M^2, \tag{1.3a, b}$$

where ν denotes the kinematic viscosity of the fluid.

The Soward & Jones (1983) theory places no restriction on the angular momentum ratio μ. However to construct their amplitude equation (see Appendix C), Harris *et al.* (2000, 2003) were obliged to restrict attention to the case of small δ which means that the angular momentum ratio is close to unity and for instability the spheres need to rotate extremely fast. As these papers explained, this is not a parameter range readily accessed by experiment. Their theory, like ours here, could be extended to encompass non-axisymmetric modes, which would be relevant, for example, to the spiral vortices. For $O(1)$ azimuthal wavenumber, however, no new terms are introduced at the levels of approximation considered and so these small asymmetries do not influence their conclusions at leading order. In this paper we wish to investigate axisymmetric finite-amplitude instability within the framework of finite δ. Indeed, our analysis will be valid for the important case when the outer sphere is at rest so that $\delta = 1$.

With regard to the onset of instability, we have already explained why the global critical Taylor number $T_{\rm crit}$ exceeds the local cylinder critical value T_c and, in terms of order of magnitude, these are related by

$$T_{\rm crit} = T_c + O(\delta^2) + O(\varepsilon). \tag{1.4}$$

For $\delta \geqslant O(\varepsilon^{1/2})$ there is a region of local instability of breadth $O(\delta R_1)$ about the equator. Significantly, the critical linear marginal mode has width $O(\varepsilon^{1/2}R_1)$ independent of the size of δ. So though the mode fills the locally unstable region when $\delta = O(\varepsilon^{1/2})$ it is much thinner when $\delta \gg \varepsilon^{1/2}$ leaving no disturbance throughout the remaining bulk of the region (see (4.8) and (4.10)). This is an unnatural feature that the nonlinear theory needs to address.

The comparability in the sizes of the correction terms in (1.4) suggested to Harris *et al.* (2000) the importance of the parameter range $\delta = O(\varepsilon^{1/2})$, for which

their $O(\delta/\varepsilon^{1/2})$ parameter $\widetilde{\kappa}$ (see equation (C1b)) is $O(1)$. For this parameter range Harris *et al.* (2003, henceforth referred to as HBS) traced complicated bifurcation sequences, which in all cases led via global bifurcations to travelling waves. Indeed, for numerically large values of $\widetilde{\kappa}$ highly subcritical finite-amplitude states were identified, suggesting that they may exist in small-amplitude form for

$$T = T_c + O\big((\delta\varepsilon)^{2/3}\big) \tag{1.5}$$

close to the local cylinder critical value T_c, in the sense that $T \to T_c$ (while $T_{\rm crit} \to T_c + O(\delta^2)$) as $\varepsilon \downarrow 0$.

In this respect an interesting comparison can be made with a related problem of thermal convection in a rotating self-gravitating sphere. That problem also exhibits the destructive effects of phase mixing with the consequence that the global critical Rayleigh number exceeds the local critical Rayleigh number (Jones, Soward & Mussa 2000). Previously, however, Soward (1977) proposed that nonlinearity could counterbalance the phase mixing and so permit small finite-amplitude motion close to the local critical value, just as in the suggestion (1.5). He uncovered localized solutions, which we will call pulses. The possibility of corresponding isolated pulse solutions for the spherical Couette flow problem at Taylor numbers close to the local critical value T_c was first investigated by Hocking (private communication, *circa* 1980) but he was unable to identify any such features. Ewen & Soward (1994a) sought localized pulse solutions for an amplitude equation which, although more general than the one of concern here, did include our fundamental Complex–Ginzberg–Landau equation (2.17) (see also (2.22)) as a special case. The form of their solutions supported Hocking's conclusion that no localized pulses could exist for the spherical Couette problem close to T_c. In view of this negative result Ewen & Soward (1994b) proceeded to investigate pulse-trains for their more general system without considering the spherical Couette case – as this flow did not allow isolated pulses it was natural to conclude that pulse-trains were also unfeasible. In contrast, some recent simulations reported by HBS for numerically large values of $\delta/\varepsilon^{1/2}$ suggest that pulse-train solutions might exist after all. An exploration of this possibility is the main objective for this paper.

We outline briefly the nature of the pulse-train solutions suggested by HBS's results in the limit $\varepsilon^{1/2} \ll \delta \ll 1$. To begin, the Taylor number T must be of size (1.5) and then there is a locally unstable region centred on the equator and of width $O((\delta\varepsilon)^{1/3}R_1)$. Inside this zone there are $O((\delta^2/\varepsilon)^{1/3})$ individual pulses each of extent $O((\varepsilon^2/\delta)^{1/3}R_1)$. Each pulse oscillates in time at a frequency proportional to the distance $R_1|\theta|$ between its centre and the equator (here $-\theta$ denotes the latitude of the pulse). The centres remain almost stationary and are separated by the same $O((\varepsilon^2/\delta)^{1/3}R_1)$ distance. As a consequence, the frequency increment $O((\delta/\varepsilon^2)^{2/3}(\nu/R_1^2))$ between any two neighbouring pulses is the same and this enables each pulse to interact coherently with its neighbours. The spatial structure of each pulse is similar and is characterized by a wave with wavelength of the same size as its pulse envelope. In view of the spatial linear increase of the temporal frequency, the envelope wave travels towards the equator at a phase speed $O((\delta/\varepsilon^2)^{2/3}|\theta|\nu/R_1)$. Notice that the narrower $O(\varepsilon R_1)$ width of the vortices means that they drift towards the equator at the slower true phase speed $O((\delta/\varepsilon)|\theta|\nu/R_1)$. Energy propagates away from the equator at the group velocity, which is also linear in $|\theta|$ and has exactly the same order of magnitude. This causes the pulse centres to separate uniformly on the time scale $O((\varepsilon/\delta)(R_1^2/\nu))$ which is long compared with the frequency increment time scale $O((\varepsilon^2/\delta)^{2/3}(R_1^2/\nu))$.

The picture just described is essentially spatially uniform and as such is the lowest-order solution on the pulse-train length scale $O((\varepsilon^2/\delta)^{1/3}R_1)$. The background on which the pulses ride is inhomogeneous on the longer length scale $O((\delta\varepsilon)^{1/3}R_1)$ of the unstable region. Indeed, because of this inhomogeneity, the pulse-train solutions found by HBS for numerically large values of $\delta/\varepsilon^{1/2}$ exhibit complicated temporal behaviour over a longer time scale. For certain parameter values HBS identified a long-time beating frequency but we shall not address such issues in this paper. Rather, we simply construct localized pulse-train solutions and then merely speculate on how this low-order structure fits into the complete picture. Our numerical evidence suggests that the structures are robust and provide the building blocks for the realized solutions which without doubt exhibit complex spatio-temporal behaviour on long length and time scales.

We have stressed the small-δ limit of almost co-rotation because we can make comparisons with the comprehensive findings of HBS, who ignored the role of the small but finite group velocity. Nevertheless our pulse-train theory together with our order-of-magnitude estimates continue to apply even when $\delta = 1$ which corresponds to the outer sphere being at rest. For $O(1)$ values of δ the continual increase of the pulse separation provides a further temporal non-uniformity.

There is continuing interest in nonlinear frequency selection mechanisms when the background state is slowly varying spatially, as it is for our spherical Couette flow problem. These problems are generally characterized by a single frequency, whose value follows after the solution of a nonlinear eigensystem. Pier, Huerre & Chomaz (2001) discuss two types of localized mode that they denote soft and steep global modes. Whereas the amplitude of the soft mode evaporates smoothly at each end, the steep mode terminates abruptly at one end, where the finite-amplitude state collapses across a front (Pier *et al.* 1998). Such solutions have been identified in a variety of physical systems (Meunier *et al.* 1997; Pier & Huerre 2001; Pier 2002). For them the role of the group velocity is vital in the frequency selection process (Couairon & Chomaz 1997; Tobias, Proctor & Knobloch 1998) and is related to the Dee & Langer (1983) criterion, which requires the group velocity to be directed into the mode across the front. Each of our pulses drifts at the group velocity away from the equator and, since there is no relative velocity, is thus more like a soft mode. Nevertheless, our pulse-train itself must terminate abruptly sufficiently far from the equator, where local conditions can no longer support the finite-amplitude state. There the final pulse in the train is followed by the zero-amplitude state and so, though frontal in appearance, this is not a genuine front in the steep mode sense. Importantly, each pulse has a distinct frequency and so the pulse-train is a generalization of the single pulse notion considered by the aforementioned authors. We stress these related ideas because we believe that in the spherical Couette flow problem frequency selection is essentially a local criterion, although it is rendered discrete by the necessity of filling the locally unstable region with pulses possessing resonating frequencies.

The remainder of our paper is organized as follows. In § 2 we recap the derivation of the amplitude equation, which governs the modulation of the complex amplitude of the Taylor vortices, from the partial differential equations for the spherical Couette flow system. We make a locally homogeneous approximation, which is valid on an intermediate length scale large compared to the pulse width $O((\varepsilon^2/\delta)^{1/3}R_1)$ but short compared to the $O((\delta\varepsilon)^{1/3}R_1)$ extent of the locally unstable region. This allows identification of the structure of the pulse-train solution within the ensuing simplified system. Spatio-temporally periodic solutions are found in § 3 and these have an elegant structure which is best addressed numerically using a Fourier series in space

and solving an initial value problem in time. The ensuing temporally periodic solutions are linked to a Fourier series in time (Appendix A), which defines the pulse structure; the link is effected by Fourier transforms (Appendix B). In § 4 we consider our two limiting cases and, first, in § 4.1, investigate the limit $\varepsilon^{1/2} \ll \delta \ll 1$ of almost co-rotation. We compare our results with computations by HBS pertaining to large values of $\delta/\varepsilon^{1/2}$. Both quantitative and qualitative comparisons are excellent and fully support the estimates that we made above. The agreement provides reassuring evidence that the pulse-trains can be modulated and so fill the complete locally unstable region albeit with complicated and possibly chaotic long-time behaviour. Though the pulse-train theory still applies when the outer sphere is at rest ($\delta = 1$), the long-time theory now involves the additional complication of the small poleward group velocity. This means that the pulses are no longer stationary but drift apart away from the equator. Since the systematic slow separation is the same for all adjacent pulse pairs we are able to show in § 4.2 that our pulse-trains can be embedded within a temporally uniformly valid approximation. In the final discussion in § 5 we comment on the spatio-temporal behaviour of the actual Taylor vortices as predicted by our solutions and examine the extent to which our findings are supported by experiments and numerical simulations. We provide an alternative survey of our work in a summary article (Soward & Bassom 2003).

2. The governing equations

We follow the formulation adopted by Soward & Jones (1983) in which the governing equations are made dimensionless by using R_1 for the unit of length and $\varepsilon^2 R_1^2/\nu$ as the unit of time. In consequence, the phase and group velocity of waves are measured in terms of $\varepsilon^{-2}\nu/R_1$ which should be contrasted with $R_1\Omega_1$, the adopted unit of fluid velocity. The system is referred to spherical polar coordinates $(r, \pi/2 + \theta, \phi)$, where $-\theta$ is the latitude and $\theta = 0$ is the equator. Relative to our spherical polar coordinates we write

$$\boldsymbol{u} \equiv \frac{1}{r\cos\theta}\left(\frac{R_M\varepsilon}{r}\frac{\partial\psi}{\partial\theta},\ -R_M\frac{\partial\psi}{\partial z},\ h\right), \qquad r \equiv 1 + \varepsilon z, \tag{2.1a, b}$$

where the inner sphere boundary is $z = 0$ and the outer is $z = 1$. Our solution

$$\psi = \Psi + \widehat{\psi}, \qquad h = H + \delta\widehat{h} \tag{2.2a, b}$$

is then composed of the basic flow (Ψ, H) and a perturbation $(\widehat{\psi}, \delta\widehat{h})$ to it.

The basic steady flow is given correct to leading order by

$$\Psi \approx (1 + \widetilde{\mu})\sin 2\theta\cos\theta\, F(z), \qquad H \approx (1 + \widetilde{\mu})\cos^2\theta\, G(z), \tag{2.3a, b}$$

where

$$F(z) = \frac{1 - \widetilde{\mu}}{5!\,(1 - \mu)}\, z^2(1 - z)^2\left[\tfrac{5}{2} - \widetilde{\delta}\left(z - \tfrac{1}{2}\right) + O(\varepsilon)\right] \tag{2.3c}$$

and

$$G(z) = \tfrac{1}{2} - \widetilde{\delta}\left(z - \tfrac{1}{2}\right) + O(\varepsilon). \tag{2.3d}$$

From the definitions (1.2) we note

$$(1 - \widetilde{\mu})/(1 - \mu) = 1 + O(\varepsilon/\delta), \qquad \widetilde{\delta} = \delta + O(\varepsilon) \tag{2.4a, b}$$

and, since we are interested in the limit $\delta \gg \varepsilon^{1/2}$, we assume that these two quantities take the values unity and δ respectively in our applications. We will also ignore the $O(\varepsilon)$ error terms in the representations (2.3c, d) of $F(z)$ and $G(z)$.

We consider perturbations to our basic state which are given locally by

$$\begin{bmatrix} \widehat{\psi} \\ \widehat{h} \end{bmatrix} = \begin{bmatrix} -\mathrm{i} f(z) \\ g(z) \end{bmatrix} \exp\left[\mathrm{i}\left(\int \frac{k}{\varepsilon}\,\mathrm{d}\theta - \omega\tau\right)\right], \tag{2.5}$$

where τ is the time. Correct to lowest order the functions $f(z)$ and $g(z)$ satisfy the equations

$$(\mathrm{D}^2 - k^2 + \mathrm{i}\omega)(\mathrm{D}^2 - k^2) f + \mathrm{i}kT \sin 2\theta\,[F'(\mathrm{D}^2 - k^2) - F'''] f + [2\mathrm{i}\sin\theta\,(G\mathrm{D} + G') - 2kG\cos\theta]\, g = 0, \tag{2.6a}$$

$$(\mathrm{D}^2 - k^2 + \mathrm{i}\omega)\, g + \mathrm{i}kTF' \sin 2\theta\, g - k\delta^{-1} TG' \cos\theta\, f = 0 \tag{2.6b}$$

and boundary conditions

$$f = \mathrm{D}f = g = 0 \qquad \text{at} \qquad z = 0 \quad \text{and} \quad z = 1, \tag{2.7}$$

where both the prime and D denote differentiation with respect to z. The solution of this eigenvalue problem determines a dispersion relation

$$\omega = \omega(\theta,\, k,\, T) \tag{2.8}$$

which we shall later solve numerically. The formulation here is entirely equivalent to (2.15) and (2.16) of Soward & Jones (1983) which itself generalizes (4.1) and (4.2) of Walton's (1978) $\delta = 1$ case.

2.1. *Local structure*

We introduce the local critical Taylor number $T_l(\theta_l)$, frequency $\omega = \omega_l(\theta_l)$ and wavenumber $k = k_l(\theta_l)$ at latitude $-\theta_l$, which are all real, satisfy (2.8) and are chosen to minimize T. These local values are characterized by a real group velocity

$$\mathrm{Im}\,\{[\omega_{,k}]_l\} = 0, \tag{2.9}$$

where the subscript notation $,\bullet$ denotes the partial derivative with respect to $\bullet$. The minimum value of T_l occurs on the equator where it determines the local critical cylinder values

$$T_c = T_l(0), \quad \omega_c = \omega_l(0) = 0, \quad k_c = k_l(0), \quad \theta_c = \theta_l = 0; \tag{2.10}$$

the subscript c, for local values at $\theta_l = 0$, is used to stress the cylinder nature of the approximation. Moreover, at the equator the phase mixing is real and so

$$\mathrm{Im}\,\{[\omega_{,\theta}]_c\} = 0. \tag{2.11}$$

Since we are only concerned with the behaviour of the solution near the equator where $|\theta_l| \ll 1$ we can expand $\omega_l \equiv \omega(\theta_l, k_l, T_l)$ as a Taylor series about the local minimum value $\omega_c \equiv \omega(\theta_c, k_c, T_c)$. Equating the real and imaginary parts determines

$$\omega_l = [\omega_{,\theta}]_c\, \theta_l + [\omega_{,\theta k}]_c\, \theta_l\, (k_l - k_c) + \ldots \tag{2.12a}$$

and

$$-[\omega_{,T}]_c\, (T_l - T_c) = \tfrac{1}{2}\{[\omega_{,\theta\theta}]_c\, \theta_l^2 + [\omega_{,kk}]_c\, (k_l - k_c)^2\} + \ldots, \tag{2.12b}$$

respectively. Likewise the vanishing of the imaginary part of the group velocity $[\omega_{,k}]_l$ yields

$$-[\omega_{,kT}]_c\,(T_l - T_c) = [\omega_{,kk}]_c\,(k_l - k_c) + \ldots \tag{2.12c}$$

so that, correct to lowest order, we have

$$k_l = k_c + O\left(\theta_l^2\right), \qquad \omega_l = [\omega_{,\theta}]_c\,\theta_l + O\left(\theta_l^3\right), \tag{2.13a, b}$$

$$-[\omega_{,T}]_c\,(T_l - T_c) = \tfrac{1}{2}\,[\omega_{,\theta\theta}]_c\,\theta_l^2 + O\left(\theta_l^4\right), \tag{2.13c}$$

in which $[\omega_{,\theta\theta}]_c/[\omega_{,T}]_c$ is real and negative. From (2.13a, b) we determine the local phase velocity

$$c_p \equiv \omega/k \approx \left\{[\omega_{,\theta}]_c/k_c\right\}\theta_l. \tag{2.13d}$$

To leading orders the other key partial derivatives required by our theory are

$$c_g \equiv [\omega_{,k}]_l \approx [\omega_{,\theta k}]_c\,\theta_l, \qquad [\omega_{,T}]_l \approx [\omega_{,T}]_c, \tag{2.14a, b}$$

$$[\omega_{,\theta}]_l \approx [\omega_{,\theta}]_c, \qquad [\omega_{,kk}]_l \approx [\omega_{,kk}]_c, \tag{2.14c, d}$$

where c_g is the local group velocity and both $-\mathrm{i}[\omega_{,T}]_c$ and $\mathrm{i}[\omega_{,kk}]_c$ are real and positive. It should be remarked however that although $-\mathrm{i}[\omega_{,T}]_l$, $[\omega_{,\theta}]_l$ and $\mathrm{i}[\omega_{,kk}]_l$ are all real at lowest order, they are fully complex when $\theta_l \neq 0$ and neglected smaller terms are taken into account.

Our objective is to show that weakly nonlinear solutions exist in the form of localized pulses which can combine to form pulse-trains. Each pulse is localized on a length scale $O(\varepsilon^{2/3})$ and oscillates at a different frequency to its neighbour, while the pulse-trains exist on a wider region width $O(\varepsilon^{1/3})$. It is possible to discuss the form of these pulse-trains in the vicinity of some location $\theta = \theta_l = O(\varepsilon^{1/3})$ on an intermediate length scale which is large compared to $O(\varepsilon^{2/3})$ but short compared to $O(\varepsilon^{1/3})$. To that end we consider solutions in the vicinity of θ_l with the structure

$$\begin{bmatrix}\widehat{\psi}\\ \widehat{h}\end{bmatrix} = b(\theta - \theta_l, \tau)\begin{bmatrix}-\mathrm{i}f_l(z)\\ g_l(z)\end{bmatrix}\exp\left\{\mathrm{i}\left[\frac{k_l(\theta-\theta_l)}{\varepsilon} - \omega_l\tau\right]\right\} + \text{c.c.}, \tag{2.15}$$

where c.c. denotes complex conjugate and $(f_l(z), g_l(z))$ is the solution of the eigenvalue problem (2.6) with the local critical values T_l, ω_l, k_l at θ_l. Application of standard multiple-scale methods shows that $b(\theta - \theta_l, \tau)$ is governed by

$$\frac{\partial b}{\partial \tau} + \varepsilon\,[\omega_{,k}]_l\,\frac{\partial b}{\partial \theta} = \{-\mathrm{i}\,[\omega_{,T}]_l(T - T_l) - \mathrm{i}\,[\omega_{,\theta}]_l(\theta - \theta_l) - \mathscr{C}|b|^2\}b + \tfrac{1}{2}\mathrm{i}\,\varepsilon^2\,[\omega_{,kk}]_l\,\frac{\partial^2 b}{\partial \theta^2}, \tag{2.16}$$

where $\mathscr{C}$ is a real constant determined by weakly nonlinear theory, but, of course, in part depends on the normalization of the local eigenfunctions $(f_l(z), g_l(z))$.

Adopting the lowest-order approximations for the partial derivatives, (2.16) may be cast in the dimensionless form

$$\frac{\partial \mathscr{A}}{\partial t} + c\,\frac{\partial \mathscr{A}}{\partial \chi} = (\lambda + \mathrm{i}\chi - |\mathscr{A}|^2)\mathscr{A} + \frac{\partial^2 \mathscr{A}}{\partial \chi^2} \tag{2.17}$$

under the change of variables

$$\tau = \frac{t}{\left(-[\omega_{,\theta}]_c\,\mu\,\varepsilon^{2/3}\right)}, \quad \theta - \theta_l = \mu\,\varepsilon^{2/3}\chi, \quad \theta_l = \frac{\mu^2\,\varepsilon^{1/3}}{\gamma}\,\Theta_l, \tag{2.18a, b, c}$$

$$c = \frac{c_g}{\left(-[\omega_{,\theta}]_c\,\mu^2\,\varepsilon^{1/3}\right)} = \frac{[\omega_{,k\theta}]_c}{(-[\omega_{,\theta}]_c\,\gamma)}\,\Theta_l, \tag{2.18d}$$

$$\lambda = \frac{(-\mathrm{i}\,[\omega_{,T}]_c)}{\left(-[\omega_{,\theta}]_c\,\mu\,\varepsilon^{2/3}\right)}(T - T_c) - \Theta_l^2, \qquad \sqrt{\mathscr{C}}\,b = (-[\omega_{,\theta}]_c\,\mu)^{1/2}\,\varepsilon^{1/3}\,\mathscr{A}, \tag{2.18e, f}$$

where

$$\gamma \equiv \frac{\{-[\omega_{,kk}]_c[\omega_{,\theta\theta}]_c\}^{1/2}}{(-2\,[\omega_{,\theta}]_c)}\,(>0), \qquad \mu \equiv \left\{\frac{\mathrm{i}\,[\omega_{,kk}]_c}{-2\,[\omega_{,\theta}]_c}\right\}^{1/3}(>0). \tag{2.19a, b}$$

An alternative representation $\theta_l = \mu\varepsilon^{2/3}\chi_l$ is sometimes helpful with relations

$$\theta = \mu\,\varepsilon^{2/3}(\chi_l + \chi), \qquad \Theta_l = \Upsilon_\varepsilon\,\chi_l, \qquad \Upsilon_\varepsilon = \varepsilon^{1/3}\,\gamma/\mu. \tag{2.20a, b, c}$$

An equation similar to (2.17) was considered by Couairon & Chomaz (1999), but with $\mathrm{i}\chi$ replaced by χ. The absence of the i is, however, crucial and leads to solutions with very different properties. A further reduction of (2.17) is obtained by adopting the coordinate moving at the group velocity c and making the additional change of variables

$$\mathscr{A}(\chi, t; \varphi, t_0) = \exp\left[\mathrm{i}\left(\tfrac{1}{2}ct^2 + \Omega t + \varphi\right)\right] a(x, t - t_0), \tag{2.21a}$$

where

$$x = \chi - \Omega - ct. \tag{2.21b}$$

Here t_0 is an arbitrary time origin (included for later convenience), while φ is an arbitrary real phase. The arbitrary real frequency Ω can be set to zero under the shift of the local origin $\theta_l \to \theta_l + \mu\varepsilon^{2/3}\Omega$, and then (2.17) is satisfied when $a(x, t)$ solves

$$\frac{\partial a}{\partial t} = (\lambda + \mathrm{i}x - |a|^2)a + \frac{\partial^2 a}{\partial x^2}. \tag{2.22}$$

2.2. *Symmetries*

Numerical investigations of initial value problems showed that a family of stable finite-amplitude pulse-train solutions of (2.22) exists with the structure

$$a(x, t; L) = \mathrm{e}^{\mathrm{i}\pi/4}\sum_{\forall n}\exp\left[\mathrm{i}\left(2n + \tfrac{1}{2}\right)Lt\right]\overline{a}\left(x - \left(2n + \tfrac{1}{2}\right)L\right) + \mathrm{e}^{-\mathrm{i}\pi/4}\sum_{\forall n}\exp\left[-\mathrm{i}\left(2n + \tfrac{1}{2}\right)Lt\right]\overline{a}\left(x + \left(2n + \tfrac{1}{2}\right)L\right), \tag{2.23a}$$

parameterized by the constant L which measures the distance between the pulses. By an appropriate choice of the phase φ in our representation (2.21), we may restrict attention to the complex function $\overline{a}(x)$, which has the symmetry property

$$\overline{a}(-x) = \overline{a}^*(x), \tag{2.23b}$$

where the asterisk denotes complex conjugate. Such solutions exist over a finite range $L_{\min}(\lambda) \leqslant L \leqslant L_{\max}(\lambda)$, which depends on λ, provided that $\lambda > \lambda_{\inf}$ where

$$L_{\min}(\lambda_{\inf}) = L_{\max}(\lambda_{\inf}) \equiv L_{\inf}. \tag{2.24a}$$

Our numerical simulations show that

$$\lambda_{\inf} \approx 2.54074 \qquad \text{with a corresponding} \qquad L_{\inf} \approx 2.11831. \tag{2.24b}$$

On substitution of (2.23a) into (2.22) we find that $\overline{a}(x)$ satisfies the equation

$$\frac{\mathrm{d}^2\overline{a}}{\mathrm{d}x^2} + (\lambda + \mathrm{i}x)\,\overline{a} = \sum_{\forall m,n} \sigma_{m,n}\,\overline{a}(x - mL)\,\overline{a}(x - nL)\,\overline{a}^*(x - (m+n)L), \qquad (2.25a)$$

where, for integer m, n,

$$\sigma_{m,n} = \begin{cases} -1, & m \text{ and } n \text{ both odd}, \\ 1, & \text{otherwise}. \end{cases} \qquad (2.25b)$$

The required solution must meet the boundary conditions $|\overline{a}| \to 0$ as $|x| \to \infty$.

We note that the representation (2.23a) of $a(x, t)$ preserves the property (2.23b) of $\overline{a}(x)$ so that

$$a(-x, t) = a^*(x, t); \qquad (2.26)$$

moreover it also possesses the two basic spatio-temporal symmetries

$$a(x + L, t \pm T) = \mp \exp(\mathrm{i}Lt)\,a(x, t), \qquad (2.27a)$$

where the time T (not to be confused with the Taylor number) satisfies

$$LT = \pi. \qquad (2.27b)$$

From these, or directly, we may deduce that

$$a(x + 2L, t) = \exp(2\mathrm{i}Lt)\,a(x, t), \qquad a(x, t + 2T) = -a(x, t) \qquad (2.28a, b)$$

and we remark that the factor $\exp(2\mathrm{i}Lt)$, which emerges from the spatial shift $2L$, could be removed by considering the function $\exp(-\mathrm{i}tx)a(x, t)$.

The symmetries of $a(x, t)$ can be translated to the properties of $\mathscr{A}(\chi, t; \varphi, t_0)$ defined by (2.21a); the most important of these concerns its nature after the time $2L/c$ during which the pulse-train drifts a distance $2L$. Result (2.28a) can be used to show that

$$\mathscr{A}\left(\chi,\ t + \frac{2L}{c};\ \varphi,\ t_0\right) = \mathscr{A}\left(\chi,\ t;\ \varphi + 2L\left[t_0 - \frac{2L}{c}\right],\ \left[t_0 - \frac{2L}{c}\right]\right), \qquad (2.29a)$$

which generalizes to

$$\mathscr{A}\left(\chi,\ t + \frac{2NL}{c};\ \varphi,\ t_0\right) = \mathscr{A}\left(\chi,\ t;\ \varphi + 2NL\left[t_0 - (N+1)\frac{L}{c}\right],\ \left[t_0 - \frac{2NL}{c}\right]\right), \qquad (2.29b)$$

for integer N. This shows that, following the time shifts, the phase angle φ is changed, as is the time origin t_0 of the function $a(x, t)$. Furthermore the symmetry (2.28a) shows that $\mathscr{A}$ is invariant under the (effective spatial) translation $\Omega \to \Omega + 2L$, while we note the trivial invariance under the change of phase $\varphi \to \varphi + 2\pi$. Finally the symmetry (2.28b) implies that

$$\mathscr{A}(\chi,\ t;\ \varphi,\ t_0) = (-1)^M \mathscr{A}(\chi,\ t;\ \varphi,\ t_0 + 2MT), \qquad (2.30a)$$

for integer M.

When there exist integers $M = M_P$ and $N = N_P$ such that $L^2/\pi c = L/Tc = M_P/N_P$ is rational, the function $\mathscr{A}$ has the property

$$\mathscr{A}\left(\chi,\ t + \frac{2N_P L}{c};\ \varphi,\ t_0\right) = (-1)^{M_P} \mathscr{A}(\chi,\ t;\ \varphi + 2N_P L t_0,\ t_0). \qquad (2.30b)$$

Ignoring the phase shift $\varphi \to \varphi + 2N_P L t_0$, we see that $\mathscr{A}(\chi, t; \varphi, t_0)$ has period $2N_P L/c = 2M_P T$. Furthermore, the phase is also periodic for particular values of t_0. Essentially we have two periods, namely $2L/c$ and $2T$, which, if incommensurate, implies that the solution lies on a torus.

3. The pulse-train

We have presented our pulse-train (2.23) as a Fourier series in time with coefficients which, under space shifts, are generated by the single amplitude function $\overline{a}(x)$. Though these individual pulse amplitudes solve (2.25), that is not the most convenient method to determine them. Instead we consider in § 3.1 an alternative Fourier series in space, which captures the class of spatially periodic functions of period $2L = 2\pi/T$ in the sense of (2.28*a*). Then, in § 3.2, following the numerical solution of initial value problems, we identify the temporally periodic sub-class of solutions (see (3.8) below) with half-period $2T$ as defined by (2.28*b*). The coefficients of this Fourier series in space are generated, after appropriate equal interval time shifts, by a single function $\widehat{A}(t)$. In Appendix A we establish the equivalence of the Fourier series in space with time-shifted coefficients $\widehat{A}$ (necessary to achieve temporal periodicity) and the Fourier series in time with space-shifted coefficients $\overline{a}$ (necessary to achieve spatial periodicity), namely the pulse structure (2.23). Then, in Appendix B, we establish the remarkable and elegant result that $\widehat{A}(t)$ is simply the Fourier transform of $\overline{a}(x)$. We illustrate the properties of our spatio-temporal periodic solution in § 3.3 with the minimum-λ solution at $\lambda = \lambda_{\text{inf}}$ (see (2.24*b*)) and offer an explanation as to why our solutions only exist over a limited range of L at given $\lambda\,(\geqslant\lambda_{\text{inf}})$.

3.1. *Spatially periodic solutions; period* $2L$

The spatial periodicity dictated by (2.28*a*) is met by the Fourier series representation

$$a(x,t) = \exp(\mathrm{i}tx) \sum_{\forall n} A_n(t)\, \exp(\mathrm{i}nTx) \tag{3.1}$$

and the substitution of (3.1) into (2.22) shows that the functions $A_n(t)$ satisfy

$$\frac{\mathrm{d}A_n}{\mathrm{d}t} - [\lambda - (t+nT)^2]A_n = -\sum_{\forall \alpha,\beta} A_{n+\alpha}(t)\, A_{n+\beta}(t)\, A^*_{n+\alpha+\beta}(t) \tag{3.2a}$$

together with the initial conditions

$$A_n(t) \to 0 \quad \text{as} \quad t + nT \downarrow -\infty. \tag{3.2b}$$

We can make some preliminary estimates for conditions that non-trivial solutions within the spatially periodic class must satisfy. To that end we average over the periodicity length $2L$ and derive the mean energy and dissipation integrals

$$\mathscr{E}(t) = \frac{1}{2L}\int_{-L}^{L} |a(x,t)|^2\, \mathrm{d}x \;=\; \sum_{\forall n} |A_n(t)|^2, \tag{3.3a}$$

$$\mathscr{D}(t) = \frac{1}{2L}\int_{-L}^{L} \left|\frac{\partial a}{\partial x}(x,t)\right|^2 \mathrm{d}x \;=\; \sum_{\forall n} (t+nT)^2\, |A_n(t)|^2, \tag{3.3b}$$

which, in view of (2.22), satisfy

$$\frac{1}{2}\frac{\mathrm{d}\mathscr{E}}{\mathrm{d}t} = \lambda\mathscr{E} - \mathscr{D} - \frac{1}{2L}\int_{-L}^{L} |a(x,t)|^4\, \mathrm{d}x. \tag{3.4}$$

On each successive time interval $(N-1/2)T \leqslant t \leqslant (N+1/2)T$, the form of (3.3$b$) implies the inequality

$$\mathscr{D} \geqslant (t-NT)^2\,\mathscr{E} \qquad (\geqslant 0) \tag{3.5a}$$

for all integer N, while in addition use of the Cauchy–Schwartz inequality gives

$$\frac{1}{2L}\int_{-L}^{L} |a(x,t)|^4\,\mathrm{d}x \geqslant \mathscr{E}^2. \tag{3.5b}$$

The substitution of these inequalities into (3.4) shows that the growth of $\mathscr{E}$ is bounded above by

$$\frac{1}{2T}\ln\left[\frac{\mathscr{E}((N+1/2)T)}{\mathscr{E}((N-1/2)T)}\right] \leqslant \frac{1}{T}\int_{-T/2}^{T/2}[(\lambda-t^2)-\mathscr{E}]\mathrm{d}t = \lambda - \frac{T^2}{12} - \frac{1}{T}\int_{-T/2}^{T/2}\mathscr{E}\,\mathrm{d}t. \tag{3.6a}$$

In view of the non-negativity of $\mathscr{E}$, this result implies that finite-amplitude solutions necessarily decay when

$$\lambda < \frac{T^2}{12} = \frac{\pi^2}{12L^2}. \tag{3.6b}$$

This energy stability argument places a lower bound on λ for the existence of finite-amplitude solutions.

3.2. *Temporally periodic solutions; half-period* $2T$

In view of the fact that all previous evidence indicates that individual pulses cannot exist close to critical (see, for example, Ewen & Soward 1994*a*), we were doubtful that pulse-train solutions are possible. In contrast, numerical simulations presented by HBS provide compelling evidence to the contrary. So, guided by HBS, we began by tackling the initial value problem (3.2) for particular values of λ and T, together with plausible choices of the amplitudes $A_n(0)$. Our expectation was that, if such pulse-train solutions exist, they have temporal period T and the structure $A_n(t) = \widehat{A}(t+nT)$. On solving the initial value problem we were rewarded to discover that the solution quickly locked onto a stable (within our spatial periodicity class) solution. The surprise however, which supported our earlier scepticism, was to find that the solution did not have the simple period T structure. Instead the solutions acquired a period $4T$ with the more complicated form

$$A_n(t) = \mu_n \widehat{A}(t+nT) \tag{3.7a}$$

where

$$\mu_0 = \mu_1 = 1, \qquad \mu_{n+2} = -\mu_n \tag{3.7b}$$

for all integer n. The resulting representation

$$a(x,t) = \exp(\mathrm{i}tx)\sum_{\forall n}\mu_n \widehat{A}(t+nT)\exp(\mathrm{i}nTx) \tag{3.8}$$

of (3.1) satisfies all the spatio-temporal symmetries (2.27) and (2.28), while we show in Appendix A that it leads to the pulse-train structure (2.23a).

The substitution of expression (3.8) into (2.22) shows that it is a solution when $\widehat{A}(t)$ satisfies

$$\frac{\mathrm{d}\widehat{A}}{\mathrm{d}t} - \left(\lambda - t^2\right)\widehat{A} = -\sum_{\forall m,n}\sigma_{m,n}\widehat{A}\,(mT+t)\,\widehat{A}\,(nT+t)\,\widehat{A}^{*}\,((m+n)\,T+t) \tag{3.9a}$$

with $\sigma_{m,n}$ defined by (2.25*b*), where we have used the identity

$$\mu_\beta\,\mu_{m+\beta}\,\mu_{n+\beta}\,\mu_{m+n+\beta} \equiv \sigma_{m,n}, \tag{3.9b}$$

which holds for all integers m, n and β. (If the simpler solution $A_n(t) = \widehat{A}(t+nT)$ had existed, it would also satisfy (3.9*a*) but with $\sigma_{m,n} \equiv 1$ for all m and n. We believe that no such solutions exist.) The fact that the expression (3.8) leads to terms with either sign in the nonlinear part of (3.9*a*) appears to be the essential ingredient necessary for non-trivial solutions satisfying the initial condition

$$\widehat{A}(t) \to 0 \quad \text{as} \quad t \downarrow -\infty. \tag{3.9c}$$

Since all the coefficients in (3.9*a*) are real, we restrict attention to real solutions $\widehat{A}(t)$. As a result a and $\partial a/\partial x$ are real and pure imaginary on the symmetry axis $x = 0$, where they take the values

$$a(0,t) = \sum_{\forall n} \mu_n \widehat{A}(t+nT) \quad \text{and} \quad \frac{\partial a}{\partial x}(0,t) = \mathrm{i} \sum_{\forall n} \mu_n(t+nT)\widehat{A}(t+nT). \tag{3.10a, b}$$

Additionally, we introduce $a_{\max}(x) \equiv \max\{|a(x,t)|\}$, where the maximization at given x is taken over the half-period $2T$; remember that $a(0, t+2T) = -a(0,t)$. Whether or not $a_{\max}(0)$ is the maximum amplitude over all x and t is unclear; nevertheless, it certainly provides a useful measure of the amplitude of our solution.

In Appendix B we show that $\widehat{A}(t)$ in the expansion (3.8) is linked to the pulse amplitude $\overline{a}(x)$ in (2.23) by

$$\widehat{A}(t) = \frac{1}{\sqrt{2}L}\,\mathscr{F}\{\overline{a}\}(t) \equiv \frac{1}{\sqrt{2}L}\int_{-\infty}^{\infty} \overline{a}(x)\exp(-\mathrm{i}tx)\,\mathrm{d}x, \tag{3.11a}$$

where we have introduced the Fourier transform $\mathscr{F}$. Use of the inverse Fourier transform $\mathscr{F}^{-1}$ leads us to the result

$$\overline{a}(x) = \sqrt{2}L\,\mathscr{F}^{-1}\{\widehat{A}\}(x) \equiv \frac{1}{\sqrt{2}T}\int_{-\infty}^{\infty} \widehat{A}(t)\exp(\mathrm{i}tx)\,\mathrm{d}t. \tag{3.11b}$$

Having made the choice that $\widehat{A}(t)$ is real, the symmetry property $\overline{a}(-x) = \overline{a}^*(x)$ (see (2.23*b*)) follows immediately. These Fourier transform results are remarkable in that they originate from the Fourier series (A 2*a*) in x (periodicity length $L/2 = 2\pi/4T$) and (A 2*b*) in t (periodicity time $4T = 4\pi/L$). The link established in Appendix A, upon which Appendix B builds, is non-trivial. Many of the subtle and elegant results stem from the factor $\exp(\mathrm{i}tx)$ in (3.1) and the necessity of superimposing the four distinct structures of which (A 1*a*) is composed.

We comment briefly on the mean energy (3.3*a*) introduced in § 3.1. Upon substitution of the ansatz (3.7) it reduces to

$$\mathscr{E}(t) = \sum_{\forall n} |\widehat{A}(t+nT)|^2 \tag{3.12}$$

with period T. Its time average

$$\langle \mathscr{E} \rangle \equiv \frac{1}{T}\int_{-T/2}^{T/2} \mathscr{E}\,\mathrm{d}t \;=\; \frac{1}{T}\int_{-\infty}^{\infty} |\widehat{A}(t)|^2\,\mathrm{d}t \tag{3.13a}$$

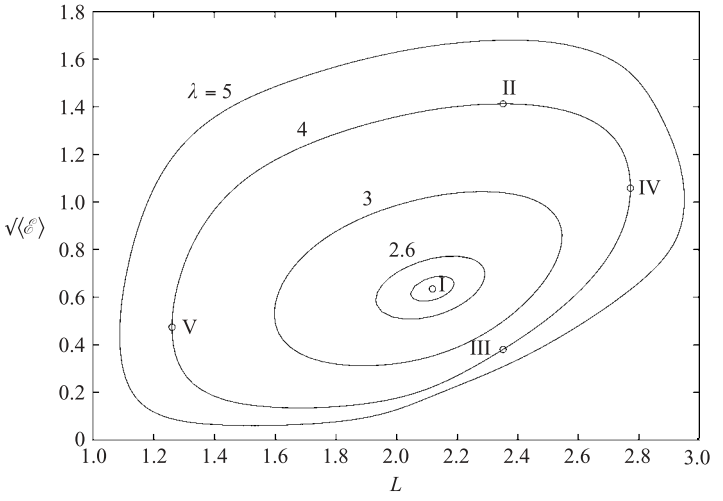

FIGURE 1. Contours of $\sqrt{\langle \mathscr{E} \rangle}$ *vs.* the pulse spacing L for the five values $\lambda = 2.55$ (innermost curve) 2.6, 3, 4 and 5 (outermost). Particular parameter combinations are identified by small circles and are labelled I–V for future reference.

gives the space–time average of $|a|^2$. An alternative representation is obtained by expressing a in the form (2.23*a*) and performing the time integration first. This yields

$$\langle \mathscr{E} \rangle \equiv \frac{1}{2\pi} \int_{-L/2}^{L/2} \int_{-T}^{T} |a(x,t)|^2 \, \mathrm{d}t \, \mathrm{d}x = \frac{1}{L} \int_{-\infty}^{\infty} |\overline{a}(x)|^2 \, \mathrm{d}x, \tag{3.13b}$$

which, in view of Parseval's theorem for Fourier transforms, is consistent with (3.13*a*).

Finally we note that the inequality (3.6*a*) limits the energy of our periodic solutions to the range

$$\langle \mathscr{E} \rangle + \frac{T^2}{12} \leqslant \lambda. \tag{3.14}$$

3.3. *The nature of the solutions*

Once we had identified a periodic structure by numerical solution of our initial value problem we investigated (3.9*a*), not on the infinite interval $-\infty < t < \infty$ but, rather, over the quarter-period T. To that end we introduced $B_n(t) = \widehat{A}(t + nT) = \mu_n A_n(t)$ and solved (3.9*a*) over the interval $0 \leqslant t \leqslant T$ subject to the periodicity conditions $B_n(0) = B_{n-1}(T)$ with, of course, $B_n(t) \to 0$ as $n \downarrow -\infty$. This numerical work was executed using the AUTO package (Doedel *et al.* 1997) on a finite set of $2N + 1$ equations based on the approximation that $B_n(t) \equiv 0$ on the interval for $|n| > N$, where N was chosen to be sufficiently large to ensure convergence.

The contours on figure 1 illustrate the value of the square-root of the mean energy $\langle \mathscr{E} \rangle$ as a function of the distance $L = \pi/T$ between pulses at various values of λ. Significantly, all solutions are of finite amplitude. That is easily understood because the linear solution of (3.9*a*) is proportional to $\exp(\lambda t - t^3/3)$ and so diverges as $t \downarrow -\infty$; hence it cannot satisfy the initial condition. The contours evaporate at $L_{\inf} \approx 2.11831$ and $\langle \mathscr{E} \rangle \approx 0.40251$ when $\lambda_{\inf} = 2.54074$ as noted in (2.24*b*). This constitutes the lowest value of λ at which a pulse-train solution could be found. Since this solution is only just activated and then only at the particular value $T = T_{\inf} \equiv \pi/L_{\inf} \approx 1.483$, it is of some interest to investigate some of its properties.

In figure 2(*a*) we show the amplitudes $A_n(t)$ (which, recall, satisfy (3.7*a*)) for a near-minimizing solution over roughly two periods (that is, $8T$). We notice that at any particular instant only three (or, at worst, perhaps four) amplitudes are non-zero to graph-plotting accuracy. This suggests that the nature of the solution might be amenable to understanding within the framework of a highly truncated system, which only involves the $m=n=0$ nonlinear term in (3.9*a*):

$$\frac{\mathrm{d}\widehat{A}}{\mathrm{d}t} = (\lambda - t^2 - |\widehat{A}|^2)\widehat{A}. \tag{3.15}$$

The amplitude of the general solution is given by

$$|\widehat{A}|^2 = \exp[2(\lambda t - t^3/3)] \Big/ \left\{2\int_{t_1}^{t} \exp[2(\lambda\tau - \tau^3/3)]\mathrm{d}\tau\right\}, \tag{3.16}$$

which exists on $t_1 < t < \infty$ where t_1 is an arbitrary constant. The solution satisfies $\widehat{A} \to 0$ as $t \to \infty$ as required but diverges to ∞, not as $t \downarrow -\infty$ (the behaviour of the linear solution) but rather as $t \downarrow t_1$. Accordingly this is not an acceptable solution. Since our solutions appear to depend on the role of the coefficients μ_n, it is hardly surprising that this severe truncation, which fails to capture that dependence, is inadequate. Indeed we believe that the alternating signs generated by the μ_n are crucial for the existence of our pulse-train solutions. Furthermore the most likely candidate for the maintenance of the pulse-trains is the nonlinear term $[\widehat{A}(T+t)]^2\widehat{A}^*(2T+t)$, associated with $\sigma_{1,1}=-1$ $(m=n=1)$, which may act as a nonlinear trigger for $\widehat{A}(t)$ when t lies roughly in the range $-2T \leqslant t \leqslant -T$.

The physical nature of our solution is described by the pulse-train (2.23), which combines pulses $\overline{a}(x \mp (2n+\frac{1}{2})L)$ each of frequency $\pm(2n+\frac{1}{2})L$. In figure 2(*b*) we illustrate the amplitude $|\overline{a}(x)|$ for our minimizing solution as well as its real and imaginary parts which possess symmetries $\mathrm{Re}\{\overline{a}(-x)\} = \mathrm{Re}\{\overline{a}(x)\}$ and $\mathrm{Im}\{\overline{a}(-x)\} = -\mathrm{Im}\{\overline{a}(x)\}$. (The graph only shows the one pulse centred at $x=0$.) Relative to this origin the pulse-train consists of that pulse together with neighbours possessing centres at $x=nL$ (integer n) and frequencies that differ by L. This means that the amplitude $a(x,t)$ with pulse centres at $x=nL+L/2$ defined by (2.23) behaves in a complicated way. We do, however, give a time-series plot of $a(0,t)$ (which is real valued, see (3.10*a*)) on figure 2(*c*), because we anticipate the interaction between the pulses centred at $x=-L/2$ and $x=L/2$ to be greatest there. Further, on figure 2(*d*) we portray phase-plane plots of $|\partial a/\partial x|$ against $|a|$ corresponding to the three locations $x=0$, $L/4$ and $L/2$. As time t proceeds the closed loops are followed in an anti-clockwise sense and completed in the half-period $2T$. The smallest loop occurs at the pulse centre $x=L/2$, where interaction with its neighbours is minimized and, in contrast, the largest loop occurs half-way between two pulses at $x=0$. (Notice that this loop appears to possess peculiar corners on the axes but this is merely a reflection of the fact that a is real and $\partial a/\partial x$ is pure imaginary; see (3.10) and figure 2(*c*).) It is remarked that contours are unchanged under the shift $x \to x+L$ and the reflection $x \to -x$, which is why we can take our representative sample within the range $0 \leqslant x \leqslant L/2$.

It is not easy to understand the mathematical properties of our pulses by thinking of them as solutions of (2.25*a*); the indirect approach involving the Fourier transform $\widehat{A}(t)$ just discussed is far more transparent. Nevertheless there are some pertinent observations that we can make concerning the necessary interaction of a pulse with its neighbours. Suppose the pulses are non-interacting, which would formally be the

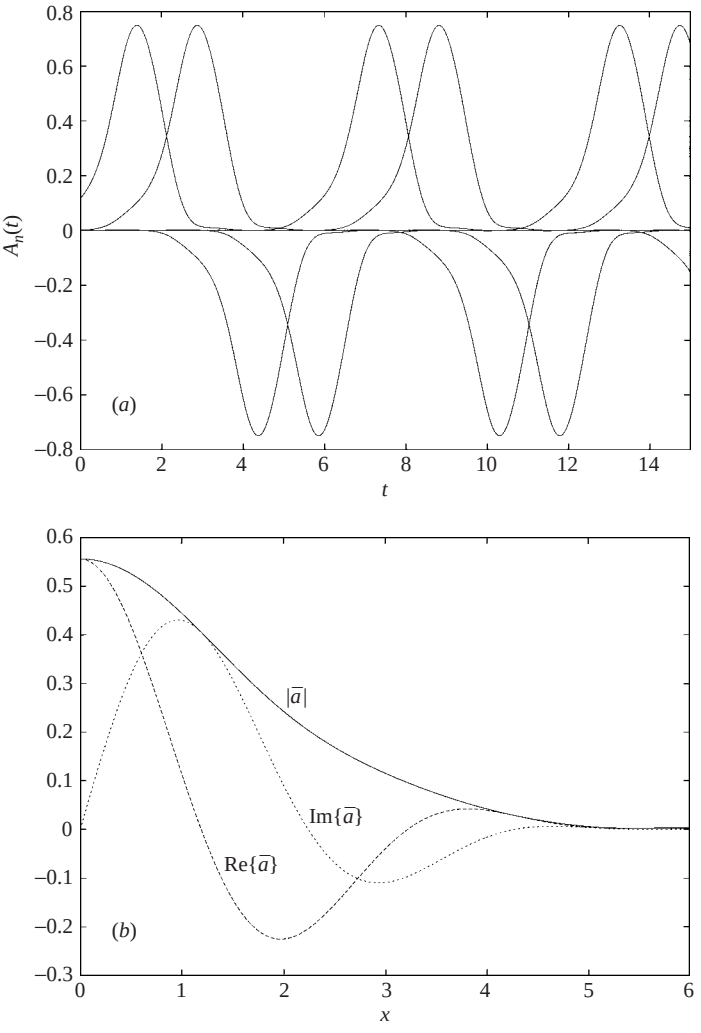

FIGURE 2(*a*, *b*). For caption see facing page.

case in the limit $L \to \infty$; then (2.25*a*) reduces to

$$\frac{\mathrm{d}^2\overline{a}}{\mathrm{d}x^2} + (\lambda + \mathrm{i}x - |\overline{a}|^2)\overline{a} = 0. \tag{3.17}$$

Hocking (private communication, *circa* 1980) reported attempting to seek localized solutions of (3.17) satisfying $|\overline{a}| \to 0$ as $|x| \to \infty$ but could find none. Later calculations by Ewen & Soward (1994*a*) also suggested that this equation admits no localized solutions. This gives a very interesting situation in which pulses can exist within an interacting train but not in isolation. Neighbour interactions play a crucial role and are necessary for the existence of the train. Moreover the phase of each pulse is carefully tuned to increase successively by $\pi/2$ and then decrease by $-\pi/2$ as indicated by the series (2.23*a*), a property linked to the alternating signs of the pairs $(\mu_{2n}, \mu_{2n+1}) = (-1)^n(1, 1)$ for integer n.

The remarks concerning the necessary interaction of $A_n(t)$ with its neighbour explains why, at given λ, there is a maximum value of T and corresponding minimum

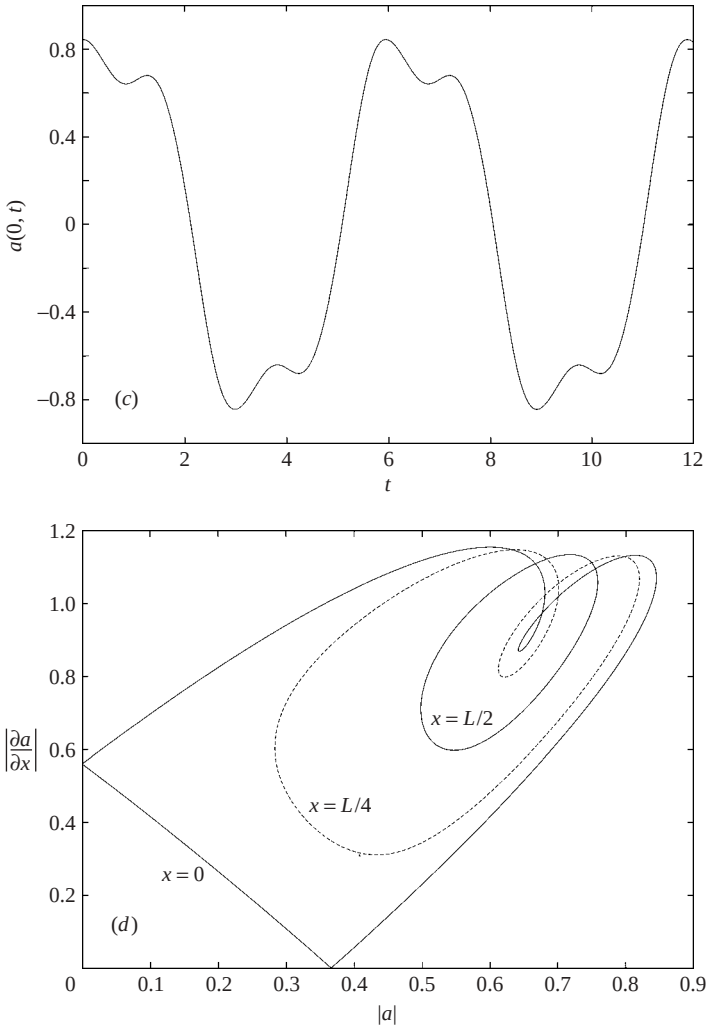

FIGURE 2. The solution at the point labelled I on figure 1, close to $\lambda_{\inf}$. Here $\lambda = 2.541$, $(L,\ \sqrt{\langle \mathscr{E} \rangle}) \approx (2.118,\ 0.63443)$ with $T \approx 1.48328$. (*a*) The functions $A_n(t)$ portrayed for successive integer n and separated by time T. (*b*) The pulse $\overline{a}(x)$ determined from the functions $A_n(t)$ using (3.7*a*) and (3.11*b*). The modulus $|\overline{a}(x)|$, real part $\mathrm{Re}\{\overline{a}(x)\}$ and imaginary part $\mathrm{Im}\{\overline{a}(x)\}$ are identified by the continuous and long and short dashed lines respectively; $\overline{a}(0) \approx 0.55635$. (*c*) The time series of $a(0,t)$ (constructed using (3.10*a*)); this function is of period $4T$. (*d*) The phase portraits of $|\partial a/\partial x(x,t)|$ *vs.* $|a(x,t)|$ for the cases $x = 0$, $L/4$ and $L/2$. The loops are followed in an anti-clockwise sense and are completed in the half-period $2T$.

value of L, namely $L_{\min}$, at which pulse-trains can persist. Likewise, the argument just given concerning the interaction of neighbouring pulses explains why there is a maximum possible value of L, namely $L_{\max}$. It is however important to appreciate that pulse neighbours are not identified by neighbouring $A_n(t)$ despite the suggestive terminology. This aspect is highlighted by the fact that our solutions exist only on some limited range $L_{\min}(\lambda) \leqslant L \leqslant L_{\max}(\lambda)$, as evident on figure 1, being terminated at small L by the large T separation of the A_n and at large L by pulse separation.

Interestingly, inside the range $L_{\min}(\lambda) < L < L_{\max}(\lambda)$ there are two solutions, one of small amplitude and the other large; see figure 1. The small-amplitude solution is

presumably unstable, while the stability of the large-amplitude solution is evidently a complicated matter outside the scope of our present analysis. We simply remark that for disturbances within the spatial periodicity class (3.1) of wavelength $2L$, the large-amplitude solutions that we report for moderate λ are likely to be stable. Certainly the solution of our initial value problem had no difficulty locking onto them. The question of sideband instability to disturbances of different wavelength is a far more difficult issue to address. Nevertheless, we anticipate that the L which maximizes the mean energy provides an excellent candidate for stability against sideband disturbances and so we focus on such solutions in the following section, particularly § 4.1.

4. Two limiting cases

In this section we apply our results to two limiting problems. In § 4.1 we consider the case of almost co-rotation $\varepsilon^{1/2} \ll \delta \ll 1$ and compare our predictions with HBS's results for $\delta = O(\varepsilon^{1/2})$. In § 4.2 we discuss the implication of our results for the situation in which the outer sphere is at rest $\delta = 1$. For the first time we show that finite-amplitude nonlinear solutions composed of pulse-trains may occur close to the local critical Taylor number T_c, subcritically to the Soward & Jones (1983) global critical value $T_{\rm crit}$.

The nature of our development is sensitive to the small value of the group velocity c. When $\varepsilon^{1/2} \ll \delta \ll 1$, the value of c is truly negligible and the apparatus set up in § 2 can be applied directly with $c = 0$. In contrast, when $\delta = 1$ or, more generally, when $\delta = O(1)$, c increases linearly with latitude. Thus although it is locally constant on the pulse length scale, c does vary on the longer length scale of the locally unstable region on which motion occurs. This means that the development in § 2, which assumes that c is constant, does not provide a uniformly valid solution on the longest length and time scales. The main purpose of § 4.2 is to show that with fairly minor modifications to the development in § 2 a uniformly valid solution can be constructed, which continues to be characterized by the pulse train (2.23a) but now with L dependent on t rather than just a constant.

4.1. *Almost co-rotation:* $\varepsilon^{1/2} \ll \delta \ll 1$

We begin by recapping the parameter values appropriate to the case of small δ, when the spheres almost co-rotate. To that end, we regard the frequency ω as a function not only of θ, k and T but now also of δ. So, in place of (2.8), we have

$$\omega = \omega(\theta,\ k,\ T,\ \delta). \tag{4.1}$$

Critical values in the narrow-gap cylinder limit are $\theta = 0$, $\delta = 0$, $k = k_0$, $T = T_0$, $\omega = \omega_0 = 0$, where

$$k_0 \approx 3.116, \quad T_0 \approx 1707.76. \tag{4.2a, b}$$

The symmetries of the problem ensure that

$$[\omega_{,\delta}]_0 = 0, \quad [\omega_{,\theta}]_0 = 0, \quad [\omega_{,\delta k}]_0 = 0, \quad [\omega_{,k\theta}]_0 = 0, \tag{4.3a, b, c, d}$$

where the zero subscript indicates that all quantities are evaluated at the cylinder critical values under the limit $\delta \downarrow 0$. About those values the local Taylor number and frequency have the expansions

$$T_l(\theta_l) = T_0 - \frac{[\omega_{,\delta\delta}]_0}{2\,[\omega_{,T}]_0}\,\delta^2 - \frac{[\omega_{,\theta\theta}]_0}{2\,[\omega_{,T}]_0}\,\theta_l^2, \qquad \omega_l(\theta_l) = [\omega_{,\delta\theta}]_0\,\delta\,\theta_l, \tag{4.4a, b, c}$$

in which

$$\frac{[\omega_{,\delta\delta}]_0}{T_0\,[\omega_{,T}]_0} \approx 0.0152, \quad \frac{\mathrm{i}\,[\omega_{,\delta\theta}]_0}{T_0\,[\omega_{,T}]_0} \approx -0.599, \quad \frac{[\omega_{,\theta\theta}]_0}{T_0\,[\omega_{,T}]_0} \approx -4.330, \qquad (4.5a, b, c)$$

together with the additional results

$$\frac{\mathrm{i}\,[\omega_{,\delta k\theta}]_0}{T_0\,[\omega_{,T}]_0} \approx 0.0553, \qquad \frac{[\omega_{,kk}]_0}{T_0\,[\omega_{,T}]_0} \approx -0.296, \quad -\mathrm{i}\,T_0\,[\omega_{,T}]_0 \approx 13.01. \qquad (4.5d, e, f)$$

Correct to leading order these determine

$$\gamma = \frac{\{-[\omega_{,kk}]_0[\omega_{,\theta\theta}]_0\}^{1/2}}{(-2\,[\omega_{,\delta\theta}]_0\,\delta)} \approx \frac{0.945}{\delta}, \qquad \mu^3 = \frac{[\omega_{,kk}]_0}{2\mathrm{i}\,[\omega_{,\delta\theta}]_0\,\delta} \approx \frac{0.247}{\delta} \qquad (4.6a, b)$$

(see (2.19)).

For small δ it is also appropriate to reinstate the $O(\theta^2)$ term in (2.17) which then becomes

$$\frac{\partial \mathscr{A}}{\partial t} + c_{,\chi}\chi\frac{\partial \mathscr{A}}{\partial \chi} = [\lambda(\chi) + \mathrm{i}\chi - |\mathscr{A}|^2]\mathscr{A} + \frac{\partial^2 \mathscr{A}}{\partial \chi^2}, \qquad \lambda(\chi) \equiv \lambda(0) - \Upsilon_\varepsilon^2\,\chi^2, \qquad (4.7a, b)$$

where χ is measured from the equator (i.e. $\theta = \mu\varepsilon^{2/3}\chi$ or, equivalently, $\chi_l = 0$ in (2.20a)) and

$$c_{,\chi} = \left(-\frac{c_g}{c_p}\right)\frac{\Upsilon_\varepsilon}{k_0\gamma} \approx 0.147\,(\delta\varepsilon)^{1/3}, \qquad \frac{c_g}{c_p} = \frac{k_0[\omega_{,\delta k\theta}]_0}{[\omega_{,\delta\theta}]_0} \approx -0.288, \qquad (4.7c, d)$$

$$\Upsilon_\varepsilon = \varepsilon^{1/3}\,\gamma/\mu \approx 1.506\,(\varepsilon/\delta^2)^{1/3}. \qquad (4.7e)$$

Most of these quoted values were determined by Soward & Jones (1983) although the size of $[\omega_{,T}]_0$ was given by Davey (1962). We recalculated and checked these results (as well as computing $[\omega_{,\delta k\theta}]_0$) by direct numerical differentiation.

There are a number of features of equation (4.7a) that are worthy of comment. Significantly, the range

$$|\chi| < \sqrt{\lambda(0)}/\Upsilon_\varepsilon = O\big(\delta^{2/3}/\varepsilon^{1/3}\big) \qquad (4.8)$$

on which the system is locally unstable ($\lambda(\chi) > 0$) decreases in width with decreasing δ. Our pulse-train analysis of § 3 was based on the idea that the local value of $\lambda(\chi)$ is approximately constant over the length scale of pulses. Strictly, that approximation is only valid when $\Upsilon_\varepsilon \ll 1$, in contrast to the assumptions underpinning HBS who considered the case $\Upsilon_\varepsilon = O(1)$. Furthermore, on the relatively long length scale (4.8) of the locally unstable region, the group velocity $c_{,\chi}\chi = O(\delta)$ is negligible and so (4.7a, b) reduces to

$$\frac{\partial \mathscr{A}}{\partial t} = \big(\lambda + \mathrm{i}\chi - \Upsilon_\varepsilon^2\,\chi^2 - |\mathscr{A}|^2\big)\mathscr{A} + \frac{\partial^2 \mathscr{A}}{\partial \chi^2}, \qquad (4.9)$$

where now $\lambda \equiv \lambda(0)$, a constant as in (2.22). Nevertheless, the small but finite group velocity will lead to non-uniformities on a relatively long time scale, which are not captured by (4.9). Indeed (4.19) below shows that the pulse centres drift apart causing the pulse centre separation L to increase very slowly with time $L(t) = L(0)\exp(c_{,\chi}\,t)$ (see (4.20b)). For $\delta \ll 1$ this is a small perturbation of little importance and so we postpone discussing the implications of $c_{,\chi} \neq 0$, until we consider the case $\delta = O(1)$ in the next subsection. With the drift ignored $c_{,\chi} \approx 0$, we have $L(t) = L$ a constant. Accordingly we can make sensible comparisons with the pulse-train-like solutions

λ	$\mathscr{A}_{\max}$	N	T	L	$-2\Omega/L$
3.375	1.3897	46	1.3585	2.3126	—
3.875	1.6047	6	1.2967	2.4228	0.5226
3.975	*1.4311*	*10*	*1.2800*	*2.4544*	*0*
4.0	1.6248	9	1.2889	2.4375	0.6647
4.5	1.7910	67	1.2600	2.4933	0.8357

TABLE 1. The characteristics of HBS's numerical solutions to equation (4.9) for $\Upsilon_\varepsilon = 1/4$ at various values of λ. The initial pitchfork bifurcation from the undisturbed basic state occurs at $\lambda = \lambda_p = 4.25$. The interpretations of the various properties are given in the text below (4.11). Data for a symmetry-preserving case $\Omega = 0$ are written in italics.

obtained by HBS for numerically small values of Υ_ε, specifically $\Upsilon_\varepsilon = 1/4$, with our pulse-train solutions valid for $\Upsilon_\varepsilon \ll 1$. In practice, HBS non-dimensionalized the amplitude equation (2.16) in a different way to that done here and they considered (C1*a*) of Appendix C rather than (4.9). Here we describe the nature of HBS's solutions relative to our units, but quantify the relationship between the formulations in Appendix C for reference.

HBS considered the bifurcation sequence for various values of Υ_ε. For small Υ_ε the initial bifurcation is via a supercritical pitchfork at $\lambda = \lambda_p \equiv (2\Upsilon_\varepsilon)^{-2} + \Upsilon_\varepsilon$ to a steady finite-amplitude state $\mathscr{A} = \mathscr{A}_s$ with the symmetry property $\mathscr{A}_s(-\chi) = \mathscr{A}_s^*(\chi)$ (cf. (2.23*b*)). This state has extremely small amplitude, which immediately following the bifurcation is given by

$$\mathscr{A}_s(\chi) \approx 2^{1/4}(\lambda - \lambda_p)^{1/2} \exp\left[-1/\left(16\Upsilon_\varepsilon^3\right) - \tfrac{1}{2}\Upsilon_\varepsilon \chi^2 + \tfrac{1}{2}\mathrm{i}\chi/\Upsilon_\varepsilon\right] \quad \text{for} \quad 0 < \lambda - \lambda_p \ll \Upsilon_\varepsilon \tag{4.10}$$

(see § 3.1, figure 6 of Harris *et al.* 2000; also our Appendix C). That loses stability via a supercritical Hopf bifurcation (e.g. HBS figure 5 with $\Upsilon_\varepsilon = 1/2\sqrt{2}$). The associated non-zero temporal mean limit cycle (HBS figure 9) expands until it becomes a homoclinic cycle with the zero-amplitude state as its vertex. That homoclinic cycle glues with the corresponding oppositely signed cycle to leave a limit cycle with zero temporal mean (HBS figure 11), much like our pulse-train solution $\mathscr{A}_{\mathrm{SP}}(\chi, t) = a(\chi, t; L)$ (see (2.23*a*)). This periodic solution continues to exhibit the symmetry $\mathscr{A}_{\mathrm{SP}}(-\chi, t) = \mathscr{A}^*_{\mathrm{SP}}(\chi, t)$ (cf. (2.26)), which HBS referred to as a Symmetry-Preserving (SP) solution. Even for the moderate value $\Upsilon_\varepsilon = 1/4$, the solution is strongly subcritical, exists on a small-amplitude solution branch down to some lesser value of λ and then advances back along a large-amplitude solution branch (HBS figure 12). This is compatible with our figure 1 which shows that for each L on $L_{\min} < L < L_{\max}$ there is both a small- and a large-amplitude solution.

The picture painted by HBS was complicated by the fact that on returning along the initially stable upper branch the solution became unstable to a second frequency (also within the SP-class), which was sometimes found to lock in the sense of Arnold tongues. Such a five-fold increase in period was detected by HBS in the case $\Upsilon_\varepsilon = 1/4$ at $\lambda = 3.975$ and identified on HBS figure 12. Its character is illustrated on HBS figure 20, which is reproduced here relative to our units in figure 3 so that we can make direct comparison with our new results. The time series graphed in figure 3(*a*) identifies the underlying period $4T \approx 5.12$ (and the corresponding frequency $L/2 \approx 1.227$, see table 1). This is also the period of one cycle in the phase portrait figure 3(*b*), which is

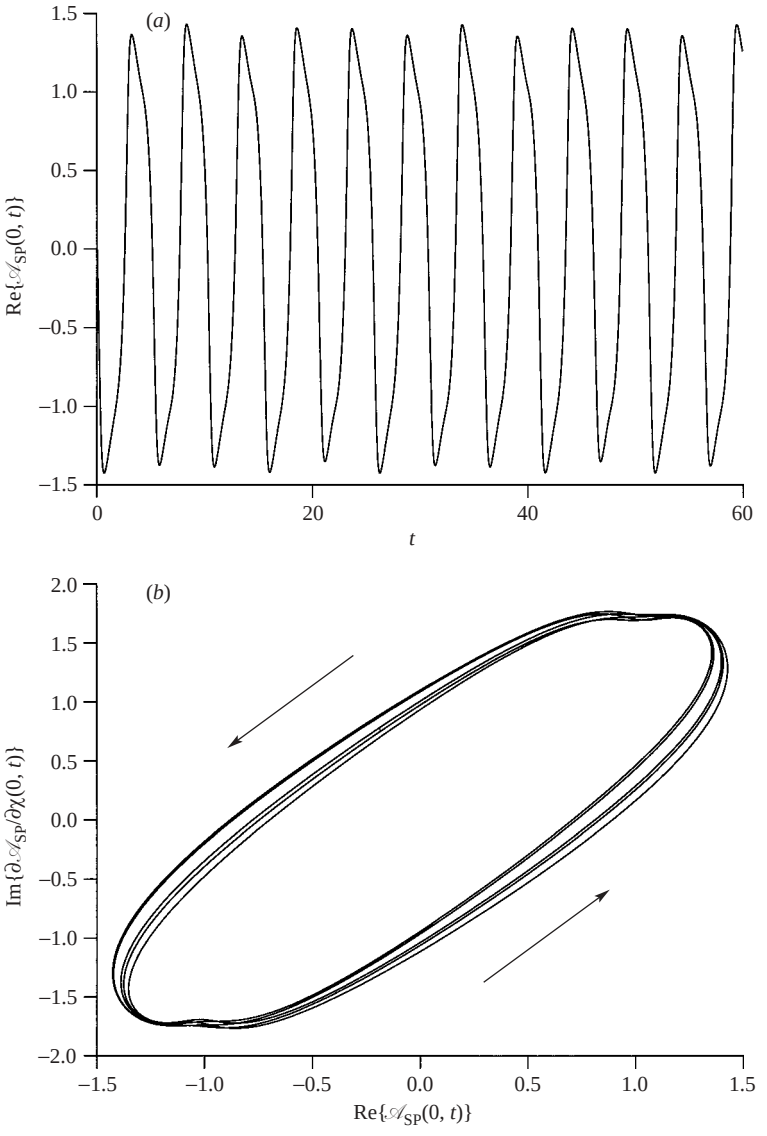

FIGURE 3. The symmetry-preserved (SP) $\Upsilon_\varepsilon = 1/4$ solution $\mathscr{A}_{SP}(\chi, t)$ of equation (4.9) at $\lambda = 3.975$ with $N = 10$ as identified in table 1. (The diagrams are based on figure 20 of HBS but recast in terms of our units.) (*a*) The time series of $\mathrm{Re}\{\mathscr{A}_{SP}(0, t)\}$ which is almost periodic on time $4T$ but is actually periodic over time $20T \approx 25.6$. (*b*) The phase portrait of $\mathrm{Im}\{\partial\mathscr{A}_{SP}/\partial\chi(0, t)\}$ against $\mathrm{Re}\{\mathscr{A}_{SP}(0, t)\}$. The five loops are followed in the anti-clockwise sense indicated.

followed in the anti-clockwise sense. Since five circuits are required before the solution repeats itself, the complete period is $20T$.

HBS also discovered that on solving (4.9) as an initial value problem the realized solution generally breaks the symmetry and exhibits the drifting phase Ω present in (2.21*a*). They deemed these Symmetry-Broken (SB) solutions and we link them to

$$\mathscr{A}_{SB}(\chi, t) = \exp(\mathrm{i}\Omega t)\, a(\chi - \Omega, t; L), \tag{4.11}$$

as defined by (2.21) and (2.23) with $c=\varphi=0$. For certain values of λ at $\Upsilon_\varepsilon=1/4$ HBS found relatively simple solutions, whose characteristics are listed in table 1. By this we mean that in addition to a frequency L and corresponding half-period $2T$ an extra beating period $2NT$ can be clearly identified over which the amplitude is modulated. (The relevant value of N is given in the third column.) In the second column, $\mathscr{A}_{\max}$ denotes the maximum of $|\mathscr{A}(\chi,t)|$ over t at $\chi=0$. In view of the drifting phase, it corresponds to our $a_{\max}(x)$ at $x=-\Omega$ and presumably underestimates the likely maximum value $a_{\max}(0)$. This shift of origin is important for comparison purposes and so we list $2x/L=-2\Omega/L$ rather than Ω in the last column.

Though HBS generally found very complicated structures unless the parameter values were carefully chosen as in table 1, the presence of an underlying frequency L was a robust feature of all their solutions and in some cases the drifting phase Ω could be determined. Finally we note that the SP-solution listed in italics was identified from within that SP-class. Though it was stable to SB-disturbances, it may not necessarily be the solution achieved following the time evolution from more general initial data.

From our small-Υ_ε point of view the pulse-train solution (4.11) is only the leading-order approximation to HBS's solutions. The value of Ω and any long-time-scale aperiodic behaviour such as the beating period are determined by a higher-order theory, which is influenced by the fact that the local value of $\lambda(\chi)$ varies spatially and is not a constant as assumed by our theory. Not only does that modulation lead to non-uniformities but the fact that no solution is possible for $\lambda(\chi)<\lambda_{\inf}$ (see (2.24*b*)) is also significant. In short, the pulse-train must terminate abruptly, dropping suddenly from a finite amplitude to zero in much the same way as at a front. Often the existence of fronts provide a mechanism for frequency selection but here our belief is that because of the robust nature of the wave-train, it maximizes its amplitude near $\chi=0$, so fixing the frequency. This viewpoint is supported by the frequency values L listed in table 1 which compare favourably with those associated with the maximum mean energies $\langle\mathscr{E}\rangle$ on figure 1.

Our large-amplitude solution for $\lambda=4$ (in the middle of the range of values itemized in table 1) at $L=2.35107$, where $\langle\mathscr{E}\rangle$ is maximized for this choice of λ, is illustrated in figure 4. There is considerable interaction between both neighbouring $A_n(t)$ (figure 4*a*) and pulses (see figure 4*b*). The characters of the time series for $\mathrm{Re}\{a(0,t)\}$ and of our phase portrait of $|\partial a/\partial x(0,t)|$ against $|a(0,t)|$ are illustrated in figures 4(*c*) and 4(*d*) respectively; both of these compare favourably with their counterparts for the SP-solution of equation (4.9) presented in figure 3, albeit after moduli are taken. Likewise the time series for $|a(L/4,t)|$ (also given on figure 4*c*) is in reasonable accord with the SB-solution time series for $|\mathscr{A}_{\mathrm{SB}}(0,t)|$, see figure 5(*a*), pertaining to the choices $\Upsilon_\varepsilon=1/4$ and $\lambda=3.875$ with $-\Omega\approx0.263\,L$ close to $L/4$. Similar good agreement is noted between the phase portrait of $|\partial a/\partial x(L/4,t)|$ against $|a(L/4,t)|$ in figure 4(*d*) and figure 5(*b*). An interesting feature is the kinks evident on both the phase portraits together with further plots given in HBS (their figures 22(*b*) and 23(*b*)). The space–time contour plot of $\mathrm{Re}\{\mathscr{A}_{\mathrm{SB}}(\chi,t)\}$ illustrated in figure 5(*c*) (based on HBS figure 21*e*) is important for it provides evidence of the pulse-train structure, though even more convincing evidence is provided by the diagnostic space–time contour plot of $|\mathscr{A}_{\mathrm{SB}}(\chi,t)|$ in HBS figure 21(*f*), which is not reproduced here. On figure 5(*c*) can be identified the frequencies of five (and only five) pulses centred at $\chi=\Omega+(n+1/2)L$ ($n=-2,\ -1,\ 0,\ 1,\ 2$), that is, at ≈-4.2673, -1.8444, 0.5783, 3.0011 and 5.4239. Four dislocations of the patterns occur half-way between the pulse centres at $\chi=\Omega+nL$ ($n=-1,\ 0,\ 1,\ 2$); only the outer two are shown unambiguously on the figure, while

the inner two are less definitive. These dislocations are close to the local maxima (except for the left-hand maximum, which is presumably a manifestation of non-uniform conditions at the end of the pulse train) and minima of $\mathscr{A}_{\rm SB}(\chi, t)$ at $t = 20$ graphed in figure 5(*d*).

We have already remarked that for fixed λ and L the evidence of figure 1 is that there are two possible pulse-trains. In figure 4 we showed the solution with the largest energy when $\lambda = 4$ (point II on figure 1) and for comparison the corresponding small-amplitude solution also with $L = 2.35107$ (point III on figure 1) is illustrated in figure 6. It is presumably unstable but does help us to understand the structure of possible solutions. We will see in the following subsection that the group velocity leads to pulse separation and a tendency for L to slowly increase with time so that the pulse-train character for larger L is of particular interest. Simply to illustrate the nature of the large- and small-L limits we plot the $A_n(t)$ and pulse profiles $\overline{a}(x)$ for $\lambda = 4$ at $L_{\max} \approx 2.77229$ in figure 7 and at $L_{\min} \approx 1.26040$ in figure 8 corresponding to the points IV and V respectively on figure 1. In the former case, with T relatively small, neighbouring $A_n(t)$ are evidently strongly interacting, whereas in the latter case they are not. Curiously, this interaction between neighbouring pulses is more marked when they are well separated with $L = L_{\max}$. Apparently, when they are close together with $L = L_{\min}$ the pulse profile become smoother. This must reflect the fact that the inverse Fourier transform $\overline{a}(x)$ of the smooth transform function $\widehat{A}(t)$ is itself smooth, though physically why $\overline{a}(x)$ should have this property is not obvious.

4.2. *Outer sphere at rest*

When the outer sphere is at rest, $\delta = 1$, we note that the cylinder critical values are

$$k_c \approx 3.1266, \qquad T_c \approx 1694.95. \tag{4.12a, b}$$

Significantly these values are smaller than the true critical values $k_{\rm crit} \approx 3.1769$ and $T_{\rm crit} \approx 1767.90$ (Soward & Jones 1983) for the onset of linear instability. From our calculations we obtain

$$-\mathrm{i}\, T_c\, [\omega_{,T}]_c \approx 13.10, \quad -[\omega_{,\theta}]_c \approx 7.693, \quad [\omega_{,k}]_c = 0, \tag{4.13a, b, c}$$

$$\mathrm{i}[\omega_{,\theta\theta}]_c \approx 53.4, \quad [\omega_{,k\theta}]_c \approx 0.69, \quad \mathrm{i}\,[\omega_{,kk}]_c \approx 3.83, \tag{4.13d, e, f}$$

where the value of $\mathrm{i}[\omega_{,T}]_c$ is in accord with Davey's (1962) result. Use of (2.19) and (2.20*c*) determines

$$\gamma \approx 0.929, \quad \mu^3 \approx 0.249, \quad \Upsilon_\varepsilon \approx 1.478\ \varepsilon^{1/3}, \tag{4.14a, b, c}$$

where our value of μ agrees with Walton's (1978) value $\mu^3 \approx 1/4.0199 \approx 0.2488$. In turn, these determine the coefficients (2.18*d*, *e*) in (2.17), which become

$$\lambda \approx 2.71\ \varepsilon^{-2/3}\, \frac{T - T_c}{T_c} - \Theta_l^2, \quad c \approx 0.096\ \Theta_l. \tag{4.15a, b}$$

All the scale estimates made in the previous section continue to hold, but simplify in the sense that $\delta = O(1)$.

As remarked in § 4.1, the region of local instability ($\lambda > 0$) expands with increasing δ and when $\delta = O(1)$ it has width $O(\varepsilon^{-1/3})$ relative to the inner χ-units (see (4.8)). That is why in § 2 we developed local solutions $\mathscr{A}$ (recall (2.21)) valid on the θ-length scale $O(\varepsilon^{2/3})$ in the neighbourhood of some $\theta = \theta_l = O(\varepsilon^{1/3})$. These pulse-train solutions drift slowly away from $\theta = 0$ at the positive group velocity c defined by (4.15*b*). Since c increases linearly with distance from the equator, our pulse solutions exhibit spatio-temporal non-uniformities on the long length $\chi = O(\varepsilon^{-1/3})$ of the locally unstable

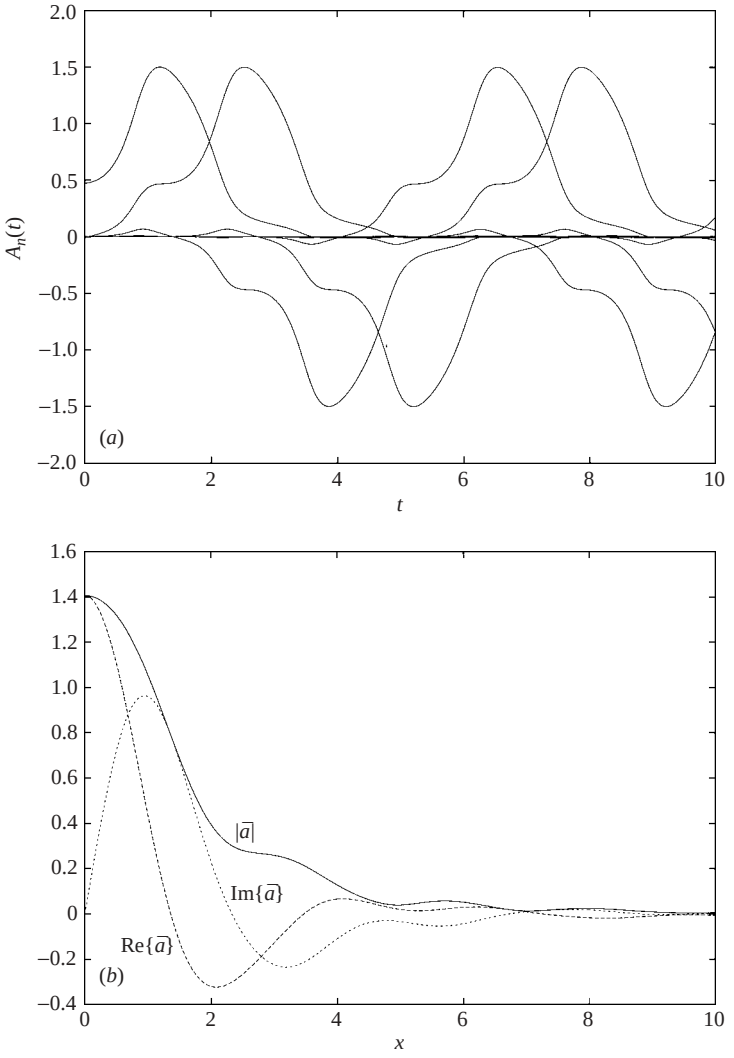

FIGURE 4(*a*, *b*). For caption see facing page.

region and the associated long time $t = O(\varepsilon^{-1/3})$. Remarkably, our pulse structure can be embedded within a spatially uniform representation and to demonstrate this we replace (2.15) with

$$\begin{bmatrix} \widehat{\psi} \\ \widehat{h} \end{bmatrix} = \exp\left(\frac{\mathrm{i}k_c\theta}{\varepsilon}\right) \sum_{n\in\mathscr{S}} \begin{bmatrix} -\mathrm{i}f_n(z) \\ g_n(z) \end{bmatrix} b_n(\theta-\theta_n,\tau)\exp\left\{\mathrm{i}\left[\frac{(-1)^n\pi}{4}+\varphi_n(\tau)\right]\right\} + \text{c.c.}, \tag{4.16a}$$

where

$$\varphi_n(\tau) = -\int_{\tau_0}^{\tau} [\omega_{,\theta}]_c \theta_n(\tau)\,\mathrm{d}\tau. \tag{4.16b}$$

Here $(f_n(z),\ g_n(z))$ represents the eigenfunction at the centre $\theta = \theta_n(\tau)$ of each individual pulse while the amplitude $b_n(\theta-\theta_n,\tau)$ is measured in the frame relative to its moving centre. Furthermore, the summation is taken over the set $\mathscr{S}$ of integers

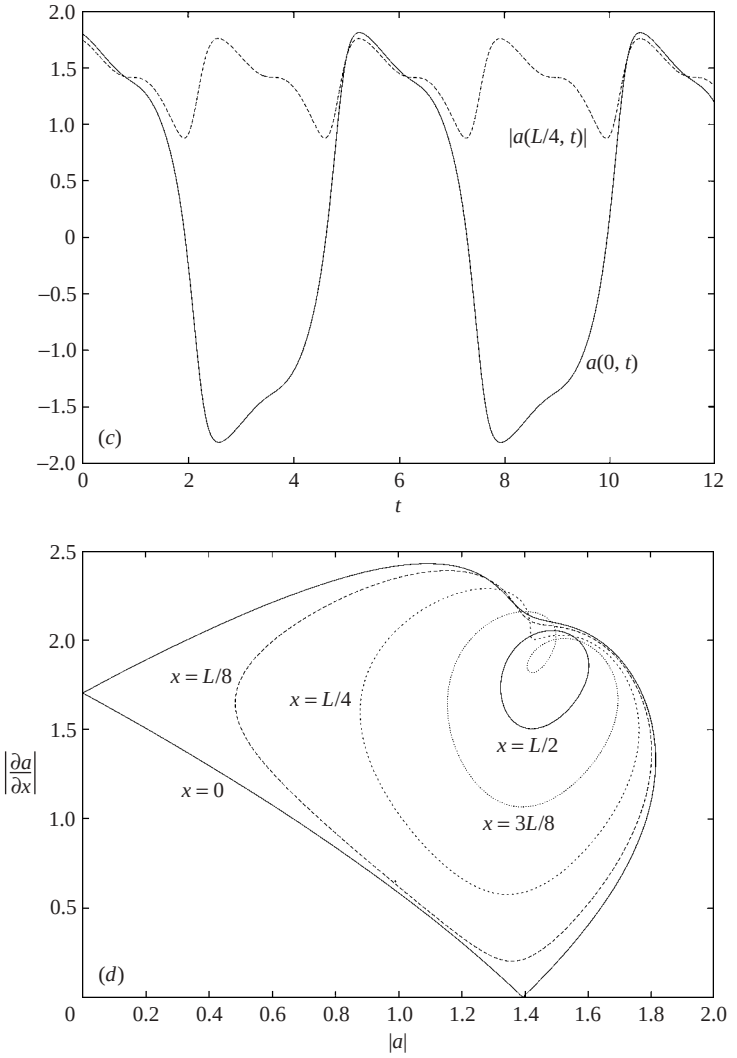

FIGURE 4. The solution at the point labelled II on figure 1 with $\lambda = 4$, $(L, \sqrt{\langle \mathscr{E} \rangle}) \approx (2.35107, 1.41246)$ and $T \approx 1.33625$. Subplots (a)–(d) correspond to their equivalents for figure 2 (with $\overline{a}(0) \approx 1.4069$) except that (c) shows the time series for both $a(0, t)$ (solid line) and $|a(L/4, t)|$ (dashed) and (d) illustrates the phase portraits for the five cases $x = nL/8$ with $n = 0, 1, \ldots, 4$.

n of order $\varepsilon^{-1/3}$ such that θ_n lies in the locally unstable region and τ_0 is some time origin independent of n. Then, under the change of variables (2.18), each pulse has the structure

$$\sqrt{\mathscr{C}}\, b_n(\theta - \theta_n) = (-[\omega_{,\theta}]_c\, \mu)^{1/2}\, \varepsilon^{1/3}\, \overline{a}(x; L, \lambda), \tag{4.17a}$$

in which $\overline{a}(x; L, \lambda)$ has the symmetry-preserving property $\overline{a}(-x) = \overline{a}^*(x)$ and solves (2.25a) as before. Relative to the scaled variables

$$x = \chi - \chi_n(t), \qquad \theta = \mu\varepsilon^{2/3}\chi, \qquad \theta_n = \mu\varepsilon^{2/3}\chi_n(t) \tag{4.17b, c, d}$$

(cf. (2.18)) the pulse separation and phase differences are

$$L(t) \equiv \chi_{n+1}(t) - \chi_n(t), \qquad \Phi(t) \equiv \varphi_{n+1}(t) - \varphi_n(t). \tag{4.17e, f}$$

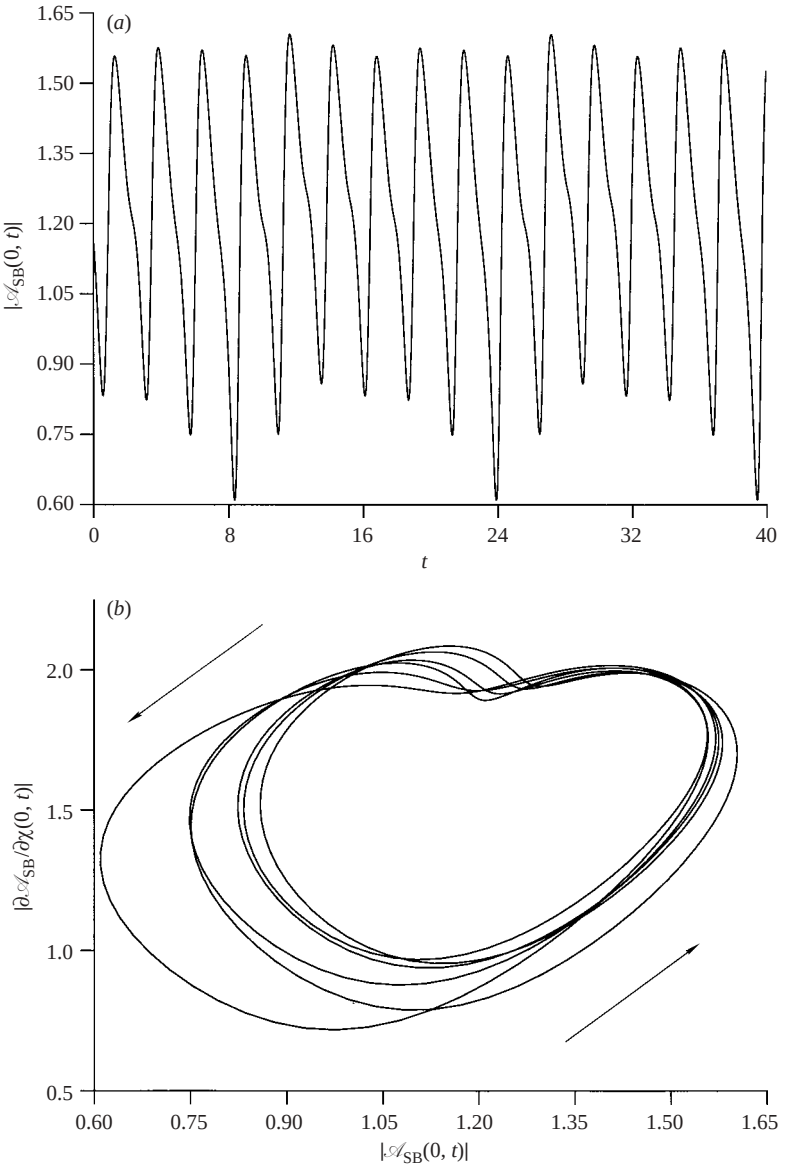

FIGURE 5(a, b). For caption see facing page.

Despite the fact that the pulse centres $\chi_n(t)$ drift polewards with time, both $L(t)$ and $\Phi(t)$ turn out to be independent of n (see (4.20) and (4.21)). This important property permits the solutions $b_n(\theta - \theta_n)$ proportional to $\overline{a}$ to remain uniformly valid as time proceeds. Strictly, the value of λ for each pulse b_n is given by the local value (4.15a) at $\theta_n(t)$, namely

$$\lambda_n(t) = \lambda(\chi_n(t)) \tag{4.18}$$

(see (4.7b)). This leads to the spatial non-uniformities already discussed in § 4.1 and whose role we consider later.

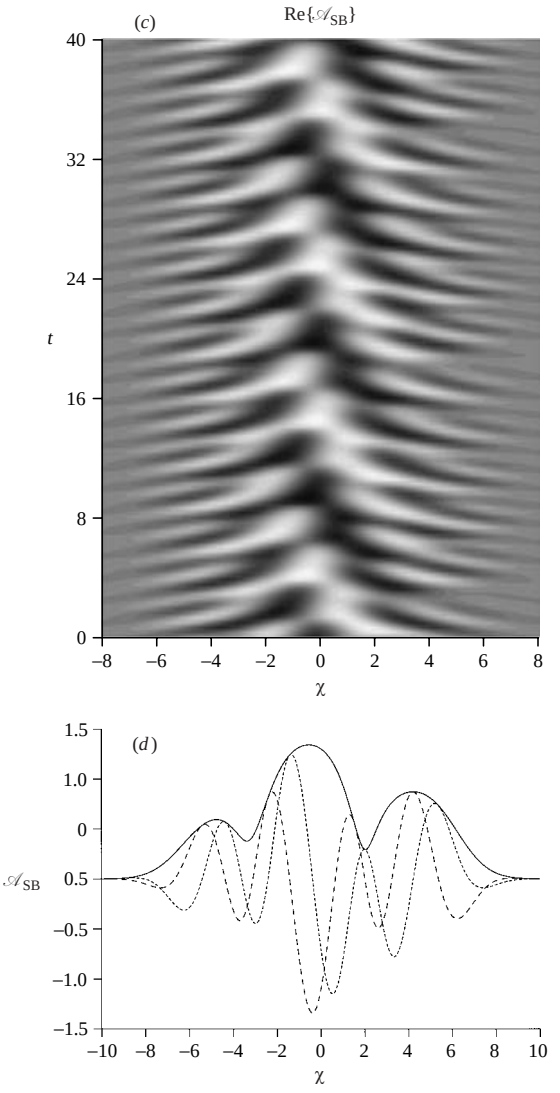

FIGURE 5. The symmetry-broken (SB) $\Upsilon_\varepsilon = 1/4$ solution of equation (4.9) for $\lambda = 3.875$ with $L \approx 2.423$, $\Omega \approx -0.633$ and $N = 6$ as identified in table 1. (Based on figure 21 of HBS but re-expressed in our units.) (*a*) The time series of $|\mathscr{A}_{SB}(0, t)|$. (*b*) The phase portrait of $|\partial\mathscr{A}_{SB}/\partial\chi(0, t)|$ *vs.* $|\mathscr{A}_{SB}(0, t)|$ followed in the anti-clockwise sense indicated. (*c*) Contours of constant $\mathrm{Re}\{\mathscr{A}_{SB}(\chi, t)\}$ with four dislocations at $\chi = \Omega + nL$ ($n = -1, 0, 1, 2$) (i.e. $\approx -3.056, -0.633, 1.790, 4.213$). (*d*) The magnitudes $|\mathscr{A}_{SB}|$ and $\mathrm{Re}\{\mathscr{A}_{SB}\}$ ($\mathrm{Im}\{\mathscr{A}_{SB}\}$) at time $t = 20$ identified by the continuous and long (short) dashed curves respectively.

The movement of $\chi_n(t)$ at the group velocity c is governed by

$$\frac{\mathrm{d}\chi_n}{\mathrm{d}t} = c_{,\chi}\chi_n, \tag{4.19a}$$

where

$$c_{,\chi} = \left(-\frac{c_g}{c_p}\right)\frac{\Upsilon_\varepsilon}{k_0\gamma} \approx 0.1419\,\varepsilon^{1/3}, \qquad \frac{c_g}{c_p} = \frac{k_c[\omega_{,k\theta}]_c}{[\omega_{,\theta}]_c} \approx -0.28 \tag{4.19b, c}$$

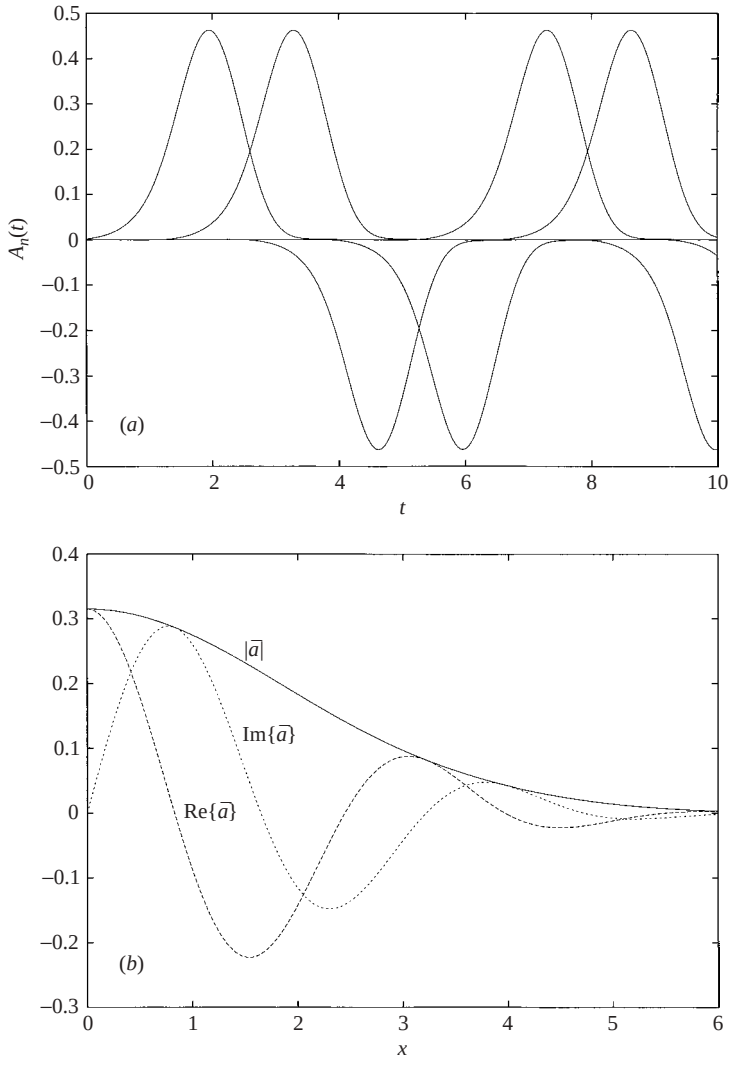

FIGURE 6(*a*, *b*). For caption see facing page.

and this solution determines

$$\chi_n(t) = \chi_n(0)\,\exp(c_{,\chi}\,t), \quad L(t) = L(0)\,\exp(c_{,\chi}\,t). \tag{4.20a, b}$$

In other words, if the initial separation $L(0) = \chi_{n+1}(0) - \chi_n(0)$ is the same for all pulses, that continues to hold as time proceeds. Furthermore, since the frequency of each pulse is also $\chi_n(t)$, the phases are given by (4.16*b*), which following the change of variables (4.17) reduces to

$$\varphi_n(t) = \int_0^t \chi_n\,\mathrm{d}t = \frac{\chi_n(t) - \chi_n(0)}{c_{,\chi}} \qquad \text{giving} \qquad \Phi(t) = \frac{L(t)}{c_{,\chi}} \tag{4.21a, b}$$

independent of n. To see the connection with our earlier representations (2.21) for $\mathscr{A}(\chi, t)$ and (2.23) for $a(x, t)$, we note that

$$\varphi_n(t) \approx \chi_n(0)\,t + \tfrac{1}{2}\,\chi_n(0)\,c_{,\chi}\,t^2 \qquad \text{for} \quad t \ll \varepsilon^{-1/3}, \tag{4.22}$$

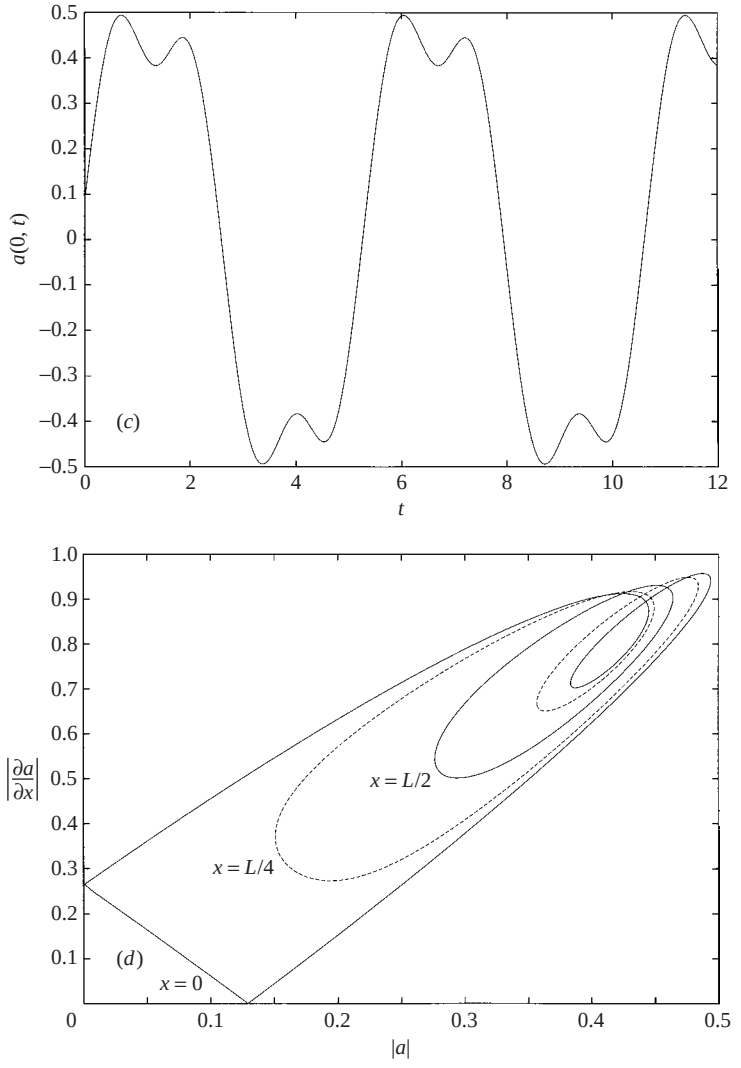

FIGURE 6. The solution at the point labelled III on figure 1. As in figure 4 we have $\lambda = 4$ and $L = 2.35107$ ($T = 1.33625$) but the energy is reduced to $\sqrt{\langle \mathscr{E} \rangle} \approx 0.38016$ and $\overline{a}(0) \approx 0.31503$. Sub-plots (*a*)–(*d*) correspond to their equivalents for figure 2.

where $\chi_n(0) = O(\varepsilon^{-1/3})$. We set $\chi_n(0) = \chi_l + \Omega + (n + \frac{1}{2})L(0)$, so that for $t = O(1)$ we have

$$\chi_n(t) - \chi_l = \Omega + \left(n + \tfrac{1}{2}\right)L(0) + \chi_l\, c_{,\chi}\, t + O\left(\varepsilon^{2/3} t^2\right), \tag{4.23a}$$

and

$$\varphi_n(t) - \chi_l\, t = \left[\Omega + \left(n + \tfrac{1}{2}\right)L(0)\right] t + \tfrac{1}{2}\,\chi_l\, c_{,\chi}\, t^2 + O\left(\varepsilon^{2/3} t^3\right), \tag{4.23b}$$

where $\chi_l = O(\varepsilon^{-1/3})$. We now obtain exact correspondence with (2.21) and (2.23), on setting $\varphi = 0$ and $t_0 = 0$ in (2.21) and identifying $\pm(2n + \frac{1}{2}) - \frac{1}{2}$ in (2.23*a*) with our n. These restrictions on φ and t_0 are not significant and non-zero values can be accommodated by transformations. Accordingly (2.21) and (2.23) gives the short $O(1)$-time-scale solution with $L = L(0)$ and $c = \chi_l c_{,\chi}$.

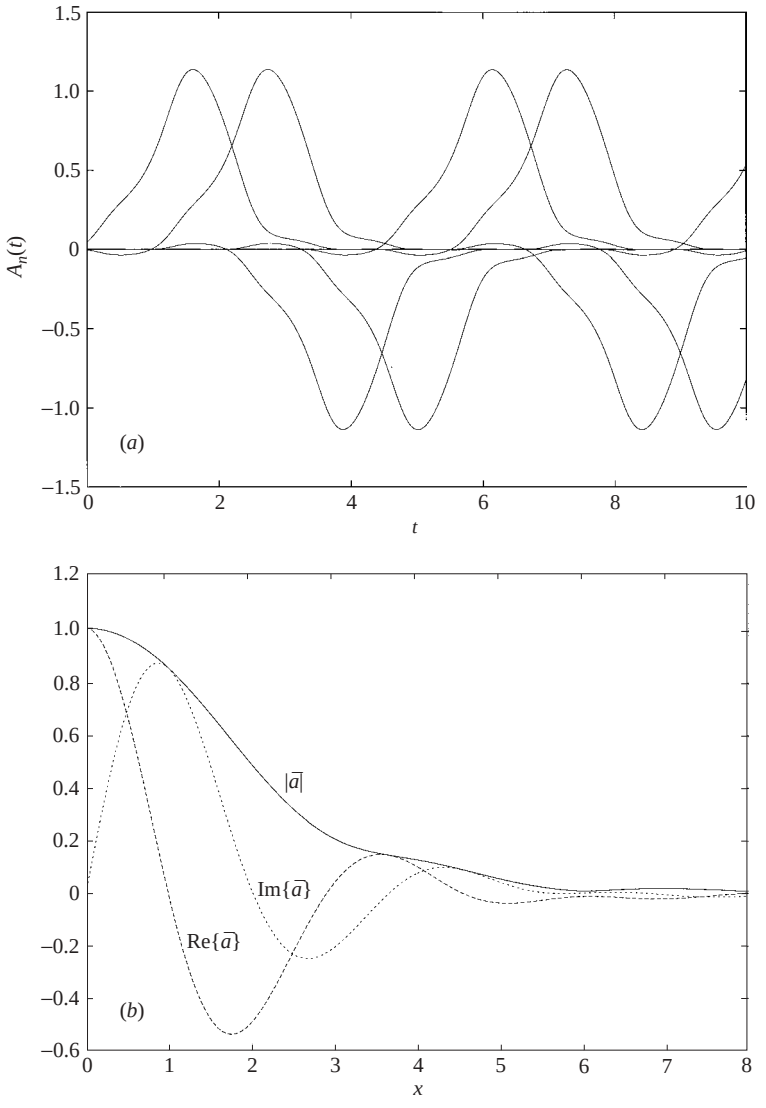

FIGURE 7. The solution at the point labelled IV on figure 1. Here $\lambda = 4$, $(L, \sqrt{\langle \mathscr{E} \rangle}) \approx (2.77229, 1.05800)$ and $T \approx 1.13321$. Sub-plots (a) and (b) correspond to their equivalents for figure 2 with $\overline{a}(0) \approx 1.0108$.

Our new solution is uniform on the longer $O(\varepsilon^{-1/3})$ time scale. The above construction has given us essentially a spatially uniform structure that evolves with pulses having separation $L(t)$, centres $\chi_n(t)$ and amplitudes $\overline{a}(x : L(t), \lambda_n(t))$. Both $L(t)$ and $\lambda_n(t)$, defined by (4.18) and (4.20), vary on the long time $O(\varepsilon^{-1/3})$. Evidently the solution $\overline{a}(x : L(t), \lambda_n(t))$ cannot remain valid on this long time for two reasons. First, the pulses at the outer edge of the locally unstable region must collapse at $\chi_n(t)$ once $L(t)$ exceeds $L_{\max}(\lambda_n(t))$. Second, as $L(t)$ increases so we must expect the pulse-train to become unstable to sideband instabilities. In short, we must anticipate chaotic behaviour on these long $O(\varepsilon^{-1/3})$ length and time scales with the underlying pulse-trains providing local structures on the shorter $O(1)$ length and time scales. A plausible scenario is that the pulse-train structures form space–time patches connected

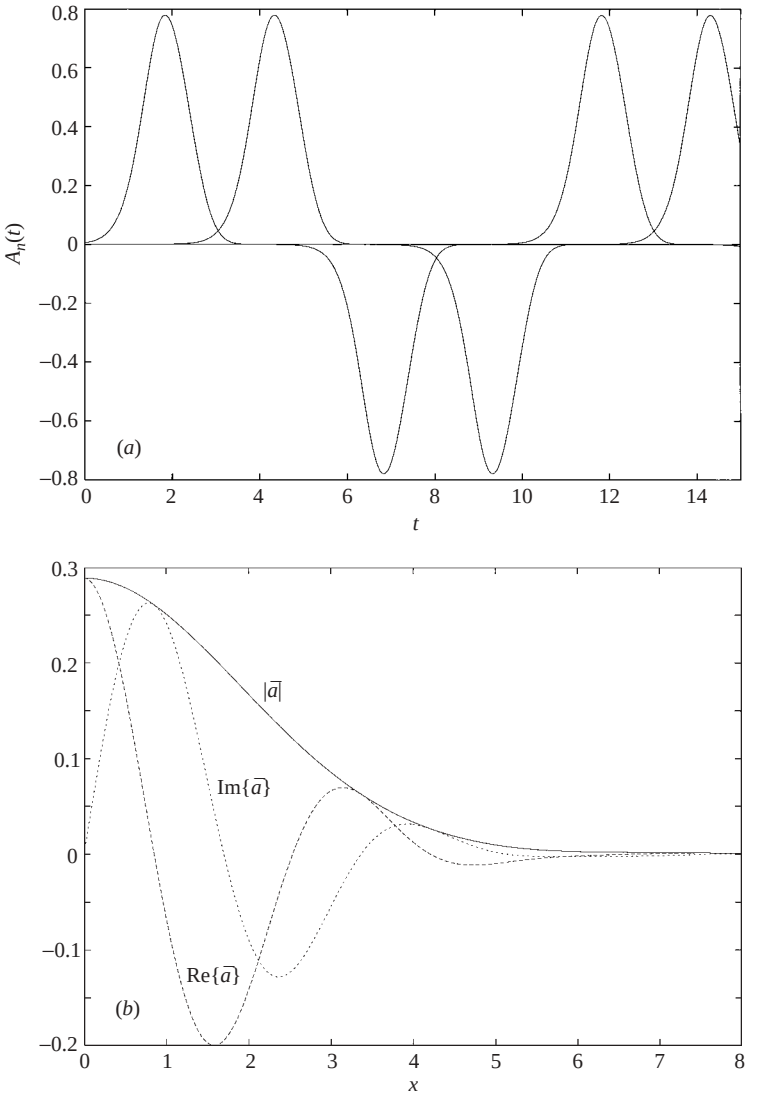

FIGURE 8. The solution at the point labelled V on figure 1. Here $\lambda = 4$, $(L, \sqrt{\langle \mathscr{E} \rangle}) \approx (1.26040, 0.47395)$ and $T \approx 2.49253$. Sub-plots (a) and (b) correspond to their equivalents for figure 2 with $\bar{a}(0) \approx 0.28905$.

chaotically one to another across dislocations where $L(T)$ suffers an abrupt change in value. We believe that they are sufficiently robust to regenerate themselves and to be space-filling in the majority of the local unstable region defined by

$$|\chi| < \sqrt{\lambda(0)}/\Upsilon_\varepsilon. \tag{4.24}$$

Remember also that the pulse-train amplitudes cannot decrease continuously to zero but must terminate abruptly at some finite size. Though this feature resembles a front, it is not one in the conventional sense as we explained towards the end of § 1. The distinction is important because in many single-frequency systems it is the frontal condition that selects the frequency which is determined by an inward group

velocity across the front, measured relative to the moving pulse. Our problem is not like this as our pulses move at the group velocity and so there is no relative velocity. Furthermore, c increases linearly with Θ_l and possesses the sign appropriate to outward propagation of energy from the equator towards the edge of the locally unstable region where the pulses eventually disappear. In that alternative sense the group velocity is directed outwards rather than inwards. Accordingly we anticipate that local conditions dominate frequency selection (i.e. the value of $L(t)$) in each patch. The final termination of pulse-trains at their outer limit might be regarded simply as yet another stronger and more severe type of dislocation.

Finally we comment briefly on the nature of each of the individual pulses that compose the wave-trains. To that end, we write $\overline{a}(x)$ in its polar form

$$\overline{a}(x) = R(x) \exp\left[\mathrm{i} \int_0^x K(x)\, \mathrm{d}x\right], \tag{4.25}$$

where the amplitude R and wavenumber K are real. The pulse solutions displayed in figures 2(b), 4(b), 7(b) and 8(b) exhibit positive K. This is likely to be a generic property and was certainly found to be the case for all the parameter values examined in HBS. The upshot is that the modulated wave, which it describes at $\chi = \chi_n(t)$, always propagates in the direction opposite to the sign of the frequency $\chi_n(t)$. That means that the waves propagate at velocity $-\chi_n(t)/K = O(\chi)$ towards $\chi = 0$, opposite to the direction of the group velocity $c = O(\varepsilon^{1/3}\chi)$. At the equator $\chi = 0$ the pulse with the nearest centre dominates and provides a wave which generally crosses the equator as illustrated in figure 5(c). Only when the solution has the symmetry-preserving property $\mathscr{A}_{\mathrm{SP}}(-\chi, t) = \mathscr{A}^*_{\mathrm{SP}}(\chi, t)$ with the equator half-way between pulse centres is there no wave propagation across the equator.

5. Concluding remarks

The trivial observation that the phase φ in (2.21) is arbitrary has important physical repercussions. Significantly the Taylor vortex pattern that (2.21) modulates may be shifted latitudinally at will by appropriate choice of φ. Only when the radial structure of the eigenfunctions $(f(z), g(z))$ in (2.5) directly influences the asymptotics can the vortex boundary location be fixed, as it will at some higher order which is outside the scope of our present study. This lack of determinism means, for example, that we cannot distinguish between cases in which the equator is a vortex boundary, a vortex centre or something asymmetric intermediate between the two. A similar lack of determinism has its parallel for non-axisymmetric disturbances, in that any modulation on an $O(1)$ azimuthal length scale introduces no new terms in our low-order amplitude equations and so has no influence on the results.

The polar representation (4.25) is helpful for an understanding of the nature of the Taylor vortex structure as implied by our solutions. The amplitude $R(x) = |\overline{a}(x)|$ clearly gives the amplitude of the envelope of the local wave. On the other hand the local wavenumber $K(x)$ only provides a small correction to the total wavenumber $\varepsilon^{-1/3}k_c + K$, which determines the local phase velocity $-\chi_n(t)/[\varepsilon^{-1/3}k_c + K]$. From that point of view its small magnitude $|K|$ has little physical significance; nevertheless its sign does. Indeed, since the signs of K and k_c are the same, the vortices propagate with velocity $-\varepsilon^{1/3}\chi_n(t)/k_c$ in the same direction as the envelope wave (4.25), which has velocity $-\chi_n(t)/K$. This is interesting because though the vortex drift velocity is the same for all vortices inside each pulse, it differs by $-\varepsilon^{1/3}[\chi_{n+1} - \chi_n(t)]/k_c = -\varepsilon^{1/3}L(t)/k_c$ between neighbours. The vortices in the outer

pulse overtake the inner pulse in such a way that one vortex is lost in an $O(1)$ time. Note that one envelope pulse wave is lost in exactly the same time, a process which is clearly evidenced by the dislocations at $\chi_n(t) + L(t)/2$ half-way between two pulse centres in figure 5(c); indeed that is the location at which the vortices evaporate.

When the spheres almost co-rotate, $\varepsilon^{1/2} \ll \delta \ll 1$, the dislocations are fixed in space. This limit is illustrated well by the HBS results portrayed in figure 5 for $\mathscr{A}_{\rm SB}(\chi, t)$ at a numerically large value of $\delta/\varepsilon^{1/2}$. Here, as in §4, χ is measured relative to the equatorial local origin $\theta_l = 0$. The grey shaded space–time contour plot of $\mathrm{Re}\{\mathscr{A}_{\rm SB}(\chi, t)\}$ in figure 5(c) shows how the envelope wave tends to propagate towards the equator as described towards the end of §4.2. As explained above, the Taylor vortices have a much shorter wavelength and so the space–time contours for the actual vortices are closely packed and are inclined at a very shallow angle to the t-axis. Thus the complicated temporal structure of $\mathscr{A}_{\rm SB}(\chi, t)$ displayed in figure 5(c) defines a smooth vortex distribution except at the dislocation half-way between the pulse centres identified by the temporal merging of two $\mathscr{A}_{\rm SB}(\chi, t)$ wave crests. Further insight is to be gleaned from the envelope of $|\mathscr{A}_{\rm SB}(\chi, t)|$ shown in figure 5(d) at some fixed time t. The local maxima and minima are close to the midpoints $\chi_n(t) + L(t)/2$ where the vortices disappear and it is here that the envelope has its maximum range of oscillation. In contrast, the amplitude at the centre of a pulse hardly varies and oscillates about $\overline{a}(0)$; these ideas are fully supported by the results in figures 2(d), 4(d) and 6(d). When, however, the outer sphere is at rest ($\delta = 1$), the dislocation drifts outwards over a distance comparable to that pulse width on the $O(1)$ time scale (see (4.19a) with $\chi_n(t) = O(\varepsilon^{-1/3})$).

Bartels (1982) obtained numerical SP-solutions for spherical Taylor–Couette flow by applying symmetry-preserving boundary conditions at the equator. His results for the narrow gap case $\varepsilon = 0.025$, with the outer sphere at rest, show complicated temporal behaviour that could be accounted for by our general description, albeit within our SP-class. That means that its space–time features bear some resemblance to figure 5(c) although modified so as to exhibit SP-characteristics such as those for $\mathscr{A}_{\rm SP}(\chi, t)$ with $\Omega = 0$ illustrated in figure 3 in the almost co-rotating limit.

For the medium gap $\varepsilon = 0.154$ Mamum & Tuckerman (1995) find subcritical finite-amplitude SP-solutions. On increasing the Taylor number they are first manifested as a pitchfork to steady states. In our small-ε narrow-gap limit, we find comparable subcritical behaviour but to time-dependent solutions, which generally possess broken symmetry. HBS found several different routes to travelling waves via various global bifurcations and so it is hardly surprising to find that on increasing ε several alternative scenarios are possible.

Sparsely distributed dislocations are apparent in non-axisymmetric vortices seen in the numerical results of Dumas & Leonard (1994) for the narrow-gap limit. These are just as our theory predicts albeit in the axisymmetric space–time context. Such modes are spiral vortices also discussed by Nakabayashi (1983), Nakabayashi & Tsuchida (1988a,b) and Sha & Nakabayashi (2001) but there in the medium-gap limit. Though non-axisymmetric, they exhibit travelling wave features comparable to ours. In their results wavy vortices are seen near the equator while, at higher latitudes, the vortices spiral away from the equator in the negative longitudinal (azimuthal) direction. If we tentatively identify the azimuthal coordinate with time, the spiral vortex states found in the papers cited immediately above resemble a wave propagating towards the equator as our waves under their pulse envelopes suggest.

Evidently the link with both laboratory and numerical experiments is somewhat tenuous. The main difficulty faced is that we have identified behaviours on three

different length and time scales. First, there is the short length scale εR_1 and time scale $\varepsilon^2 R_1^2/\nu$ of the vortices. Second, there is modulation on the intermediate length scale $\varepsilon^{2/3} R_1$ of the pulses for which the relevant time scale $\varepsilon^{4/3} R_1^2/\nu$ is inversely proportional to the frequency increment between neighbouring pulses. This is the space–time range over which our analysis is valid. Third, the pulses exist and are spatially modulated on a relatively wide locally unstable region width $O(\varepsilon^{1/3} R_1)$, though this is still short compared to the $O(R_1)$ length associated with the distance between the pole and the equator. The long time associated with temporal modulation caused by pulse separation due to the group velocity is $O(\varepsilon^{4/3} R_1^2/\nu)$. We can only speculate on the complicated spatio-temporal evolution over these longest scales. It would be difficult to conduct experiments at a sufficiently small ε such that these scale separations can be distinguished.

Though we have not proved that the pulse-trains persist on the longest time scales we have shown how the basic pulse unit can support the existence of its neighbour. Indeed the essential idea is that at every location there is a preferred frequency which increases linearly with respect to distance from the equator. Moreover the initial value calculation for spatially periodic solutions undertaken in § 3.1 was formulated (see the factor $\exp(\mathrm{i}tx)$ in (3.1)) to accommodate that preference. Nevertheless the realized temporally periodic forms to which the solution settles after the transients decay possess the discrete set rather than a continuous distribution of frequencies. Each pulse is localized in the vicinity of the point at which the frequency is preferred. Furthermore, the constant frequency jump between neighbouring pulses is essential for their mutual resonance. The fact that the pulse-train solutions emerge naturally as the solution to an initial value problem suggests that they are robust. At the outset of the current investigation it was far from clear whether pulse-trains were even possible. Our demonstration of their existence provides an affirmative answer to the long-outstanding question as to whether subcritical finite-amplitude solutions can occur in the vicinity of the local critical Taylor number.

We are grateful to the referees for their comments and suggestions for improving the presentation of this work. Part of this study was conducted while A. P. B. was at the School of Mathematics, University of New South Wales. He is indebted to the Australian Research Council and the Royal Society whose support made this visit possible. Further thanks are due to staff of the School (especially Peter Blennerhassett) and to the staff and students of New College, UNSW for their hospitality.

Appendix A. The equivalence of the time and space Fourier series representations

In order to establish the link between the Fourier time (2.23a) and space (3.8) series representations, we take advantage of the property $\mu_{n+4} = \mu_n$ for all integer n and write (3.8) in the alternative form

$$a(x,t) = \sum_{\alpha=0}^{3} \mu_\alpha \, a_\alpha(x,t), \tag{A 1a}$$

in which

$$a_\alpha(x,t) = \exp(\mathrm{i}tx) \sum_{\forall n} \widehat{A}(t + (4n+\alpha)T) \exp(\mathrm{i}(4n+\alpha)Tx). \tag{A 1b}$$

Our strategy is to consider each of the four Fourier space series $a_\alpha(x,t)$ and to recast them as Fourier time series. We then reassemble the sum (A 1a) as the Fourier time series (2.23a).

For the particular case $\alpha = 0$ the function

$$a_0(x,t) = \sum_{\forall n} \widehat{A}(t+4nT)\,\exp(\mathrm{i}(t+4nT)x) \tag{A 2a}$$

clearly has the property $a_0(x,t) = a_0(x,t+4T)$ and so is periodic on the time $4T$. It follows that $a_0(x,t)$ can be represented as the Fourier sum

$$a_0(x,t) = \sum_{\forall m} \overline{a}_m(x)\,\exp\left(\mathrm{i}m\frac{L}{2}t\right), \tag{A 2b}$$

for some functions $\overline{a}_m(x)$. Since $\exp(-\mathrm{i}tx)a_0(x,t)$ as defined by (A 2a) is spatially periodic on the length $\pi/2T \equiv L/2$, it has the symmetry property

$$\exp(-\mathrm{i}tx)\,a_0(x,t) = \exp\left[-\mathrm{i}t\left(x+\frac{L}{2}\right)\right] a_0\left(x+\frac{L}{2},t\right). \tag{A 3a}$$

Explicit substitution of the formula (A 2b) for a_0 on both sides gives

$$\sum_{\forall m} \overline{a}_m(x)\exp\left[-\mathrm{i}\left(x-m\frac{L}{2}\right)t\right] = \sum_{\forall m} \overline{a}_m\left(x+\frac{L}{2}\right)\exp\left[-\mathrm{i}\left(x-(m-1)\frac{L}{2}\right)t\right], \tag{A 3b}$$

which is satisfied when

$$\overline{a}_m\left(x+\frac{L}{2}\right) = \overline{a}_{m-1}(x) \tag{A 4a}$$

for all x and all integer m. This recurrence relation has the solution

$$\overline{a}_m(x) = \frac{1}{2\sqrt{2}}\,\overline{a}\left(x-m\frac{L}{2}\right), \tag{A 4b}$$

where as yet $\overline{a}(x)$ remains an undetermined function and the factor $1/2\sqrt{2}$ has been introduced for convenience, see (A 7). Substitution of this result into (A 2b) determines the Fourier time series

$$\exp(-\mathrm{i}tx)\,a_0(x,t) = \frac{1}{2\sqrt{2}}\sum_{\forall m}\exp\left[-\mathrm{i}\left(x-m\frac{L}{2}\right)t\right]\overline{a}\left(x-m\frac{L}{2}\right). \tag{A 5}$$

Since $a_\alpha(x,t) = a_0(x,t+\alpha T)$, it is readily established from (A 1b), (A 5) and noting $LT = \pi$, that

$$\exp(-\mathrm{i}tx)\,a_\alpha(x,t) = \frac{1}{2\sqrt{2}}\sum_{\forall m}\exp\left(\mathrm{i}m\alpha\frac{\pi}{2}\right)\exp\left[-\mathrm{i}\left(x-m\frac{L}{2}\right)t\right]\overline{a}\left(x-m\frac{L}{2}\right) \tag{A 6}$$

for each α. The substitution of (A 6) into (A 1a) and use of the identity

$$\frac{1}{2\sqrt{2}}\sum_{\alpha=0}^{3}\mu_\alpha\exp\left(\mathrm{i}n\alpha\frac{\pi}{2}\right) = \begin{cases} 0, & n \text{ even}, \\ \mathrm{e}^{\mathrm{i}\pi/4}, & n = \ldots,-3,1,5,\ldots, \\ \mathrm{e}^{-\mathrm{i}\pi/4}, & n = \ldots,-5,-1,3,\ldots, \end{cases} \tag{A 7}$$

recovers the pulse-train structure (2.23a).

Appendix B. The Fourier transform relationship (3.11*a*) of $\widehat{A}(t)$ to $\overline{a}(x)$

To establish (3.11*a*) it is sufficient to consider the spatially periodic function $\exp(-\mathrm{i}tx)\,a_0(x,t)$ defined by (A 2*a*). We multiply it by $\exp(-4\mathrm{i}nTx)$ and integrate over its periodicity length $\pi/2T = L/2$ to obtain the finite Fourier transform

$$\widehat{A}(t+4nT) = \frac{2}{L}\int_{-L/4}^{L/4} a_0(x,t)\,\exp[-\mathrm{i}(t+4nT)x]\,\mathrm{d}x. \tag{B 1a}$$

Since the expression (A 2*a*) has temporal period $4T$, we substitute $a_0(x,t) = a_0(x,t+4nT)$ into (B 1*a*) and then make the change of variable $t+4nT \to t$. In this way we deduce

$$\widehat{A}(t) = \frac{2}{L}\int_{-L/4}^{L/4} a_0(x,t)\,\exp(-\mathrm{i}tx)\,\mathrm{d}x. \tag{B 1b}$$

Finally, substitution of the formula (A 5) for $a_0(x,t)$ into (B 1*b*) yields

$$\widehat{A}(t) = \frac{1}{\sqrt{2}L}\int_{-L/4}^{L/4}\sum_{\forall m}\exp\left[-\mathrm{i}\left(x - m\frac{L}{2}\right)t\right]\overline{a}\left(x - m\frac{L}{2}\right)\mathrm{d}x, \tag{B 1c}$$

which, on reversing the order of integration and summation, gives (3.11*a*).

Appendix C. The relationship between the HBS problem and equation (4.9)

In order to prevent confusion between our notation here and that used by HBS, all quantities appearing in HBS are designated by a tilde. With this convention, equation (1.3) of HBS becomes

$$\frac{\partial\widetilde{a}}{\partial\widetilde{t}} = (\widetilde{\lambda} + 2\mathrm{i}\widetilde{\kappa}\widetilde{x} - \widetilde{x}^2 - |\widetilde{a}|^2)\widetilde{a} + \frac{\partial^2\widetilde{a}}{\partial\widetilde{x}^2}, \tag{C 1a}$$

which is retrieved from (4.9) upon making the transformations $t = \Upsilon_\varepsilon^{-1}\widetilde{t}$, $\chi = \Upsilon_\varepsilon^{-1/2}\widetilde{x}$, $\lambda = \Upsilon_\varepsilon\widetilde{\lambda}$ and $\mathscr{A} = \Upsilon_\varepsilon^{1/2}\widetilde{a}$. Under this reduction the coefficient of the quadratic spatial term χ^2 in (4.9) becomes unity while the $\mathrm{i}\chi$ term acquires the factor

$$2\widetilde{\kappa} \equiv \Upsilon_\varepsilon^{-3/2} \approx 0.541\,\delta/\varepsilon^{1/2}. \tag{C 1b}$$

This is the dimensionless parameter used by HBS and, in a sense, plays a role in their work similar to that of Υ_ε here.

Whereas HBS used the representation (2.23) for symmetry-preserved solutions, they effectively shifted their frequency by $L/2$ to describe symmetry-broken states, thereby removing the halves in the expansion (2.23*a*). Accordingly, we translate the data listed in HBS's table 2 for the case $\widetilde{\kappa} = 4$ (equivalently $\Upsilon_\varepsilon = 1/4$) into our table 1 with the scalings mentioned together with $L = \Upsilon_\varepsilon\widetilde{\omega}$ and $\Omega = \Upsilon_\varepsilon(\widetilde{\Omega} - \widetilde{\omega}/2)$.

REFERENCES

Bartels, F. 1982 Taylor vortices between two concentric rotating spheres. *J. Fluid Mech.* **119**, 1–25.

Bühler, K. 1990 Symmetric and asymmetric Taylor vortex flow in spherical gaps. *Acta Mech.* **81**, 3–38.

Couairon, A. & Chomaz, J.-M. 1997 Absolute and convective instabilities, front velocities and global modes in nonlinear systems. *Physica D* **108**, 236–276.

Couairon, A. & Chomaz, J.-M. 1999 Fully nonlinear global modes in slowly varying flows. *Phys. Fluids* **11**, 3688–3703.

Davey, A. 1962 The growth of Taylor vortices in flow between rotating cylinders. *J. Fluid Mech.* **14**, 336–368.

Dee, G. & Langer, J. S. 1983 Propagating pattern selection. *Phys. Rev. Lett.* **50**, 383–386.

Doedel, E. J., Champneys, A. R., Fairgrieve, T. F., Kuznetsov, Y. A., Sandstede, B. & Wang, X. 1997 AUTO97: Continuation and bifurcation software for ordinary differential equations. Available via anonymous ftp from `ftp://ftp.cs.concordia.ca/pub/doedel/auto`.

Dumas, G. & Leonard, A. 1994 A divergence-free spectral expansions method for three-dimensional flows in spherical-gap geometries. *J. Comput. Phys.* **111**, 205–219.

Ewen, S. A. & Soward, A. M. 1994*a* Phase mixed rotating magnetoconvection and Taylor's condition II. Travelling pulses. *Geophys. Astrophys. Fluid Dyn.* **77**, 231–262.

Ewen, S. A. & Soward, A. M. 1994*b* Phase mixed rotating magnetoconvection and Taylor's condition III. Wave trains. *Geophys. Astrophys. Fluid Dyn.* **77**, 263–283.

Harris, D., Bassom, A. P. & Soward, A. M. 2000 An inhomogeneous Landau equation with application to spherical Couette flow in the narrow gap limit. *Physica D* **137**, 260–276.

Harris, D., Bassom, A. P. & Soward, A. M. 2003 Global bifurcation to travelling waves with application to narrow gap spherical Couette flow. *Physica D* **177**, 122–174 (referred to herein as HBS).

Heyvaerts, J. & Priest, E. R. 1983 Coronal heating by phase-mixed shear Alfvén waves. *Astron. Astrophys.* **117**, 220–234.

Hocking, L. M. 1981 The instability of flow in the gap between two prolate spheroids. Part II. Arbitrary axis ratio. *Q. J. Mech. Appl. Maths* **34**, 475–488.

Hocking, L. M. & Skiepko, J. 1981 The instability of flow in the gap between two prolate spheroids. Part I. Small axis ratio. *Q. J. Mech. Appl. Maths* **34**, 57–68.

Huerre, P. & Monkewitz, P. A. 1990 Local and global instabilities in spatially developing flows. *Annu. Rev. Fluid Mech.* **14**, 473–537.

Jones, C. A., Soward, A. M. & Mussa, A. I. 2000 The onset of thermal convection in a rapidly rotating sphere. *J. Fluid Mech.* **405**, 157–179.

Mamum, C. K. & Tuckerman, L. S. 1995 Asymmetry and Hopf bifurcation in spherical Couette flow. *Phys. Fluids* **7**, 80–91.

Marcus, P. S. & Tuckerman, L. S. 1987*a* Simulation of flow between concentric rotating spheres. Part 1. Steady states. *J. Fluid Mech.* **185**, 1–30.

Marcus, P. S. & Tuckerman, L. S. 1987*b* Simulation of flow between concentric rotating spheres. Part 2. Transitions. *J. Fluid Mech.* **185**, 31–65.

Meunier, N., Proctor, M. R. E., Sokoloff, D., Soward, A. M. & Tobias, S. M. 1997 Asymptotic properties of a nonlinear $\alpha\omega$-dynamo wave: period, amplitude and latitude dependence. *Geophys. Astrophys. Fluid Dyn.* **86**, 249–285.

Nakagayashi, K. 1983 Transition of Taylor-Görtler vortex flow in spherical Couette flow. *J. Fluid Mech.* **132**, 209–230.

Nakagayashi, K. & Tsuchida, Y. 1988*a* Spectral study of the laminar-turbulent transition in spherical Couette flow. *J. Fluid Mech.* **194**, 101–132.

Nakagayashi, K. & Tsuchida, Y. 1988*b* Modulated and unmodulated travelling azimuthal waves on the toroidal vortices in a spherical Couette system. *J. Fluid Mech.* **195**, 495–522.

Pier, B. 2002 On the frequency selection of finite-amplitude vortex shedding in the cylinder wake. *J. Fluid Mech.* **458**, 407–417.

Pier, B. & Huerre, P. 2001 Nonlinear self-sustained structures and fronts in spatially developing wake flows. *J. Fluid Mech.* **435**, 145–174.

Pier, B., Huerre, P. & Chomaz, J.-M. 2001 Bifurcation to fully nonlinear synchronized structures in slowly varying media. *Physica D* **108**, 49–96.

Pier, B., Huerre, P., Chomaz, J.-M. & Couairon, A. 1998 Steep nonlinear global modes in spatially developing media. *Phys. Fluids* **10**, 2433–2435.

Sha, W. & Nakagayashi, K. 2001 On the structure and formation of spiral Taylor-Görtler vortices in spherical Couette flow. *J. Fluid Mech.* **431**, 323–345.

Soward, A. M. 1977 On the finite amplitude thermal instability of a rapidly rotating fluid sphere. *Geophys. Astrophys. Fluid Dyn.* **9**, 19–74.

Soward, A. M. & Bassom, A. P. 2003 Finite amplitude subcritical instability in narrow-gap spherical Couette flow. In contributions at the *13th International Couette-Taylor Workshop, Universitat Politècnica de Catalunya; Nonlinear Dynamics in Fluids* (ed. F. Marqués &

A. Meseguer), pp. 190–193. International Center for Numerical Methods in Engineering (CIMNE), Barcelona, Spain (ISBN 84-95999-28-5).

Soward, A. M. & Jones, C. A. 1983 The linear stability of flow in the gap between two concentric rotating spheres. *Q. J. Mech. Appl. Maths* **36**, 19–41.

Tobias, S. M., Proctor, M. R. E. & Knobloch, E. 1998 Convective and absolute instabilities of fluid flows in finite geometry. *Physica D* **113**, 43–72.

Walton, I. C. 1978 The linear instability of flow in the narrow spherical annulus. *J. Fluid Mech.* **86**, 673–693.

Wimmer, M. 1976 Experiments on a viscous fluid flow between concentric rotating spheres. *J. Fluid Mech.* **78**, 317–335.

Wimmer, M. 1981 Experiments on the stability of viscous flow between two concentric rotating spheres. *J. Fluid Mech.* **103**, 117–131.