Beyond the Fokker-Planck equation: pathwise control of noisy bistable systems

2002 J. Phys. A: Math. Gen. 35 2057

(http://iopscience.iop.org/0305-4470/35/9/301)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 2057–2091 PII: S0305-4470(02)29774-7

Beyond the Fokker–Planck equation: pathwise control of noisy bistable systems

Nils Berglund[1,3] **and Barbara Gentz**[2]

[1] Department of Mathematics, ETH Zürich, ETH Zentrum, 8092 Zürich, Switzerland
[2] Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany

E-mail: berglund@math.ethz.ch and gentz@wias-berlin.de

Received 8 October 2001, in final form 17 January 2002
Published 22 February 2002
Online at stacks.iop.org/JPhysA/35/2057

Abstract

We introduce a new method, allowing one to describe slowly time-dependent Langevin equations through the behaviour of individual paths. This approach yields considerably more information than the computation of the probability density. In particular, scaling laws can be obtained easily. The main idea is to show that for sufficiently small noise intensity and slow time dependence, the vast majority of paths remain in small space–time sets, typically in the neighbourhood of potential wells. The size of these sets often has a power-law dependence on the small parameters, with universal exponents. The overall probability of exceptional paths is exponentially small, with an exponent also showing power-law behaviour. The results cover time spans up to the maximal Kramers time of the system. We apply our method to three phenomena characteristic for bistable systems: stochastic resonance, dynamical hysteresis and bifurcation delay, where it yields precise bounds on transition probabilities, and the distribution of hysteresis areas and first-exit times. We also discuss the effect of coloured noise.

PACS numbers: 05.10.Gg, 02.50.-r, 75.60.-d, 92.40.Cy

Mathematics Subject Classification: 37H20, 60H10, 34E15, 82C31

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Noise is often used to model the effect of fast degrees of freedom, which are too involved to describe otherwise. In statistical physics and solid state physics, for instance, the influence of a heat bath is represented by stochastic Glauber dynamics or a Langevin equation. In meteorological and climate models, the effect of fast modes (e.g. short-wavelength

[3] Current address: CPT–CNRS, Campus de Luminy, F-13288 Marseille cedex 9, France.

modes neglected in a Galerkin approximation) is often described by noise [Ha, Ar2]. As a consequence, stochastic differential equations are widely used to model systems of physical interest, including ferromagnets [Mar], lasers [Ri, HL, HN], neurons [Tu, Lo], glacial cycles [BPSV], oceanic circulation [Ce], biomolecules [SHD] and others.

Among the simpler stochastic models in use is the Langevin equation with additive white noise

$$\mathrm{d}x_t = -\nabla V(x_t, \lambda)\,\mathrm{d}t + \sigma G(\lambda)\,\mathrm{d}W_t. \tag{1.1}$$

Here V is a potential, W_t denotes a standard vector-valued Wiener process (i.e. a Brownian motion), and σ measures the noise intensity. For now, we consider λ as a fixed parameter, but in what follows we will be concerned with situations where λ varies slowly in time. Of course, one may be interested in situations where noise enters in a different way, for instance G depending on x as well [BG4], or coloured noise.

There exist different methods to characterize the dynamics of the Langevin equation (1.1). A popular approach is to determine the probability density $p(x, t)$ of x_t, which gives all information on the instantaneous state of the system. For instance, the probability that x_t belongs to a subset $\mathcal{D}$ of phase space is given by the integral of $p(x, t)$ over $\mathcal{D}$. The density is given by the normalized solution of the Fokker–Planck equation

$$\frac{\partial}{\partial t} p(x, t) = \nabla \cdot (\nabla V(x, \lambda) p(x, t)) + \nabla \cdot D(\lambda) \nabla p(x, t) \tag{1.2}$$

where $D = (\sigma^2/2) GG^T$ is the diffusion matrix. In particular, in the isotropic case $GG^T = \mathbb{1}$, (1.2) admits the stationary solution

$$p_0(x) = \frac{1}{N} \mathrm{e}^{-2V(x)/\sigma^2} \tag{1.3}$$

where N is the normalization. For small σ, the stationary distribution is sharply peaked around the minima of the potential. For an arbitrary initial distribution, however, the Fokker–Planck equation cannot be solved in general, and one has to rely on spectral methods, WKB approximations and the like.

Even if we have obtained a solution of (1.2), this approach still has serious shortcomings. The reason for this is that the probability density only gives an instantaneous picture of the system. If, for instance, we want to compute correlation functions such as $\mathbb{E}\{x_s x_t\}$ for $0 < s < t$, solving the Fokker–Planck equation with initial condition x_0 is not enough: we need to solve it for *all* initial conditions (x_s, s). Quantities such as the supremum of $\|x_t\|$ over some time interval are even harder to handle.

It is important to take into account that the stochastic differential equation (1.1) does not only induce a probability distribution of x_t, but also generates a measure on the paths, which contains much more information. For almost any realization $W_t(\omega)$ of the Brownian motion, and any deterministic initial condition x_0, the solution $\{x_t(\omega)\}_{t \geqslant 0}$ of (1.1) is a continuous function of time (though not differentiable). The random variable $x_t(\omega)$ for fixed t is only one of many interesting random quantities which can be associated with the stochastic process.

First-exit times form an important class of such alternative random variables, and have been studied in detail. If $\mathcal{D}$ is a (measurable) subset of phase space, the first-exit time of x_t from $\mathcal{D}$ is defined as

$$\tau_{\mathcal{D}} := \inf\{t > 0 \colon x_t \notin \mathcal{D}\}. \tag{1.4}$$

For instance, if $\mathcal{D}$ is a set of the form $\{x \colon V(x) \leqslant V_1\}$, containing a unique equilibrium point which is stable, then the distribution of $\tau_{\mathcal{D}}$ is asymptotically exponential, with expectation behaving in the small-noise limit like Kramers' time

$$T_{\text{Kramers}} = \mathrm{e}^{2(V_1 - V_0)/\sigma^2} \qquad \text{where } V_0 := \min_{x \in \mathcal{D}} V(x). \tag{1.5}$$

A mathematical theory allowing one to estimate first-exit times for general n-dimensional systems (with a drift term not necessarily deriving from a potential) has been developed by Freidlin and Wentzell [FW]. In specific situations, more precise results are available, as for instance the following. If $\mathcal{D}$ contains a unique, stable equilibrium point, subexponential corrections to the asymptotic expression (1.5) are known, even for possibly time-dependent drift terms [Az, FJ]. The case where $\mathcal{D}$ contains a saddle as a unique equilibrium point has been considered by Kifer in the seminal paper [Ki]. The situation where $\mathcal{D}$ contains a stable equilibrium in its interior and a saddle on its boundary is dealt with in [MS1], employing the method of matched asymptotic expansions. The limiting behaviour of the distribution of the first-exit time from a neighbourhood of a unique stable equilibrium point as well as from a neighbourhood of a saddle point has been obtained by Day [Day1, Day2].

We remark that more recently, another approach has been introduced, which mimics concepts from the theory of dissipative dynamical systems [CF1, Schm, Ar1]. The main idea is that for a given realization of the noise (i.e. in a quenched picture), paths with different initial conditions may converge to an attractor, which has similar properties as deterministic attractors. Of course, in experiments, this random attractor is only visible if we manage to repeat the experiment many times with the same realization of noise.

The method of choice to study the Langevin equation (1.1) also depends on the timescale we are interested in. Consider, for instance, a one-dimensional double-well potential $V(x) = -\frac{1}{2}ax^2 + \frac{1}{4}bx^4$, $a, b > 0$, with a barrier of height $H = a^2/4b$. Assume that x_0 is concentrated at the bottom $x = \sqrt{a/b}$ of the right-hand potential well. If the noise is sufficiently weak, paths are likely to stay in the right-hand well for a long time. The distribution of x_t will first approach a Gaussian in a time of order

$$T_{\text{relax}} = \frac{1}{c} \tag{1.6}$$

where $c = 2a$ is the curvature at the bottom of the well (the variance of the Gaussian is approximately $\sigma^2/(2c)$). With overwhelming probability, paths will remain in metastable equilibrium inside the same potential well, for all times significantly shorter than Kramers' time $T_{\text{Kramers}} = \mathrm{e}^{2H/\sigma^2}$. Only on longer timescales will the density of x_t approach the bimodal stationary density (1.3). The dynamics will thus be very different on the timescales $t \ll T_{\text{relax}}$, $T_{\text{relax}} \ll t \ll T_{\text{Kramers}}$ and $t \gg T_{\text{Kramers}}$. Random attractors can be reached only in the last regime. In particular, results in [CF2] stating that for the double-well potential V, the random attractor almost surely consists of one (random) point apply to the regime $t \gg T_{\text{Kramers}}$.

In this work, we are interested in situations where the parameter λ varies slowly in time, that is, we consider stochastic differential equations (SDEs) of the form

$$\mathrm{d}x_t = -\nabla V(x_t, \lambda(\varepsilon t))\,\mathrm{d}t + \sigma G(\lambda(\varepsilon t))\,\mathrm{d}W_t. \tag{1.7}$$

Such situations occur if the system under consideration is slowly forced, for instance by an external magnetic field, by climatic changes, or by a varying energy supply. Note that the probability density of x_t still obeys a Fokker–Planck equation, but there will be no stationary solution in general.

The main idea of our approach to equations of the form (1.7) is the following. For sufficiently small ε and σ, and on an appropriate timescale, we show that the paths $\{x_t\}_{t\geqslant 0}$ can be divided into two classes. The first class consists of those paths which remain in certain space–time sets, typically in the neighbourhood of potential wells (but in some cases, they may switch potential wells). The geometry and size of these sets depends on noise intensity and shape of the potential. The second class consists of the remaining paths, which we do not try to describe in detail, but whose overall probability is small, typically exponentially small in some combination of ε and σ. In this way we obtain, in particular, concentration

results on the density of x_t without solving the Fokker–Planck equation, but we also obtain global information on the stochastic process, including correlations, first-exit times, transition probabilities and, for instance, the shape of hysteresis cycles.

The slow time dependence of λ introduces a new timescale. If, for instance, λ is periodic with period 1, then (1.7) depends periodically on time with period $T_{\text{forcing}} = 1/\varepsilon$. Since the shape of the potential changes in time, the definitions (1.5) and (1.6) of Kramers and relaxation times no longer make sense. We can, however, define these timescales by

$$T_{\text{Kramers}}^{(\max)} = \mathrm{e}^{2H_{\max}/\sigma^2} \qquad \text{and} \qquad T_{\text{relax}}^{(\min)} = \frac{1}{c_{\max}} \tag{1.8}$$

where $H_{\max}$ and $c_{\max}$ denote, respectively, the maximal values of barrier height and curvature of a potential well over one period. Our results typically apply to the regime

$$T_{\text{relax}}^{(\min)} \ll T_{\text{forcing}} \ll T_{\text{Kramers}}^{(\max)}, \tag{1.9}$$

that is, we require that $\varepsilon \ll c_{\max}$ and $\sigma^2 \ll 2H_{\max}/|\log \varepsilon|$. In this regime, the process has time to approach a metastable 'equilibrium' during one period, but not the stationary distribution given by (1.3). The minimal curvature and barrier height, however, are allowed to become small, or even to vanish.

We can describe the paths' behaviour on time intervals including many periods of the forcing, as long as they are significantly shorter than the maximal Kramers time. We find it convenient to measure time in units of the forcing period. After scaling t by a factor ε, the relevant timescales become

$$T_{\text{relax}}^{(\min)} = \frac{\varepsilon}{c_{\max}} \qquad T_{\text{forcing}} = 1 \qquad T_{\text{Kramers}}^{(\max)} = \varepsilon \mathrm{e}^{2H_{\max}/\sigma^2}. \tag{1.10}$$

When scaling the Brownian motion, we should keep in mind its diffusive nature, which implies that in the new units, its standard deviation grows as $\sqrt{t/\varepsilon}$. Equation (1.7) thus becomes

$$\mathrm{d}x_t = -\frac{1}{\varepsilon}\nabla V(x_t, \lambda(t))\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,G(\lambda(t))\,\mathrm{d}W_t. \tag{1.11}$$

In the deterministic case $\sigma = 0$, we will sometimes write this equation in the form $\varepsilon\dot{x} = -\nabla V(x, \lambda(t))$, which is customary in singular perturbation theory. Throughout this paper, we will assume that V, G and λ are smooth functions, that $V(x, \lambda)$ grows at least like $\|x\|^2$ for large $\|x\|$, and that the matrix elements of G, as well as λ, have uniformly bounded absolute values. A large part of this paper is devoted to one-dimensional systems, and we come back to the multidimensional case in section 6.

The aim of this paper is to illustrate our methods by applying them to a number of physically interesting examples. We thus emphasize the conceptual aspects, and refrain from giving mathematical details of the proofs, which will be presented elsewhere [BG1, BG2, BG3].

We start, in section 2, by explaining the basic ideas in the simplest situation, where most paths remain concentrated near the bottom of a potential well. We also briefly describe the dynamics near a saddle. The three subsequent sections are devoted to more interesting cases in which paths may jump between the wells of a double-well potential. Note that we restrict our attention to double-well potentials only to keep the presentation simple, but the dynamics in more complicated multi-well potentials can be described by the same approach.

In section 3, we apply our method to the phenomenon of stochastic resonance, where noise allows for transitions between potential wells which would be impossible in the deterministic case. We compute the threshold noise intensity needed for transitions to be likely, and show that most paths are close, in a natural geometrical sense, to a periodic function. This provides an alternative quantitative measure of the signal's periodicity to the commonly used signal-to-noise ratio. Although stochastic resonance has been widely studied, to our best knowledge the

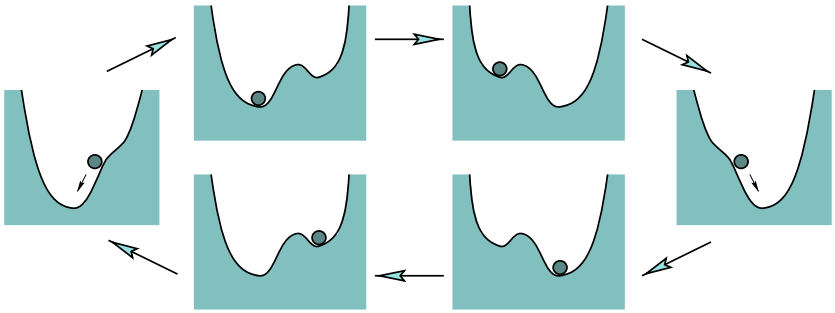

Figure 1. Example of a periodically forced double-well potential. In the limit of infinitely slow forcing, an overdamped particle tracks the bottom of potential wells. However, its position is not determined by the instantaneous value of the forcing alone. The system describes a hysteresis cycle, whose shape depends on the frequency of the forcing. Noise may kick the particle over the potential barrier, and thus influences the shape of the hysteresis cycle.

subtle dependence of the threshold noise intensity and the probability of exceptional paths on the small parameters were unknown so far.

Section 4 is devoted to hysteresis, which is also characteristic for forced bistable systems. For sufficiently strong forcing, one of the two potential wells disappears, causing trajectories to switch between potential wells even when no noise is present (figure 1). As a result, even in the adiabatic limit, the instantaneous value of the parameter λ does not suffice to determine the state of the system: solutions follow hysteresis cycles. In absence of noise, their area is known to scale in a nontrivial way with the frequency of the forcing. Our methods allow us to characterize the distribution of the random hysteresis area for positive noise intensity. In particular, we show that for noise intensities above an amplitude-dependent threshold, the typical area no longer depends, to leading order, on the amplitude or frequency of the forcing. We believe that a pathwise study as presented here is inevitable in order to obtain results on the size of random hysteresis cycles.

In section 5, we consider the effect of additive noise on systems with spontaneous symmetry breaking, i.e. when the potential transforms from single to double well. In the deterministic case, solutions are known to track the saddle for a considerable time before falling into one of the potential wells, a phenomenon known as bifurcation delay. We characterize the effect of additive noise on this delay, and on the probability to choose one or the other potential well after the bifurcation. Furthermore, we give results on the concentration of paths near potential wells which allow, in particular, one to determine the optimal relation between speed of parameter drift and noise intensity for an experimental determination of the bifurcation diagram. Advantage over previously known results is the description of the whole sample path—which is what the observer would typically see.

Finally, section 6 contains some generalizations. We first discuss analogous results to those of section 2 for multidimensional potentials. This formalism allows us to treat the effect of the simplest kind of coloured noise, given by an Ornstein–Uhlenbeck process, in a natural way. We conclude by discussing the dependence of previously discussed phenomena on noise colour.

2. Near wells and saddles

We start by discussing situations in which the noise intensity is sufficiently small, compared to the depth of a given potential well, for paths to remain concentrated near the bottom of the

well during a long time interval. In section 2.1, the solvable linear case is used to compare the information provided by the Fokker–Planck equation and by the pathwise approach. In section 2.2, we show that our method naturally extends to the nonlinear case. We briefly describe the dynamics near a saddle in section 2.3.

2.1. Linear case

It is instructive to consider first the case of a linear force (that is, of a quadratic potential), which can be solved completely. A general one-dimensional, time-dependent quadratic potential can be written as

$$V(x,t) = \tfrac{1}{2}c(t)(x - x^\star(t))^2 \tag{2.1}$$

where $x^\star(t)$ is the location of the potential minimum, and $c(t)$ is the curvature of the potential. The SDE (1.7) takes the form

$$\mathrm{d}x_t = \frac{1}{\varepsilon}a^\star(t)(x_t - x^\star(t))\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}g(t)\,\mathrm{d}W_t \tag{2.2}$$

where $a^\star(t) = -c(t)$. Throughout this paper, we will use $a^\star$ to denote the linearization of the force at an equilibrium point $x^\star$, with $a^\star < 0$ if the equilibrium is stable, and $a^\star > 0$ if it is unstable. In this subsection and the following one, we consider the stable case, and assume that $a^\star(t) \leqslant -a_0$ for all t under consideration, where a_0 is a positive constant. For simplicity, we shall assume that the functions $x^\star(t)$ and $a^\star(t)$, as well as $g(t)$, are real-analytic.

Let us first investigate the probability density $p(x,t)$ of x_t. The Fokker–Planck equation being linear, it can be easily solved. Assume for simplicity that the distribution of x_0 is Gaussian, with expectation $\mathbb{E}\{x_0\}$ and (possibly zero) variance $\mathrm{var}\{x_0\}$. (In addition, we always assume the initial distribution to be independent of the Brownian motion.) Then x_t has a Gaussian distribution for any $t > 0$, with density

$$p(x,t) = \frac{1}{\sqrt{2\pi\,\mathrm{var}\{x_t\}}}\exp\left\{-\frac{(x - \mathbb{E}\{x_t\})^2}{2\mathrm{var}\{x_t\}}\right\} \tag{2.3}$$

where expectation and variance of x_t obey the ODEs

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\{x_t\} = \frac{1}{\varepsilon}a^\star(t)(\mathbb{E}\{x_t\} - x^\star(t)) \tag{2.4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{var}\{x_t\} = \frac{2}{\varepsilon}a^\star(t)\mathrm{var}\{x_t\} + \frac{\sigma^2}{\varepsilon}g(t)^2. \tag{2.5}$$

Note that $\mathbb{E}\{x_t\}$ coincides with the deterministic solution x_t^{det} of equation (2.2) for $\sigma = 0$, with initial condition $x_0^{\mathrm{det}} = \mathbb{E}\{x_0\}$:

$$\mathbb{E}\{x_t\} = x_t^{\mathrm{det}} = x_0^{\mathrm{det}}\mathrm{e}^{\alpha^\star(t)/\varepsilon} - \frac{1}{\varepsilon}\int_0^t \mathrm{e}^{\alpha^\star(t,s)/\varepsilon}a^\star(s)x^\star(s)\,\mathrm{d}s \tag{2.6}$$

where we use the notations

$$\alpha^\star(t,s) = \int_s^t a^\star(u)\,\mathrm{d}u, \qquad \alpha^\star(t) = \alpha^\star(t,0). \tag{2.7}$$

Our stability assumption $a^\star(t) \leqslant -a_0\ \forall t$ implies that $\alpha^\star(t,s) \leqslant -a_0(t-s)$ for $t > s$. Hence the first term on the right-hand side of (2.6) decreases exponentially fast: it is at most of order ε after time $\varepsilon|\log\varepsilon|/a_0$, at most of order ε^2 after time $2\varepsilon|\log\varepsilon|/a_0$, and so on.

We expect x_t^{det} to follow adiabatically the slowly drifting bottom of the potential well. To make this apparent, we evaluate the second term on the right-hand side of (2.6) by integration by parts:

$$-\frac{1}{\varepsilon}\int_0^t \mathrm{e}^{\alpha^\star(t,s)/\varepsilon}a^\star(s)x^\star(s)\,\mathrm{d}s = x^\star(t) - x^\star(0)\mathrm{e}^{\alpha^\star(t)/\varepsilon} - \int_0^t \mathrm{e}^{\alpha^\star(t,s)/\varepsilon}\dot{x}^\star(s)\,\mathrm{d}s. \tag{2.8}$$

By successive integrations by parts, we find that the general solution of (2.4) can be written as

$$\mathbb{E}\{x_t\} = x_t^{\mathrm{det}} = \bar{x}_t^{\mathrm{det}} + (x_0^{\mathrm{det}} - \bar{x}_0^{\mathrm{det}})\mathrm{e}^{\alpha^\star(t)/\varepsilon} \tag{2.9}$$

where $\bar{x}_t^{\mathrm{det}}$ is a particular solution of (2.4), admitting the asymptotic expansion[4]

$$\bar{x}_t^{\mathrm{det}} = x^\star(t) + \varepsilon\frac{\dot{x}^\star(t)}{a^\star(t)} + \varepsilon^2\frac{1}{a^\star(t)}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{x}^\star(t)}{a^\star(t)}\right) + \cdots. \tag{2.10}$$

Since $a^\star(t)$ is negative, $\bar{x}_t^{\mathrm{det}}$ tracks the bottom $x^\star(t)$ of the potential well with a small lag: $\bar{x}_t^{\mathrm{det}} < x^\star(t)$ if $x^\star(t)$ moves to the right, and $\bar{x}_t^{\mathrm{det}} > x^\star(t)$ if $x^\star(t)$ moves to the left. The particular solution (2.10) is called the *adiabatic solution* or *slow solution*. Relation (2.9) expresses the fact that all solutions of (2.4) are attracted exponentially quickly by the adiabatic solution $\bar{x}_t^{\mathrm{det}}$.

The variance of x_t can be computed in a similar way. The solution of (2.5) is given by

$$\mathrm{var}\{x_t\} = \mathrm{var}\{x_0\}\mathrm{e}^{2\alpha^\star(t)/\varepsilon} + \frac{\sigma^2}{\varepsilon}\int_0^t \mathrm{e}^{2\alpha^\star(t,s)/\varepsilon} g(s)^2\,\mathrm{d}s. \tag{2.11}$$

The behaviour of the variance is very similar to the behaviour of the deterministic solution (2.6): the initial condition $\mathrm{var}\{x_0\}$ is forgotten exponentially fast, and in analogy with (2.9) and (2.10), we can write

$$\mathrm{var}\{x_t\} = \bar{v}(t) + (\mathrm{var}\{x_0\} - \bar{v}(0))\mathrm{e}^{2\alpha^\star(t)/\varepsilon} \tag{2.12}$$

where $\bar{v}(t)$ is a particular solution of (2.5), which admits the asymptotic expansion

$$\bar{v}(t) = \frac{\sigma^2}{2|a^\star(t)|}\left[g(t)^2 + \varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{g(t)^2}{2a^\star(t)}\right) + \cdots\right]. \tag{2.13}$$

The fact that x_t has a Gaussian distribution implies in particular that for any $t > 0$,

$$\mathbb{P}\left\{|x_t - x_t^{\mathrm{det}}| > h\sqrt{\mathrm{var}\{y_t\}}\right\} = 2\int_{h\sqrt{\mathrm{var}\{y_t\}}}^{\infty} p(x_t^{\mathrm{det}} + y, t)\,\mathrm{d}y \leqslant \mathrm{e}^{-h^2/2}. \tag{2.14}$$

Hence, the distribution of x_t is concentrated in an interval of width $\sqrt{2\mathrm{var}\{y_t\}}$ around x_t^{det}, which behaves asymptotically like $\sigma g(t)/\sqrt{|a^\star(t)|}$ by (2.12) and (2.13). In words, the spreading of x_t is proportional to the noise intensity and inversely proportional to the square root of the curvature of the potential: flatter potentials give rise to a larger spreading of the distribution.

Up to now, we have only studied the probability density of x_t. However, even if the density is concentrated near the bottom of the well at all times, this does not exclude that the path $\{x_s\}_{0\leqslant s\leqslant t}$ makes occasional excursions away from $x^\star$. From now on, we consider the initial condition x_0 as deterministic, so that the path depends only on the realization of the Brownian motion. The solution of the SDE (2.2) can be written as

$$x_t = x_t^{\mathrm{det}} + y_t \qquad y_t = \frac{\sigma}{\sqrt{\varepsilon}}\int_0^t \mathrm{e}^{\alpha^\star(t,s)/\varepsilon} g(s)\,\mathrm{d}W_s. \tag{2.15}$$

The process y_t is a generalization of an Ornstein–Uhlenbeck process, with time-dependent damping and diffusion. Ideally, we would like to estimate the probability that the path leaves a strip of (time-dependent) width proportional to $\sqrt{\mathrm{var}\{x_t\}}$, and centred at x_t^{det}. This turns out to be difficult because the variance may change quickly for t very close to 0, due to the first term on the right-hand side of (2.11). To avoid these technical complications, we use a strip of width proportional to $\sqrt{\bar{v}(t)}$, defined by

$$\mathcal{B}(h) = \left\{(x, t) : |x - x_t^{\mathrm{det}}| < h\sqrt{\bar{v}(t)}\right\}. \tag{2.16}$$

[4] The asymptotic series does not converge in general, but it admits expansions to any order in ε, with a remainder which can be controlled.

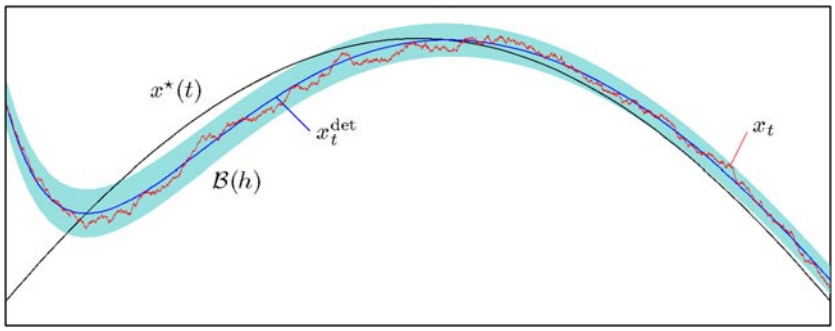

Figure 2. A sample path x_t of the linear equation (2.2) for $a^\star(t) = -4 + 2\sin(4\pi t)$, $x^\star(t) = \sin(2\pi t)$ and $g(t) \equiv 1$. Parameter values are $\varepsilon = 0.04$ and $\sigma = 0.025$. The shaded region is the set $\mathcal{B}(h)$ for $h = 3$, centred at the deterministic solution $x_t^{\rm det}$ starting at the same point as x_t. The width of $\mathcal{B}(h)$ is of order $h\sigma/\sqrt{a^\star(t)}$. Proposition 2.1 states that the probability of x_t leaving $\mathcal{B}(h)$ before times of order 1 decays roughly like $\mathrm{e}^{-h^2/2}$.

Note that $\mathcal{B}(h)$ coincides, up to order ε, with the set of points where $V(x,t)$ is smaller than $V(x_t^{\rm det},t) + (\frac{1}{2}h\sigma g(t))^2$. Showing that the path $\{x_s\}_{0\leqslant s\leqslant t}$ is likely to remain in $\mathcal{B}(h)$ is equivalent to showing that the *first-exit time* of x_s from $\mathcal{B}(h)$, defined by

$$\tau_{\mathcal{B}(h)} = \inf\{s \geqslant 0 \colon (x_s, s) \notin \mathcal{B}(h)\}, \tag{2.17}$$

is unlikely to be smaller than t. In fact, the following probabilities are equivalent (the superscripts refer to the initial condition):

$$\begin{aligned}\mathbb{P}^{0,x_0}\left\{\tau_{\mathcal{B}(h)} < t\right\} &= \mathbb{P}^{0,x_0}\{\exists s \in [0,t) \colon (x_s,s) \notin \mathcal{B}(h)\} \\ &= \mathbb{P}^{0,x_0}\left\{\sup_{0\leqslant s<t} \frac{|x_s - x_s^{\rm det}|}{\sqrt{\bar v(s)}} \geqslant h\right\}.\end{aligned} \tag{2.18}$$

The following result is a straightforward consequence of standard exponential bounds on the supremum of stochastic integrals, extended to integrals as appearing in (2.15). The proof, given in [BG1, proposition 3.4] for constant g, also applies here[5].

Proposition 2.1. *For all t and $h > 0$,*

$$\mathbb{P}^{0,x_0}\left\{\tau_{\mathcal{B}(h)} < t\right\} \leqslant C(t,\varepsilon)\mathrm{e}^{-\kappa h^2} \tag{2.19}$$

where

$$C(t,\varepsilon) = \frac{|\alpha^\star(t)|}{\varepsilon^2} + 2 \qquad \textit{and} \qquad \kappa = \tfrac{1}{2} - \mathcal{O}(\varepsilon). \tag{2.20}$$

Note that only the prefactor $C(t,\varepsilon)$ is time dependent. The exponential factor $\mathrm{e}^{-\kappa h^2}$ is small as soon as $h \gg 1$, so that the paths $\{x_s\}_{0\leqslant s\leqslant t}$ are concentrated in a neighbourhood of order $\sqrt{\bar v(s)}$ of the deterministic solution up to time t (see figure 2). More precisely, paths are unlikely to leave the strip $\mathcal{B}(h)$ of width $h\sqrt{\bar v}$ before time t provided $h^2 \gg \log C(t,\varepsilon)$.

The bound (2.19) is useful for times significantly shorter than Kramers' time, which is of order $\varepsilon\mathrm{e}^{2H/\sigma^2}$ (to reach points where the potential has value H). For longer times, the prefactor $C(t,\varepsilon)$ becomes sufficiently large to counteract the term $\mathrm{e}^{-\kappa h^2}$ for any reasonable h, which is

[5] The generalization of the proof to g bounded away from zero is trivial, but the result also holds if g vanishes. In fact, a sufficient condition is that $\bar v(t)/\bar v(s) = 1 + \mathcal{O}(\varepsilon)$ whenever $t - s = \mathcal{O}(\varepsilon^2)$, which can be checked using the asymptotic expansion (2.13).

natural, as we cannot expect paths to remain concentrated near $x_t^{\rm det}$ on such long timescales. On polynomial timescales of order σ^{-k}, however, large excursions are very unlikely.

The estimate (2.19) has been designed to yield an optimal exponent for noise intensities scaling as a power of ε. We do not expect the prefactor $C(t,\varepsilon)$ to be optimal, but for times and noise intensities polynomial in ε, it leads to subexponential corrections. However, if we do not care for the precise exponent, (2.20) can be replaced by

$$C(t,\varepsilon) = \frac{|\alpha^\star(t)|}{\varepsilon} + 2 \qquad \text{and} \qquad \kappa > 0. \tag{2.21}$$

The denominator ε in C is due to the fact that we work in slow time.

2.2. Nonlinear case

We consider now the motion in a more general, nonquadratic potential of the form

$$V(x,t) = \tfrac{1}{2}c(t)(x - x^\star(t))^2 + \mathcal{O}((x - x^\star(t))^3) \tag{2.22}$$

admitting a local minimum at $x^\star(t)$. As before, we assume that the curvature $c(t)$ is bounded below by a positive constant for all times t. We do not exclude, however, that V has other (possibly deeper) potential wells than the one at $x^\star(t)$.

The probability density can no longer be computed exactly in general, although it seems plausible that a distribution initially concentrated near $x^\star(0)$ will remain concentrated in metastable equilibrium near $x^\star(t)$, on a certain timescale. In fact, proposition 2.1 naturally extends to the nonlinear case.

Consider first the deterministic case $\sigma = 0$. A result due to Gradšteĭn and Tihonov [Gr,Ti], which is related to the adiabatic theorem of quantum mechanics, states that

- there exists a particular solution $\bar{x}_t^{\rm det}$ of the deterministic equation $\varepsilon\dot{x} = -\partial_x V(x,t)$ tracking the bottom of the potential well at a distance of order ε;
- any solution $x_t^{\rm det}$ starting in a neighbourhood of $x^\star(0)$ (in fact, inside the potential well) approaches $\bar{x}_t^{\rm det}$ exponentially fast in t/ε.

Let us fix a deterministic initial condition $x_0 = x_0^{\rm det}$ such that $x_t^{\rm det}$ is attracted by $\bar{x}_t^{\rm det}$. We introduce the notations

$$a(t) = -\frac{\partial^2 V}{\partial x^2}(x_t^{\rm det}, t) \qquad \alpha(t,s) = \int_s^t a(u)\,{\rm d}u \qquad \text{and} \qquad \alpha(t) = \alpha(t,0) \tag{2.23}$$

for the curvature of the potential at $x_t^{\rm det}$ and the analogue quantities to (2.7). Note that Tihonov's result implies that $a(t)$ asymptotically approaches $-c(t) + \mathcal{O}(\varepsilon)$. The difference $y_t = x_t - x_t^{\rm det}$ satisfies an SDE of the form

$${\rm d}y_t = \frac{1}{\varepsilon}[a(t)y_t + b(y_t,t)]\,{\rm d}t + \frac{\sigma}{\sqrt{\varepsilon}}g(t)\,{\rm d}W_t \tag{2.24}$$

where $b(y,t) = \mathcal{O}(y^2)$ describes the effect of nonlinearity. We define again a strip $\mathcal{B}(h)$ as in (2.16), with

$$\bar{v}(t) = \bar{v}(0){\rm e}^{2\alpha(t)/\varepsilon} + \frac{\sigma^2}{\varepsilon}\int_0^t {\rm e}^{2\alpha(t,s)/\varepsilon} g(s)^2\,{\rm d}s. \tag{2.25}$$

Our results work for any $\bar{v}(0)$ larger than a positive constant independent of ε. A convenient choice is $\bar{v}(0) = \sigma^2 g(0)^2/(2|a^\star(0)|)$: then the fact that $x_t^{\rm det}$ approaches exponentially fast a neighbourhood of order ε of $x^\star(t)$ implies that

$$\bar{v}(t) = \sigma^2\left[\frac{g(t)^2}{2|a^\star(t)|} + \mathcal{O}(\varepsilon) + \mathcal{O}(|x_0 - x^\star(0)|{\rm e}^{-{\rm const}\ t/\varepsilon})\right]. \tag{2.26}$$

Proposition 2.1 generalizes to:

Theorem 2.2. *There exists a constant h_0, independent of σ and ε, such that for all $h \leqslant h_0/\sigma$,*

$$\mathbb{P}^{0,x_0}\left\{\tau_{\mathcal{B}(h)} < t\right\} \leqslant C(t,\varepsilon)\mathrm{e}^{-\kappa h^2} \tag{2.27}$$

where

$$C(t,\varepsilon) = \frac{|\alpha(t)|}{\varepsilon^2} + 2 \qquad \textit{and} \qquad \kappa = \tfrac{1}{2} - \mathcal{O}(\varepsilon) - \mathcal{O}(\sigma h). \tag{2.28}$$

The interpretation is the same as in the linear case: paths are concentrated, for times significantly shorter than Kramers' time, in a strip of width proportional to $\sqrt{\bar{v}(t)}$ around x_t^{det}.

The proof is identical to the one of [BG1, theorem 2.4]. The main idea is to show that if the solution of the equation linearized around x_t^{det} remains in a strip $\mathcal{B}(h)$, then the solution of the nonlinear equation (2.24) almost surely remains in the slightly larger strip $\mathcal{B}(h[1 + \mathcal{O}(\sigma h)])$.

The main difference between the nonlinear and the linear case is the condition $h \leqslant h_0/\sigma$, which stems from the requirement that the linear term $a(t)y_t$ in equation (2.24) should dominate the nonlinear term $b(y_t, t)$ for all $(x_t, t) \in \mathcal{B}(h)$. Because of this condition, the result (2.27) is useful for $\sigma^2 \ll \kappa h_0^2/\log C(t,\varepsilon)$. It is, however, possible to derive bounds for larger deviations under additional assumptions on the potential:

Proposition 2.3. *Assume that there are constants $L_0 > 0$, $K > 0$ and $n \geqslant 2$ such that*

$$x\frac{\partial V}{\partial x}(x,t) \geqslant K|x|^n \tag{2.29}$$

whenever $|x| \geqslant L_0$ and $t \geqslant 0$. Then there exist constants $C, \kappa > 0$ such that

$$\mathbb{P}^{0,x_0}\left\{\sup_{0\leqslant s\leqslant t}|x_s| \geqslant L\right\} \leqslant C\left(\frac{t}{\varepsilon}+1\right)\mathrm{e}^{-\kappa L^n/\sigma^2} \tag{2.30}$$

for all $t \geqslant 0$, $L \geqslant L_0$ and $|x_0| \leqslant L_0/2$.

This result is a generalization of [BG3, proposition 4.3], where the case $n = 4$ was treated.

We remark in passing that the bounds (2.27) and (2.30) are sufficient to provide estimates on the moments of the distribution of x_t, without solving the Fokker–Planck equation (cf [BG3, corollary 4.6]):

Corollary 2.4. *Assume that (2.29) holds for some $n \geqslant 2$. Then*

$$\mathbb{E}^{0,x_0}\{|x_t - x_t^{\mathrm{det}}|^{2k}\} \leqslant (2k-1)!!M^k\bar{v}(t)^k \tag{2.31}$$

for some constant M, all integers k and all $t \geqslant 0$, provided $\sigma \leqslant c_0/\log(1 + t/\varepsilon)$ for a sufficiently small constant c_0.

Note that the Cauchy–Schwarz inequality immediately implies bounds on odd moments as well. The bounds on the moments are those of a Gaussian distribution with variance $M\bar{v}(t)$, even if the potential V has multiple wells. The reason is that on the timescale under consideration, solutions of the SDE do not have enough time to cross a potential barrier and reach another potential well. If V has deeper potential wells than the one at $x^\star(t)$, the metastable distribution of x_t will differ radically from the stationary distribution of the frozen system, which has most of its mass concentrated in the deepest potential well.

2.3. Escape from a saddle

Assume now that the potential $V(x,t)$ admits a saddle at $x^\star(t)$ for all times under consideration. In the deterministic case, a particular solution $\hat{x}_t^{\mathrm{det}}$ is known to track the saddle at a distance of order ε, separating the basins of attraction of two neighbouring potential wells. Trajectories starting near $\hat{x}_t^{\mathrm{det}}$ will depart from it exponentially fast, but if the initial separation $|x_0 - \hat{x}_0^{\mathrm{det}}|$

is exponentially small, the time required to reach a distance of order one from the saddle may be quite long.

Noise will help kicking x_t away from $\hat{x}_t^{\mathrm{det}}$, and thus reduce the time necessary to leave a neighbourhood of the saddle. In order to describe this effect, we consider the deviation $y_t = x_t - \hat{x}_t^{\mathrm{det}}$, which satisfies the SDE

$$\mathrm{d}y_t = \frac{1}{\varepsilon}[\hat{a}(t)y_t + \hat{b}(y_t, t)]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}g(t)\,\mathrm{d}W_t \tag{2.32}$$

where $\hat{a}(t) \geqslant a_0 > 0$ is the curvature of the potential at $\hat{x}_t^{\mathrm{det}}$, and $\hat{b}(y,t) = \mathcal{O}(y^2)$. We now describe the dynamics in a small neighbourhood of the adiabatic solution tracking the saddle, where diffusion prevails over drift[6], defined by

$$\mathcal{B}(h) = \left\{(x,t)\colon |x - \hat{x}_t^{\mathrm{det}}| < \frac{h\sigma g(t)}{\sqrt{2\hat{a}(t)}}\right\}. \tag{2.33}$$

The following result, which is proved in the same way as [BG1, proposition 3.10], shows that the first-exit time $\tau_{\mathcal{B}(h)}$ of x_t from $\mathcal{B}(h)$ is likely to be small.

Theorem 2.5. *Assume that g is bounded below by $L\varepsilon$ for a sufficiently large constant L. Then for all $h \leqslant 1$ and all initial conditions $(x_0, 0) \in \mathcal{B}(h)$,*

$$\mathbb{P}^{0,x_0}\{\tau_{\mathcal{B}(h)} > t\} \leqslant C\exp\left\{-\frac{\kappa}{h^2}\frac{1}{\varepsilon}\int_0^t \hat{a}(s)\,\mathrm{d}s\right\} \tag{2.34}$$

where C, κ are positive constants.

This result shows that x_t will leave a neighbourhood of size $\sigma g(t)/\sqrt{2\hat{a}(t)}$ of $\hat{x}_t^{\mathrm{det}}$ typically after a time of order $\varepsilon/\hat{a}(0)$. Once this neighbourhood has been left, the drift term starts prevailing over the diffusion term, and one can show, (although this is not trivial,) that the typical time needed to leave a neighbourhood of order one of the saddle is of order $\varepsilon|\log\sigma|$. We will state a similar result in section 5 when discussing the dynamics after passing through a pitchfork bifurcation point.

3. Stochastic resonance

Up to now, we have considered situations in which the potential has bounded curvature near its (isolated) extrema, so that for sufficiently small noise intensities and not too long timescales, most paths are concentrated near the bottom of the well they started in.

Not surprisingly, interesting phenomena occur when the condition on the curvature is violated. Two cases can be considered:

- *Avoided bifurcation*: the potential well becomes flatter, but the curvature does not vanish completely; sufficiently strong noise, however, may drive solutions to another potential well. This mechanism is responsible in particular for the phenomenon of *stochastic resonance*.
- *Bifurcation*: the curvature at the bottom of the well vanishes, say at time 0; for $t > 0$, new potential wells may be created (e.g. pitchfork bifurcation) or not (e.g. saddle–node bifurcation).

[6] A possible way to compare the influence of drift and noise is by Itô's formula, which yields, for $\hat{b} = 0$, $\mathrm{d}(y_t^2) = (1/\varepsilon)[2\hat{a}(t)y_t^2 + \sigma^2 g(t)^2]\,\mathrm{d}t + (2\sigma/\sqrt{\varepsilon})g(t)y_t\,\mathrm{d}W_t$. The expression in brackets is dominated by the noise term $\sigma^2 g(t)^2$ for all (x_t, t) in $\mathcal{B}(1)$, while the deterministic term $2\hat{a}(t)y_t^2$ prevails outside this domain.

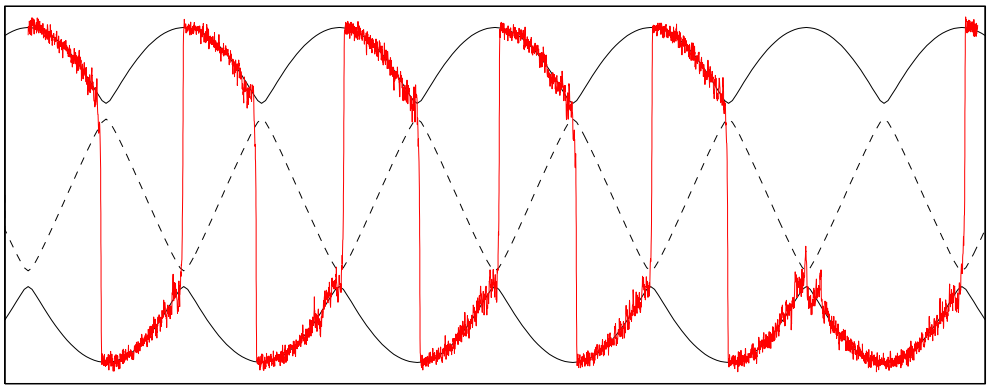

Figure 3. A sample path of the equation of motion (3.2) in an asymmetrically perturbed double-well potential. Parameter values are $\varepsilon = 0.0025$, $a_0 = 0.005$ and $\sigma = 0.065$, which is just above the threshold for transitions to be likely. The upper and lower full curves show the locations of the potential wells, while the broken curve marks the location of the saddle.

We will discuss the possible phenomena in the case of a Ginzburg–Landau potential

$$V(x,t) = \tfrac{1}{4}x^4 - \tfrac{1}{2}\mu(t)x^2 - \lambda(t)x. \tag{3.1}$$

However, the precise form of the potential and the fact that it has at most two wells are not essential.

The potential (3.1) has two wells if $27\lambda^2 < 4\mu^3$ and one well if $27\lambda^2 > 4\mu^3$. Crossing the lines $27\lambda^2 = 4\mu^3$, $\mu > 0$, corresponds to a saddle–node bifurcation, and crossing the point $\lambda = \mu = 0$ corresponds to a pitchfork bifurcation. Equilibrium points are solutions of the equation $x^3 - \mu(t)x - \lambda(t) = 0$; we will denote stable equilibria by $x^\star_\pm$, and the saddle, when present, by $x^\star_0$.

In this section, we investigate situations with avoided bifurcations, in which the potential always has two wells, but the barrier between them becomes low periodically. Bifurcation phenomena will be discussed in the next two sections.

3.1. The mechanism of stochastic resonance

Let us consider the case where μ is a positive constant, say $\mu = 1$, and $\lambda(t)$ varies periodically, say $\lambda(t) = -A\cos(2\pi t)$. If $|\lambda| < \lambda_c = 2/(3\sqrt{3})$, then V is a double-well potential. We thus assume that $A < \lambda_c$.

In the absence of noise, the existence of the potential barrier prevents the solutions from switching between potential wells. If noise is present, but there is no periodic driving ($A = 0$), solutions will cross the potential barrier at random times, whose expectation is given by Kramers' time $\varepsilon e^{2H/\sigma^2}$, where H is the height of the barrier ($H = 1/4$ in this case).

Interesting things happen when both noise and periodic driving $\lambda(t)$ are present. Then the potential barrier will still be crossed at random times, but with a higher probability near the instants of minimal barrier height (i.e. when t is an integer or half-integer). This phenomenon produces peaks in the power spectrum of the signal, hence the name *stochastic resonance* (SR).

If the noise intensity is sufficiently large compared to the minimal barrier height, transitions become likely twice per period (back and forth), so that the signal x_t is close, in some sense, to a periodic function (figure 3). The amplitude of this oscillation may be considerably larger than the amplitude of the forcing $\lambda(t)$, so that the mechanism can be used to amplify weak periodic

signals. This phenomenon is also known as *noise-induced synchronization* [SNA, NSAS]. Of course, too large noise intensities will spoil the quality of the signal.

The mechanism of SR was originally introduced as a possible explanation of the close-to-periodic appearance of the major Ice Ages [BSV, BPSV]. Here the (quasi-)periodic forcing is caused by variations in the Earth's orbital parameters (Milankovich factors), and the additive noise models the fast unpredictable fluctuations caused by the weather. Meanwhile, SR has been detected in a large number of systems (see, for instance, [MW, WM, GHM] for reviews), including ring lasers, electronic devices and even the sensory system of crayfish and paddlefish [N&].

Despite the many applications of SR, its mathematical description has remained incomplete for two decades, although several limiting cases have been studied in detail. The first approaches considered either potentials that are piecewise constant in time [BSV], or two-level systems with discrete space [ET, McNW]. Continuous-time equations have been mainly investigated through the Fokker–Planck equation, using methods from spectral theory [Fox, JH1] or linear response theory [JH2]. The main contribution of these approaches is an estimation of the signal-to-noise ratio (SNR) of the power spectrum, as a function of the noise intensity. The SNR is one of the possible quantitative measures of the signal's periodicity, and behaves roughly as $\mathrm{e}^{-H/\sigma^2}/\sigma^4$, which is maximal for $\sigma^2 = H/2$. In [MS2], an action functional is used to extend Kramers' result to the case of small-amplitude forcing.

A description of individual paths has been given for the first time in Freidlin's recent paper [Fr1]. His results apply to a general class of n-dimensional potentials, in the case where the period $1/\varepsilon$ of the driving scales as Kramers' time e^{2H/σ^2}. The fact that the minimal barrier height H is considered as constant while σ tends to zero, implies that the results only hold for exponentially long driving periods. As quantitative measure of the signal's periodicity, the L^p-distance[7] between paths $\{x_t\}_{t\geqslant 0}$ and a deterministic, periodic limit function $\phi(t)$ is used. This limit function simply tracks the bottom of a potential well, and jumps to the deeper well each time the potential barrier becomes lowest. The L^p-distance is shown to converge to zero in probability as σ goes to zero. However, Freidlin's techniques do not yield estimates on the speed of this convergence, or its dependence on p.

Our techniques allow us to provide such estimates for one-dimensional potentials, with a more natural distance than the L^p-distance: in fact, we simply use the geometrical distance between the paths and the limit function, considered as curves in the (t, x)-plane. The analysis given below also includes situations in which the minimal barrier height becomes small in the small-noise limit.

3.2. Pathwise description

For the Ginzburg–Landau potential (3.1), $\mu = 1$ and $\lambda(t) = -A\cos(2\pi t)$, the SDE takes the form

$$\mathrm{d}x_t = \frac{1}{\varepsilon}[x_t - x_t^3 - A\cos(2\pi t)]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t. \tag{3.2}$$

We assume that $A < \lambda_c$, so that there are always two stable equilibria at $x_\pm^\star(t)$ and a saddle at $x_0^\star(t)$. We introduce a parameter $a_0 = \lambda_c - A$ which measures the minimal barrier height: at $t = 0$, the barrier height is of order $a_0^{3/2}$ for small a_0, and the distance between $x_+^\star$ and the saddle at $x_0^\star$ is of order $\sqrt{a_0}$. At $t = \frac{1}{2}$, the left-hand potential well at $x_-^\star$ is likewise close to the

[7] The L^p-distance between x_t and $\phi(t)$ is the integral of $\|x_t - \phi(t)\|^p$ over a given time interval (raised to the power $1/p$). Note that in contrast to a small supremum norm, a small L^p-norm of $x_t - \phi(t)$ does not exclude that x_t makes large excursions away from $\phi(t)$, as long as these excursions are sufficiently short.

saddle. In order for transitions to become possible on a timescale which is not exponentially large, we allow a_0 to become small with ε.

Assume that we start at time $t_0 = -1/4$ in the basin of attraction of the right-hand potential well. Results from section 2 show that transitions are unlikely for $t \ll 0$. Also, for $0 \ll t \ll 1/2$, paths will be concentrated either near $x_+^\star$ or near $x_-^\star$. This allows us to define the *transition probability* as

$$P_{\text{trans}} = \mathbb{P}^{t_0,x_0}\left\{x_{t_1} < 0\right\} \qquad t_0 = -1/4,\ t_1 = 1/4. \tag{3.3}$$

The properties of P_{trans} do not depend sensitively of the choices of t_0 and t_1, as long as $-1/2 \ll t_0 \ll 0 \ll t_1 \ll 1/2$. Also the level 0 can be replaced by any level lying between $x_-^\star(t)$ and $x_+^\star(t)$ for all t. Denoting by $a \vee b$ the maximum of two real numbers a and b, our main result can be formulated as follows.

Theorem 3.1 ([BG2, theorems 2.6 and 2.7]). *For the noise intensity, there is a threshold level $\sigma_c = (a_0 \vee \varepsilon)^{3/4}$ with the following properties:*

(1) *If $\sigma < \sigma_c$, then*

$$P_{\text{trans}} \leqslant \frac{C}{\varepsilon} \mathrm{e}^{-\kappa\sigma_c^2/\sigma^2} \tag{3.4}$$

for some $C, \kappa > 0$. Paths are concentrated in a strip of width $\sigma/(\sqrt{|t|} \vee \sigma_c^{1/3})$ around the deterministic solution tracking $x_+^\star(t)$ (figure 4(a)).

(2) *If $\sigma > \sigma_c$, then*

$$P_{\text{trans}} \geqslant 1 - C\mathrm{e}^{-\kappa\sigma^{4/3}/(\varepsilon|\log\sigma|)} \tag{3.5}$$

for some $C, \kappa > 0$. Transitions are concentrated in the interval $[-\sigma^{2/3}, \sigma^{2/3}]$. Moreover, for $t \leqslant -\sigma^{2/3}$, paths are concentrated in a strip of width $\sigma/\sqrt{|t|}$ around the deterministic solution tracking $x_+^\star(t)$, while for $t \geqslant \sigma^{2/3}$, they are concentrated in a strip of width $\sigma/\sqrt{t}$ around a deterministic solution tracking $x_-^\star(t)$ (figure 4(b)).

The crossover is quite sharp: for $\sigma \ll \sigma_c$, transitions between potential wells are very unlikely, while for $\sigma \gg \sigma_c$, they are very likely. By 'concentrated in a strip of width w', we mean that the probability that a path leaves a strip of width hw decreases like $\mathrm{e}^{-\kappa h^2}$ for some $\kappa > 0$. The 'typical width w' is our measure of the deviation from the deterministic periodic function, which tracks one potential well in the small-noise case, and switches back and forth between the wells in the large-noise case.

Theorem 3.1 implies in particular that for the periodic signal's amplification to be optimal, the noise intensity σ should exceed the threshold σ_c. Larger noise intensities will increase both spreading of paths (especially just before they cross the potential barrier) and the size of transition window, and thus spoil the output's periodicity.

The results of theorem 3.1 remain valid when the parameter a_0 controlling the minimal barrier height is not small. Note that (3.4) allows for an appreciable transition probability as soon as $\mathrm{e}^{\kappa\sigma_c^2/\sigma^2}$ is smaller than a constant times the period $1/\varepsilon$. This is consistent with the well known resonance condition of the small-amplitude forcing regime, which states that the noise intensity is optimal when the Kramers time equals half the driving period. Thus, the exponent $\kappa\sigma_c^2$ should converge to twice the barrier height in the limit $A \to 0$. For a minimal height of order 1, the optimal noise intensity behaves as $1/|\log\varepsilon|^{1/2}$. Unless the driving period is exponentially large, at optimal noise level transitions from the shallower well to the deeper one will already occur some time before the barrier is lowest. With increasing noise intensity, transitions in the 'wrong' direction, namely back into the shallower well, become less unlikely.

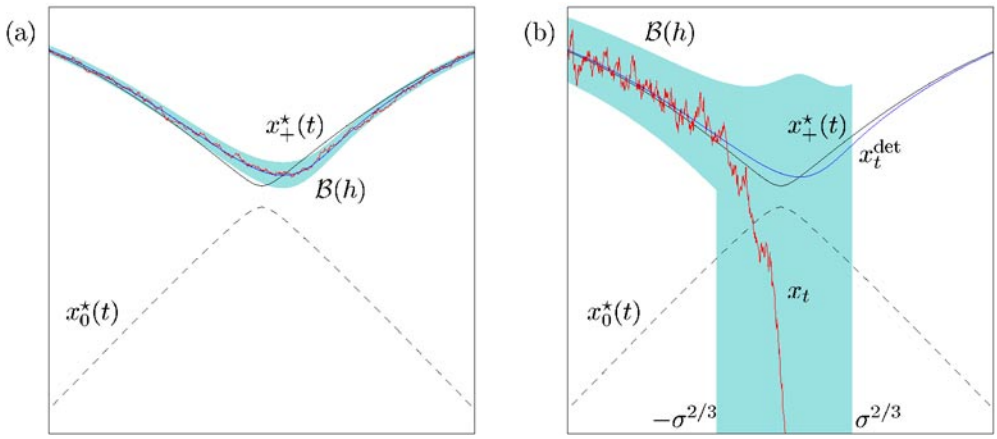

Figure 4. Sample paths of equation (3.2) in a neighbourhood of time 0, when the barrier height is minimal, for two different noise intensities. Full curves mark the location $x_+^\star(t)$ of the right-hand potential well, broken curves the location $x_0^\star(t)$ of the saddle. Parameter values are $\varepsilon = 0.0125$, $a_0 = 0.002$ and (*a*) $\sigma = 0.012$, (*b*) $\sigma = 0.07$. In case (*a*), the path remains in the set $\mathcal{B}(h)$, shown here for $h = 3$, which is centred at the deterministic solution $x_t^{\rm det}$. In case (*b*), x_t remains in $\mathcal{B}(h)$ as long as the width of $\mathcal{B}(h)$ is smaller than the distance between $x_t^{\rm det}$ and the saddle, that is, for $t \ll -\sigma^{2/3}$. The path jumps to the left-hand potential well during the time interval $[-\sigma^{2/3}, \sigma^{2/3}]$.

Hence larger driving amplitudes improve the quality of the response because they allow one to work at lower noise intensities.

Let us return to the case of small a_0. We observe that for $\sigma_c \ll \sigma \ll 1$, paths are concentrated in the right-hand well when $\sin(2\pi t) < 0$, and in the left-hand well when $\sin(2\pi t) > 0$. They switch between wells near integer and half-integer times, when the barrier is lowest. The distribution of x_t is thus shifted in time with respect to the stationary density $\mathrm{e}^{-2V(x,t)/\sigma^2}/N$ of the frozen system, which has most of its mass concentrated in the deeper well and therefore favours the left-hand well whenever $\cos(2\pi t) > 0$, and favours the right-hand well whenever $\cos(2\pi t) < 0$. Since paths may jump from the shallower to the deeper well at a time of order $\sigma^{2/3}$ before the instant of lowest potential barrier, increasing σ reduces the time during which the system is in metastable equilibrium in the shallower potential well. This phenomenon has also been observed in [Ta], where the probability density at a fixed time t is approximated by its projection on the first two eigenfunctions of the Fokker–Planck operator: for sufficiently strong noise, the density tracks the instantaneous stationary density, but it lags behind for weaker noise intensity.

It has been proposed to identify SR and synchronization with a phase-locking mechanism [NSAS], but the definition of a phase remained problematic. The fact that the majority of paths are contained in small space–time sets makes it possible to associate a random phase with those paths. For instance, when $\sigma > \sigma_c$, most paths circle the origin of the (λ, x)-plane counterclockwise (see figure 8(*c*)). For all paths not containing the origin, one can define a random phase φ_t by $\tan(\varphi_t) = x_t/\lambda(t)$. Although φ_t fluctuates, it is likely to increase at an average rate of 2π per cycle.

Part of our results should appear quite natural. If $a_0 > \varepsilon$, the threshold noise level $\sigma_c = a_0^{3/4}$ behaves like the square root of the minimal barrier height, which is consistent with the SNR being optimal for $\sigma^2 = H/2$. However, σ_c saturates at $\varepsilon^{3/4}$ for all $a_0 \leqslant \varepsilon$. Hence, even driving amplitudes arbitrarily close to λ_c cannot increase the transition probability. This is a rather subtle dynamical effect, mainly due to the fact that even if the barrier vanishes at

$t = 0$, it is lower than $\varepsilon^{3/2}$ during too short a time interval for paths to take advantage. The situation is the same as if there were an 'effective potential barrier' of height proportional to σ_c^2.

Another remarkable fact is that for $\sigma > \sigma_c$, neither the transition probability nor the width of the transition windows depend on the driving amplitude to leading order. In the remainder of this section, we are going to explain some ideas of the proof of theorem 3.1, which will hopefully clarify some of the surprising properties given above.

The first step is to understand the behaviour of the solutions of (3.2) in the deterministic case $\sigma = 0$. This problem belongs to the field of dynamical bifurcations, the theory of which is relatively well developed. We follow here a framework allowing one to determine scaling laws of solutions near bifurcation points, which has been presented in [BK]. The main idea is that when approaching a bifurcation point $(t^\star, x^\star)$, the distance between the adiabatic solution and the static equilibrium branch scales as $\varepsilon/|t - t^\star|^\rho$ for $t \leqslant t^\star - \varepsilon^\nu$, and as ε^μ for $|t - t^\star| \leqslant \varepsilon^\nu$. The rational numbers ρ, ν and $\mu = 1 - \rho\nu$ are universal exponents, which can be deduced from the Newton polygon of the bifurcation point.

Tihonov's theorem implies that away from the avoided bifurcation point at $t = 0$, solutions $x_t^{\det}$ of the deterministic equation track the equilibrium branch $x_+^\star(t)$ adiabatically at a distance of order ε. It is thus sufficient to understand what happens in a small neighbourhood of the almost-bifurcation. Note that if $\lambda = \lambda_c$, the right-hand well and the saddle merge at $x = 1/\sqrt{3}$. It is thus helpful to consider the translated variable $y_t^{\det} = x_t^{\det} - 1/\sqrt{3}$, which obeys the differential equation

$$\varepsilon \dot{y} = ct^2 - \sqrt{3}y^2 + a_0 + \text{higher order terms} \qquad c = 2\pi^2\lambda_c. \tag{3.6}$$

Consider the 'worst' case $a_0 = 0$. Then the right-hand side of (3.6) describes a transcritical bifurcation between the equilibria $y_{+,0}^\star = \pm\sqrt{c}3^{-1/4}|t| + \mathcal{O}(t^2)$. One shows that the solution $y_t^{\det}$ tracks $y_+^\star(t)$ at a distance scaling as $\varepsilon/|t|$ for $|t| \geqslant \sqrt{\varepsilon}$, and as $\sqrt{\varepsilon}$ for $|t| \leqslant \sqrt{\varepsilon}$. In fact, $y_t^{\det}$ never approaches the saddle closer than a distance of order $\sqrt{\varepsilon}$, which is because the term $-\sqrt{3}y^2$ only dominates during a short time interval of order $\sqrt{\varepsilon}$.

The same qualitative behaviour holds for $0 < a_0 \leqslant \varepsilon$. For $a_0 > \varepsilon$, one can show that $x_t^{\det}$ tracks $x_+^\star(t)$ at a distance never exceeding $\mathcal{O}(\varepsilon/\sqrt{a_0})$. Since $x_+^\star(0) - x_-^\star(0)$ is of order $\sqrt{a_0}$, $x_t^{\det}$ never approaches the saddle closer than a distance of order $\sqrt{a_0}$.

Let us now consider the random motion near $x^{\det}$ for $\sigma > 0$. We denote, as usual, by $a(t)$ the curvature of the potential at the deterministic solution $x_t^{\det}$. By (3.6), $a(t)$ behaves, near $t = 0$, as $-y_t^{\det}$, which we know to behave as $-(|t| \vee \sqrt{a_0})$ for $a_0 > \varepsilon$, and as $-(|t| \vee \sqrt{\varepsilon})$ for $a_0 \leqslant \varepsilon$. It turns out that theorem 2.2 can be extended to the present situation, to show that paths are concentrated in a strip around $x_t^{\det}$. The width of this strip is again related to the standard deviation of a linearized process, and behaves as

$$\frac{\sigma}{\sqrt{|a(t)|}} \asymp \frac{\sigma}{\sqrt{|t|} \vee \sigma_c^{1/3}}. \tag{3.7}$$

However, this property only holds under the condition that the spreading is always smaller than the distance between $x_t^{\det}$ and the saddle, which scales as $|t| \vee \sigma_c^{2/3}$. We thus have to require $\sigma < |t|^{3/2} \vee \sigma_c$, so that

- if $\sigma < \sigma_c$, the condition is always satisfied, and thus (3.4) follows from the generalization of theorem 2.2 with $h = \sigma_c/\sigma$ (figure 4(*a*));
- if $\sigma > \sigma_c$, the condition is only satisfied for $t \leqslant -\sigma^{2/3}$ (figure 4(*b*)).

It thus remains to understand what happens for $t \geqslant -\sigma^{2/3}$ if $\sigma > \sigma_c$. Here the main idea is that during the time interval $[-\sigma^{2/3}, \sigma^{2/3}]$, the process x_t has a certain number of trials to reach the saddle. If x_t reaches the saddle, it has roughly probability $1/2$ of moving in each direction. If it moves far enough towards the left, it will most probably fall into the left-hand

potential well, and is unlikely to come back for the remaining half-period. If it moves to the right, it has failed to overcome the barrier, but can try again during the next excursion.

One can define a typical time Δt for an excursion as the time needed for a path starting near $x_+^\star$ to reach and overcome the barrier with non-negligible probability, say with probability $1/3$. One can show, by comparison with suitable linearized processes, that this typical time is determined by the condition

$$|\alpha(t, t+\Delta t)| = \int_t^{t+\Delta t} |a(s)|\,\mathrm{d}s = \text{const}\,\varepsilon|\log\sigma|. \tag{3.8}$$

To obtain this, one first checks that the curvature of the potential is the same, up to sign reversal, at the adiabatic solutions tracking the bottom of the well and the saddle. To overcome the saddle, the integral in (3.8) must be of order ε, similarly as in theorem 2.5. The factor $|\log\sigma|$ is needed to reach a distance of order 1 from the saddle. Now the maximal number N of trials during the interval $[-\sigma^{2/3}, \sigma^{2/3}]$ is given by

$$N = \text{const}\frac{|\alpha(\sigma^{2/3}, -\sigma^{2/3})|}{\varepsilon|\log\sigma|} \geqslant \text{const}\frac{\sigma^{4/3}}{\varepsilon|\log\sigma|}. \tag{3.9}$$

Finally, the Markov property implies that the probability *not* to overcome the saddle during N trials is bounded by

$$\left(\frac{2}{3}\right)^N = \exp\left\{-N\log\left(\frac{3}{2}\right)\right\} \leqslant \exp\left\{-\text{const}\frac{\sigma^{4/3}}{\varepsilon|\log\sigma|}\right\} \tag{3.10}$$

which proves (3.5). An important point to note is that the transition probability is not determined by the curvature of the potential at the saddle, but by the curvature at the deterministic solution tracking the saddle, which may be larger for small a_0.

3.3. Symmetric potentials

Another case of interest is the Ginzburg–Landau potential (3.1) with $\lambda \equiv 0$ and $\mu(t) = a_0 + 1 - \cos(2\pi t)$, $a_0 > 0$. Then $V(x,t)$ is always symmetric, with minima at $x_\pm^\star(t) = \pm\sqrt{\mu(t)}$ and a barrier height $\frac{1}{4}\mu(t)^2$ becoming small at integer times. The associated SDE is

$$\mathrm{d}x_t = \frac{1}{\varepsilon}[(a_0 + 1 - \cos(2\pi t))x_t - x_t^3]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t. \tag{3.11}$$

As before, transitions between potential wells are most likely when the barrier is lowest. We can thus define a transition probability as in (3.3), with $-1 \ll t_0 \ll 0 \ll t_1 \ll 1$. We again assume that the process starts in the right-hand well. In this case, the result corresponding to theorem 3.1 is as follows:

Theorem 3.2 ([BG2] theorems 2.2–2.4). *There is a threshold noise level $\sigma_c = a_0 \vee \varepsilon^{2/3}$ with the following properties:*

(1) *If $\sigma < \sigma_c$, then*

$$P_{\text{trans}} \leqslant \frac{C}{\varepsilon}\mathrm{e}^{-\kappa\sigma_c^2/\sigma^2} \tag{3.12}$$

for some $C, \kappa > 0$. Paths are concentrated in a strip of width $\sigma/(|t| \vee \sqrt{\sigma_c})$ around the deterministic solution tracking $x_+^\star(t)$.

(2) *If $\sigma > \sigma_c$, then*

$$P_{\text{trans}} \geqslant \tfrac{1}{2} - C\mathrm{e}^{-\kappa\sigma^{3/2}/(\varepsilon|\log\sigma|)} \tag{3.13}$$

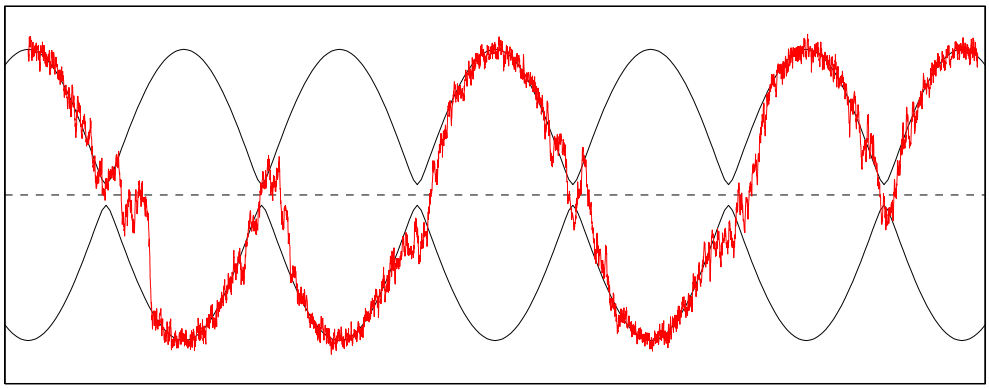

Figure 5. A sample path of the equation of motion (3.11) in a periodically modulated symmetric double-well potential. Parameter values are $\varepsilon = 0.005$, $a_0 = 0.005$ and $\sigma = 0.075$, which is above the threshold for transitions to occur with probability close to $1/2$. The upper and lower full curves show the locations of the potential wells, while the broken line marks the location of the saddle.

for some $C, \kappa > 0$. Transitions are concentrated in the interval $[-\sqrt{\sigma}, \sqrt{\sigma}\,]$. Moreover, for $t \leqslant -\sqrt{\sigma}$, paths are concentrated in a strip of width $\sigma/|t|$ around the deterministic solution tracking $x_+^\star(t)$, while for $t \geqslant \sqrt{\sigma}$, they are concentrated in a strip of width σ/t around a deterministic solution tracking either $x_+^\star(t)$ or $x_-^\star(t)$.

The main difference with respect to the previous case is that due to the symmetry, P_{trans} can never exceed $1/2$. The limiting process obtained by letting σ go to zero but keeping $\sigma > \sigma_c$ is no longer a deterministic function, but a 'Bernoulli' process, choosing between the left and the right potential well with probability $1/2$ at integer times (figure 5).

Another difference lies in the distribution of barrier crossing times in the transition window. In the asymmetric case, paths may overcome the saddle as soon as $t \geqslant -\sigma^{2/3}$, and are unlikely to return to the shallower well. In the symmetric case, paths may jump back and forth between both wells up to time $\sqrt{\sigma}$, before settling for a potential well.

3.4. Modulated noise intensity

Other mechanisms leading to SR have been examined, for instance periodic forcing which is not deterministic, but affects the noise intensity, a situation arising in power amplifiers [D&]. This case can be analysed by the same method as the previous ones, but the results are quite different.

Let us consider the motion in a static symmetric double-well potential, described by the SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon}[\mu_0 x_t - x_t^3]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}} g(t)\,\mathrm{d}W_t \tag{3.14}$$

where $\mu_0 > 0$ is fixed and $g(t)$ is periodic. We assume that $g(t) \geqslant \varepsilon|\log\sigma|/\mu_0$ for all t. Theorem 2.2 shows that for sufficiently weak noise, paths starting at time t_0 at the bottom $\sqrt{\mu_0}$ of the right-hand potential well remain concentrated in a strip around $\sqrt{\mu_0}$, with width proportional to $\sigma g(t)/(2\sqrt{\mu_0})$. This holds as long as the spreading is smaller than a constant times the distance $\sqrt{\mu_0}$ between well and saddle. The probability to cross the saddle before time t is bounded by

$$P(t) = C(t,\varepsilon)\exp\left\{-\frac{\kappa}{\sigma^2}\left(\frac{2\mu_0}{\hat{g}(t)}\right)^2\right\} \qquad \text{where } \hat{g}(t) = \sup_{t_0\leqslant s\leqslant t} g(s) \quad (3.15)$$

and $\kappa > 0$. Thus if g reaches its maximum $\hat{g}(0)$ at time 0, the probability of seeing a transition during one period satisfies

$$P_{\text{trans}} \leqslant C(1,\varepsilon)\mathrm{e}^{-\kappa\sigma_c^2/\sigma^2} \qquad \text{for } \sigma \leqslant \sigma_c = \frac{2\mu_0}{\hat{g}(0)}. \quad (3.16)$$

Note that here, as the potential is static, the threshold value for the noise intensity can be guessed from Kramers' time, assuming constant g. Taking into account that g is *not* necessarily constant, we see that for $\sigma > \sigma_c$, transitions are likely to happen in the time interval during which $\sigma g(t) \geqslant 2\mu_0$. For instance, if $g(t)$ behaves quadratically near its unique maximum, this transition window is given by

$$t^2 \leqslant \text{const}\; g(0)\left(1 - \frac{\sigma_c}{\sigma}\right). \quad (3.17)$$

In contrast to the previous cases, however, the transition times are less concentrated, in the sense that for times $t_0 < t_1 < t_2$ before the transition window,

$$P(t_1) \simeq P(t_2)^{(\hat{g}(t_2)/\hat{g}(t_1))^2}. \quad (3.18)$$

A similar argument as in the previous cases shows that for $\sigma > \sigma_c$,

$$P_{\text{trans}} \geqslant \frac{1}{2} - \text{const}\;\exp\left\{-\kappa\frac{2\mu_0\Delta}{\varepsilon|\log\sigma|}\right\} \quad (3.19)$$

where Δ is the length of the transition window.

If the potential is made asymmetric, so that a constant term $\lambda_0 > 0$ is added to the drift term in (3.14), the critical noise intensities needed to reach the saddle from the shallower left-hand well and the deeper right-hand well will be different (one can check that their ratio is $1 + \mathcal{O}(\lambda_0/\mu_0^{3/2})$). As a consequence, transitions from the shallower to the deeper well will be likely as soon as the noise intensity exceeds the smaller threshold, while transitions in both directions are likely when the noise intensity exceeds the larger threshold. When the noise intensity drops below the smaller threshold again, x_t will be in the deeper right-hand well with larger probability. The net effect is that x_t will visit both wells near integer times, but has only a small probability of remaining in the shallow well after the transition window. Thus the periodic signal is not amplified in the same way as discussed before, but nevertheless we observe an amplification mechanism which allows one to read off at which times the threshold is exceeded.

4. Hysteresis

Hysteresis is another characteristic phenomenon of bistable systems. Let us consider again the motion in the Ginzburg–Landau potential (3.1) with $\mu \equiv 1$ and $\lambda(t) = -A\cos(2\pi t)$, but without imposing the restriction $A < \lambda_c$. In the deterministic case $\sigma = 0$ the equation of motion reads

$$\varepsilon\frac{\mathrm{d}x_t}{\mathrm{d}t} = x_t - x_t^3 + \lambda(t). \quad (4.1)$$

We may ask the question: How does x_t behave, as a function of $\lambda(t)$, in the adiabatic limit $\varepsilon \to 0$? Intuitively, x_t will always track the bottom of a potential well. A naive way to see this is to formally set ε equal to zero in (4.1): we obtain the algebraic equation $x - x^3 + \lambda = 0$, which admits three branches of solutions (see figure 6); $X_+^\star(\lambda)$ and $X_-^\star(\lambda)$ correspond to potential wells, and $X_0^\star(\lambda)$ to a saddle, which exists only for $|\lambda| < \lambda_c$.

If the amplitude A is smaller than the critical value λ_c, there are always two potential wells separated by a barrier. Hence x_t will always track the bottom of the same well in the limit $\varepsilon \to 0$, so that the instantaneous value of λ is sufficient to determine the state (provided we know in which potential well the process started).

If A is larger than λ_c, however, a saddle–node bifurcation point is crossed whenever $|\lambda|$ reaches λ_c from below: the potential well tracked by x_t disappears, so that x_t jumps to the other well, which is unaffected by the bifurcation (figure 1). As a result, the state x_t is not uniquely defined by the instantaneous value of λ if $|\lambda| \leqslant \lambda_c$: x_t tracks the bottom $X_-^\star(\lambda(t))$ of the left-hand well if λ increases, and the bottom $X_+^\star(\lambda(t))$ of the right-hand well if λ decreases. This phenomenon is called *hysteresis*. The hysteresis cycle consists of the branches $X_\pm^\star(\lambda)$, $|\lambda| \leqslant A$, and two vertical lines on which $|\lambda| = \lambda_c$. It encloses an area $\mathcal{A}_0 = 3/2$, called the *static hysteresis area*. In many applications, x and λ are thermodynamically conjugated variables, and the hysteresis area represents the energy dissipation per cycle.

4.1. Dynamical hysteresis and scaling laws

Consider now what happens when ε is small but positive. The solutions of equation (4.1) will not react instantaneously to changes in the potential, so that the shape of hysteresis cycles is modified. It is important to understand the ε-dependence of quantities such as the average of x_t over one period, the value of λ when x_t changes sign, and the area enclosed by the hysteresis cycle. It is known that there are constants $\gamma_1 > \gamma_0 > 0$ such that the following properties hold:

- If $A \leqslant \lambda_c + \gamma_0\varepsilon$, then x_t cannot change sign (except during the very first period, if the process does not start near the bottom of a well). There exist two stable periodic orbits, one tracking each potential well (figure 6(*a*)). Each encloses an area $\mathcal{A}$ satisfying
$$\mathcal{A} \asymp A\varepsilon, \tag{4.2}$$
and the average of x_t over each cycle is nonzero. The notation $\asymp$ is a shorthand to indicate that $c_- A\varepsilon \leqslant \mathcal{A} \leqslant c_+ A\varepsilon$ for some constants $c_\pm > 0$ independent of ε and A.
- If $A \geqslant \lambda_c + \gamma_1\varepsilon$, then x_t changes sign twice per period. All orbits are attracted by the same periodic orbit (figure 6(*b*)), corresponding to a hysteresis cycle with zero average and area $\mathcal{A}$ satisfying
$$\mathcal{A} - \mathcal{A}_0 \asymp \varepsilon^{2/3}(A - \lambda_c)^{1/3}. \tag{4.3}$$
When x_t changes sign, the parameter λ satisfies $|\lambda| - \lambda_c \asymp \varepsilon^{2/3}(A - \lambda_c)^{1/3}$.
- If $\lambda_c + \gamma_0\varepsilon < A < \lambda_c + \gamma_1\varepsilon$, several hysteresis cycles may coexist, some of them satisfying (4.2) and others satisfying (4.3).

The scaling law (4.3) was first derived in [JGRM] for $A - \lambda_c$ of order 1, where equation (4.1) was used to model a bistable laser. The case where A is close to λ_c has been analysed in [BK].

An equation qualitatively similar to (4.1) describes the dynamics of a Curie–Weiss model of a ferromagnet, subject to a periodic magnetic field $\lambda(t)$, in the limit of infinitely many spins [Mar]. The transition between the small and large amplitude regimes has been called 'dynamic phase transition' in [TO].

The magnetization obeys a deterministic differential equation only in the limit of infinite system size. The effect of the number N of spins being finite can be modelled, in first approximation, by an additive white noise of intensity proportional to $1/\sqrt{N}$ [Mar]. It is thus of major importance to understand the effect of additive noise on the properties of hysteresis cycles.

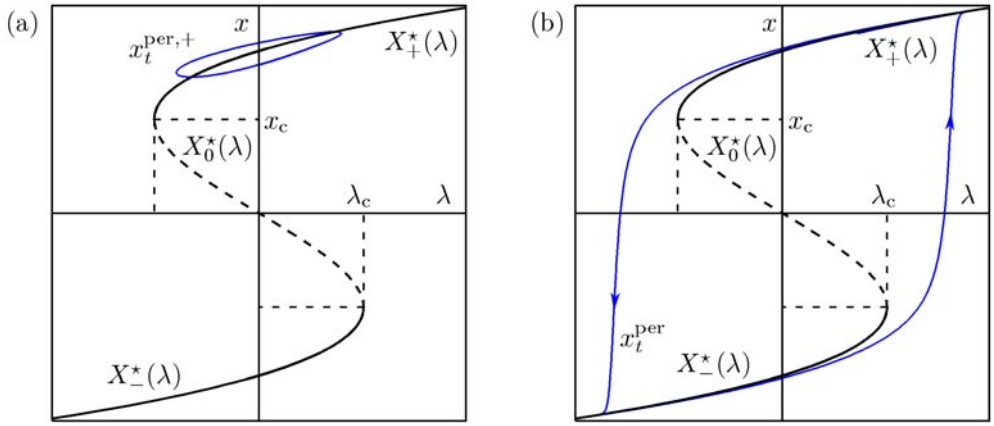

Figure 6. Periodic solutions of the deterministic equation (4.1), (*a*) in a case where the amplitude A of $\lambda(t)$ is smaller than λ_c, and (*b*) in a case where it is larger than λ_c. The enclosed area scales as εA in case (*a*), and like $\mathcal{A}_0 + \varepsilon^{2/3}(A-\lambda_c)^{1/3}$ in case (*b*), where $\mathcal{A}_0$ is the static hysteresis area. Potential wells $X_\pm^*(\lambda)$ are displayed as full curves, the saddle $X_0^\star(\lambda)$ as a broken curve.

Langevin equations have already been studied for multi-dimensional Ginzburg–Landau potentials. Then, however, the mechanism leading to hysteresis is different, because there is no potential barrier between stable states. Numerical simulations [RKP] suggested that the area of hysteresis cycles should follow the scaling law $\mathcal{A} \asymp \varepsilon^{1/3}A^{2/3}$, while various theoretical arguments indicate that $\mathcal{A} \asymp \varepsilon^{1/2}A^{1/2}$ [DT, SD, ZZ]. It is not clear whether such a scaling law exists for the Ising model [SRN].

For clarity, we will continue interpreting x_t as the magnetization and $\lambda(t)$ as the magnetic field. There exist, however, many other instances where the dynamics is described by a periodically forced Langevin equation. For instance, in models for the Atlantic thermohaline circulation, x_t represents the salinity difference between high and middle latitude, and $\lambda(t)$ represents the atmospheric freshwater flux [St, Ra]. The effect of additive noise on this system has been investigated, for instance in [Ce], while the properties of hysteresis cycles were considered in particular in [Mo].

The fact that additive noise may create relaxation oscillations has been discussed in [Fr2], where the motion of a light particle in a randomly perturbed field is investigated with the help of large-deviation theory.

4.2. The effect of additive noise

We consider the Langevin equation

$$\mathrm{d}x_t = \frac{1}{\varepsilon}[x_t - x_t^3 - A\cos(2\pi t)]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t \tag{4.4}$$

where $A > 0$. We denote $A - \lambda_c$ by a_0, but in contrast to section 3 (where a_0 had opposite sign), we do not impose that a_0 is a small parameter, and we allow positive as well as negative a_0. Let us fix a deterministic initial condition ($t_0 = -1/2$, $x_0 > 0$), such that the solution x_t^{det} of the deterministic equation (4.1) with $x_{t_0}^{\mathrm{det}} = x_0$ is attracted by the right-hand potential well.

We are interested in the quantity

$$\mathcal{A}(\varepsilon,\sigma) = -\int_{-1/2}^{1/2} x_t\lambda'(t)\,\mathrm{d}t \tag{4.5}$$

measuring the area enclosed by x_t in the (λ, x)-plane during one period (x_t does not necessarily form a closed loop, but $\mathcal{A}$ still represents the energy dissipation). $\mathcal{A}(\varepsilon, 0)$ is the area enclosed by x_t^{det}, and behaves as (4.2) or (4.3). For positive σ, $\mathcal{A}(\varepsilon, \sigma)$ is a random variable, the distribution of which we want to characterize.

Another random quantity of interest is the value λ^0 of the magnetic field when x_t changes sign for the first time:

$$\lambda^0 = \lambda(\tau^0) \qquad \tau^0 = \inf\{t > t_0 : x_t \leqslant 0\}. \tag{4.6}$$

Results from section 3 already allow us to make some predictions for the case $a_0 < 0$. For $\sigma \ll \sigma_c = (|a_0| \vee \varepsilon)^{3/4}$, x_t is unlikely to switch between potential wells, so that $\mathcal{A}(\varepsilon, \sigma)$ will be concentrated near the deterministic value $\mathcal{A}(\varepsilon, 0)$, which is of order ε (figure 8(*a*)). For $\sigma \gg \sigma_c$, x_t is likely to cross the potential barrier at a random time τ^0 which typically behaves as $-\sigma^{2/3}$. The corresponding field λ^0 behaves as $\lambda_c - \sigma^{4/3}$ (figure 8(*c*)). Thus additive noise of sufficient intensity will lead to a hysteresis area which is smaller, by an amount of order $\sigma^{4/3}$, than the static area $\mathcal{A}_0$.

The same behaviour can be shown to hold for positive a_0 up to order ε. In this case, the potential barrier vanishes during a short time interval, which is too short, however, for x_t to notice. For $a_0 > \varepsilon$, there is a similar transition between a small-noise regime (figure 8(*b*)), where x_t is likely to track the deterministic solution, and a large-noise regime, where it typically crosses the potential barrier some time before the barrier vanishes. The threshold value of σ delimiting both regimes is again deduced from the variance of the equation linearized around x^{det}, and turns out to be $\sigma_c = (\varepsilon\sqrt{a_0})^{1/2}$. For $\sigma > \sigma_c$, the typical value λ^0 of the field when x_t changes sign is again found to behave as $\lambda_c - \sigma^{4/3}$.

We thus obtain the existence of three distinct parameter regimes, with qualitatively different behaviour of typical hysteresis cycles. We summarize the main results in the following theorem, and give some additional details afterwards. Many estimates contain logarithmic dependencies on a_0, σ and ε. In order not to overburden notations, we will assume that σ and a_0 behave like a power of ε (a_0 may also be a constant), and denote $|\log \varepsilon|$ by ℓ_ε. The regimes are those indicated in figure 7, but some results are only valid if we exclude a logarithmic layer near the boundary, for instance case 2 corresponds to $a_0 > \gamma_1\varepsilon$ and $\sigma \leqslant \text{const}\ (\varepsilon\sqrt{a_0})^{1/2}/\ell_\varepsilon$.

Theorem 4.1 ([BG3, theorems 2.3–2.5]).

- *Case 1: (Small-amplitude regime). The distribution of the random area $\mathcal{A}(\varepsilon, \sigma)$ is concentrated near the deterministic value $\mathcal{A}(\varepsilon, 0) \asymp A\varepsilon$. There are two subcases to consider:*
  - *In case 1a, $\mathcal{A}(\varepsilon, \sigma)$ can be written as the sum of a Gaussian random variable with variance of order $\sigma^2\varepsilon$, centred at $\mathcal{A}(\varepsilon, 0)$, and a random remainder. The remainder has expectation and standard deviation of order $\sigma^2\ell_\varepsilon$ at most.*
  - *In case 1b, the distribution of $\mathcal{A}(\varepsilon, \sigma)$ is more spread out. The expectation and standard deviation of $\mathcal{A}(\varepsilon, \sigma) - \mathcal{A}(\varepsilon, 0)$ are at most of order $\sigma^2\ell_\varepsilon$, which may exceed $\mathcal{A}(\varepsilon, 0)$.*
- *Case 2: (Large-amplitude regime). The distribution of $\mathcal{A}(\varepsilon, \sigma)$ is concentrated near the deterministic value $\mathcal{A}(\varepsilon, 0)$ which satisfies (4.3).*
  - *In case 2a, $\mathcal{A}(\varepsilon, \sigma)$ can be written as the sum of a Gaussian random variable with variance of order $\sigma^2(\varepsilon\sqrt{a_0})^{1/3}$, centred at $\mathcal{A}(\varepsilon, 0)$, and a random remainder. The remainder has expectation and standard deviation of order $\sigma^2\ell_\varepsilon(\varepsilon\sqrt{a_0})^{-2/3}$ at most.*
  - *In case 2b, we can only show that the distribution of $\mathcal{A}(\varepsilon, \sigma)$ is concentrated in an interval of width $(\varepsilon\sqrt{a_0})^{2/3}\ell_\varepsilon$ around $\mathcal{A}(\varepsilon, 0)$.*

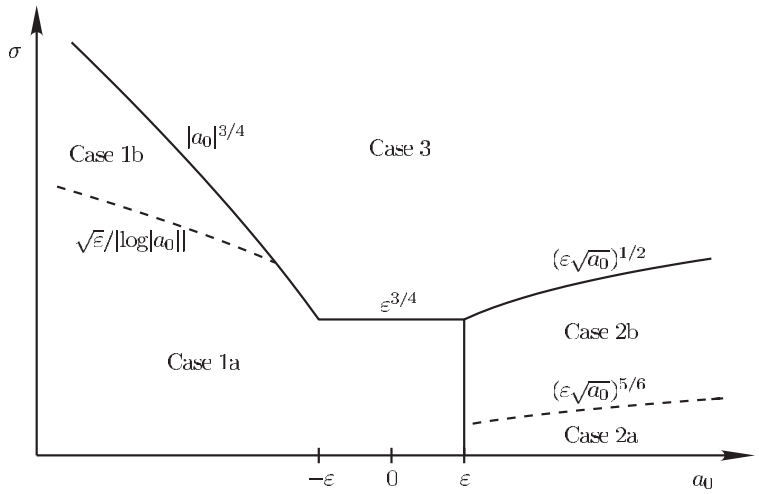

Figure 7. Definition of the parameter regimes for hysteresis cycles, shown in the plane ($a_0 = A - \lambda_c, \sigma$) for a fixed value of ε. The behaviour of the hysteresis area $\mathcal{A}(\varepsilon, \sigma)$ in each regime is described in theorem 4.1. Typical hysteresis cycles are shown in figure 8.

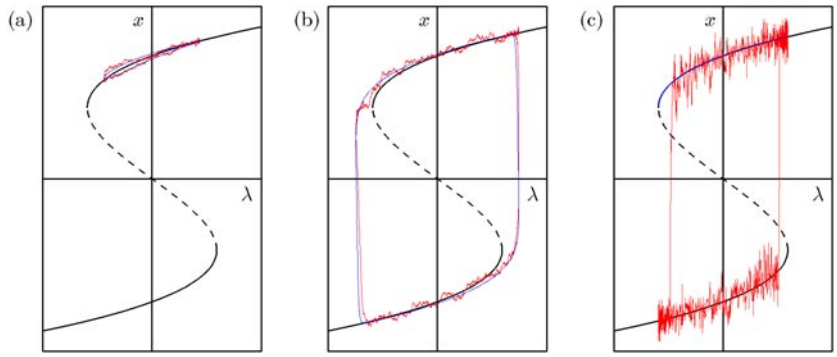

Figure 8. Typical random hysteresis 'cycles' in the three parameter regimes of figure 7. Deterministic solutions are shown for comparison. (*a*) Case 1, small-amplitude regime (here $\varepsilon = 0.05$, $a_0 = -0.1$, $\sigma = 0.025$): paths typically stay close to the deterministic cycle, which tracks a potential well. (*b*) Case 2, large-amplitude regime (here $\varepsilon = 0.02$, $a_0 = 0.1$, $\sigma = 0.05$): typical paths are close to the deterministic cycle, which switches between potential wells. (*c*) Case 3, large-noise regime (here $\varepsilon = 0.001$, $a_0 = 0$, $\sigma = 0.16$): paths are likely to cross the potential barrier when $|\lambda|$ is of order $\lambda_c - \sigma^{4/3}$.

- *Case 3: (Large-noise regime). The distribution of $\mathcal{A}(\varepsilon, \sigma)$ is concentrated near a (deterministic) reference area $\hat{\mathcal{A}}$ satisfying $\hat{\mathcal{A}} - \mathcal{A}_0 \asymp -\sigma^{4/3}$. The standard deviation of $\mathcal{A}(\varepsilon, \sigma)$ is at most of order $\sigma^{4/3}\ell_\varepsilon^{2/3}$, and its expectation belongs to an interval*

$$[\hat{\mathcal{A}} - \mathcal{O}(\sigma^{4/3}\ell_\varepsilon^{2/3}), \hat{\mathcal{A}} + \mathcal{O}(\sigma^2\ell_\varepsilon^2) + \mathcal{O}(\varepsilon\ell_\varepsilon)]. \tag{4.7}$$

  In the case where $a_0 \geqslant \varepsilon$ and $\sigma \leqslant a_0^{3/4}$, the term $\mathcal{O}(\varepsilon\ell_\varepsilon)$ has to be replaced by $\mathcal{O}(\varepsilon\sqrt{|a_0|}\,\ell_\varepsilon/\sigma^{2/3})$. In both cases, the distribution decays faster to the right of $\hat{\mathcal{A}}$ than to the left.

In regimes 1 and 2, the main effect of additive noise is to broaden the distribution of the area, which remains concentrated, however, around the corresponding deterministic value. In

regime 3, on the other hand, the hysteresis area obeys a completely new scaling law, which is determined by the noise intensity rather than by the frequency and amplitude of the driving field.

The Gaussian behaviour of $\mathcal{A}$ in cases 1a and 2a is obtained in the following way. The deviation $y_t = x_t - x_t^{\mathrm{det}}$ from the deterministic solution satisfies an equation of the form (2.24) with $g \equiv 1$, whose solution obeys the integral equation

$$y_t = \frac{\sigma}{\sqrt{\varepsilon}} \int_{t_0}^{t} \mathrm{e}^{\alpha(t,s)/\varepsilon} \, \mathrm{d}W_s + \frac{1}{\varepsilon} \int_{t_0}^{t} \mathrm{e}^{\alpha(t,s)/\varepsilon} b(y_s, s) \, \mathrm{d}s, \tag{4.8}$$

where $b(y, s) = \mathcal{O}(y^2)$. For small values of y_t, the first term dominates the second one. Its contribution to $\mathcal{A}(\varepsilon, \sigma) - \mathcal{A}(\varepsilon, 0) = -\int y_u \lambda'(u) \, \mathrm{d}u$ can be written as

$$\frac{\sigma}{\sqrt{\varepsilon}} \int_{t_0}^{t_0+1} \gamma(t_0 + 1, s) \, \mathrm{d}W_s \qquad \text{where } \gamma(t, s) := -\int_s^t \mathrm{e}^{\alpha(u,s)/\varepsilon} \lambda'(u) \, \mathrm{d}u. \tag{4.9}$$

The variance of this term is given by

$$\sigma^2 \varepsilon \Gamma(t_0 + 1, t_0) \qquad \text{where } \Gamma(t, t_0) := \frac{1}{\varepsilon^2} \int_{t_0}^{t} \gamma(t, s)^2 \, \mathrm{d}s. \tag{4.10}$$

The integral $\Gamma(t_0 + 1, t_0)$ depends only on properties of the deterministic solution x_t^{det} via the curvature $a(t)$. The auxiliary function $\gamma(t, s)$ can be evaluated by integration by parts, its leading term behaving as $-\varepsilon\lambda'(s)/|a(s)|$.

In case 1, $\Gamma(t_0 + 1, t_0)$ is of order 1, and thus the contribution of the linear term to the variance of the area is of order $\sigma^2\varepsilon$. In case 1a, one can show that the Gaussian term dominates the distribution of $\mathcal{A}(\varepsilon, \sigma)$ near $\mathcal{A}(\varepsilon, 0)$, in the sense that

$$\mathbb{P}\{|\mathcal{A}(\varepsilon, \sigma) - \mathcal{A}(\varepsilon, 0)| \geqslant H\} \leqslant \frac{C}{\varepsilon} \mathrm{e}^{-\kappa H^2/(\sigma^2\varepsilon)} \tag{4.11}$$

holds for some constants $C, \kappa > 0$, and for all H smaller than a constant times $\sqrt{\varepsilon}(|a_0| \vee \varepsilon)^{4/3}$ if $|a_0| \leqslant \varepsilon^{2/3}/\ell_\varepsilon^{4/3}$, and all H smaller than $\varepsilon/\ell_\varepsilon$ if $|a_0| \geqslant \varepsilon^{2/3}/\ell_\varepsilon^{4/3}$. Note that the upper bound (4.11) is exponentially small for the maximal value of H, except on the upper boundary of region 1a.

In case 1b, the Gaussian term no longer dominates, but one can still show that

$$\mathbb{P}\{|\mathcal{A}(\varepsilon, \sigma) - \mathcal{A}(\varepsilon, 0)| \geqslant H\} \leqslant \frac{C}{\varepsilon} \mathrm{e}^{-\kappa H/(\sigma^2\ell_\varepsilon)} \tag{4.12}$$

up to $H = \text{ const } |a_0|^{3/2}\ell_\varepsilon$. Again, (4.12) is exponentially small except on the upper boundary of region 1b.

Estimates (4.11) and (4.12) control the tails of the distribution of $\mathcal{A}(\varepsilon, \sigma)$ in a neighbourhood of $\mathcal{A}(\varepsilon, 0)$. The quartic growth of the potential $V(x, t)$ for large $|x|$ implies, on the other hand, that

$$\mathbb{P}\{|\mathcal{A}(\varepsilon, \sigma) - \mathcal{A}(\varepsilon, 0)| \geqslant H\} \leqslant \frac{C}{\varepsilon} \mathrm{e}^{-\kappa H^4/\sigma^2} \tag{4.13}$$

for all H larger than some constant (of order 1). In fact, this estimate holds in *all* parameter regimes, since it does not depend on the details of the potential near $x = 0$. This still leaves a gap between the domains of validity of (4.11) and (4.12), and of (4.13), which is due to the existence of a second potential well. In fact, the distribution of the hysteresis area will not be unimodal. Sample paths are unlikely to cross the potential barrier, but if they do so, then most probably near the instants of minimal barrier height, in which case they enclose an area of order 1. Hence the density of $\mathcal{A}(\varepsilon, \sigma)$ will have a large peak near $\mathcal{A}(\varepsilon, 0)$, and a small peak near areas of order 1 (more precisely, near $\int_{-A}^{A} (X_+^\star(\lambda) - X_-^\star(\lambda)) \, \mathrm{d}\lambda$), (see figure 9).

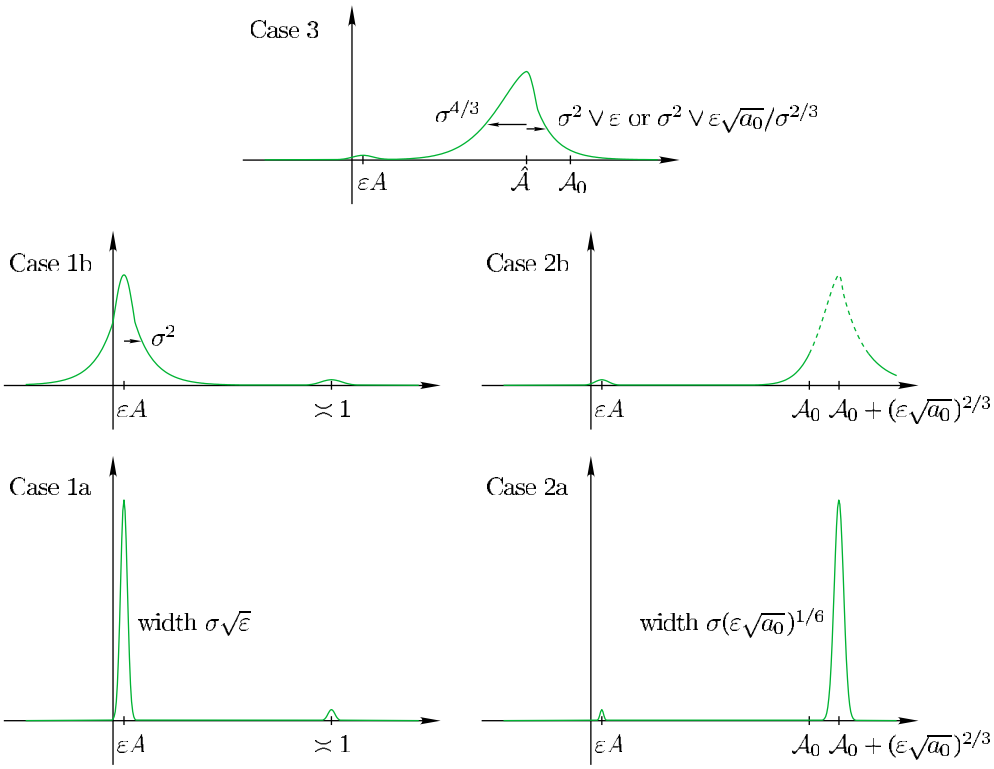

Figure 9. Sketches of the distribution of the hysteresis area $\mathcal{A}(\varepsilon, \sigma)$ in the different parameter regimes. Regime 1a: the area is concentrated near the deterministic area $\mathcal{A}(\varepsilon, 0) \asymp A\varepsilon$. There is a small probability of observing areas of order 1. Regime 1b: the distribution is more spread out. Regime 2a: the area is concentrated near the deterministic area $\mathcal{A}(\varepsilon, 0)$ of order $\mathcal{A}_0 + (\varepsilon\sqrt{a_0})^{2/3}$. Regime 2b: the area is concentrated in an interval of width $(\varepsilon\sqrt{a_0})^{2/3}$ around $\mathcal{A}_0 + (\varepsilon\sqrt{a_0})^{2/3}$, but we do not control the distribution in this interval. The broken curve shows an extrapolation of the estimates outside this interval. Regime 3: the area is concentrated near $\hat{\mathcal{A}}$, which is of order $\mathcal{A}_0 - \sigma^{4/3}$. The distribution decays faster to the right than to the left.

In case 2a, the distribution of $\mathcal{A}(\varepsilon, \sigma)$ near $\mathcal{A}(\varepsilon, 0)$ is again dominated by a Gaussian, stemming from the linearization of the SDE around x_t^{det}. The integral $\Gamma(t_0 + 1, t_0)$ in (4.10) is found to behave like $\varepsilon^{-2/3} a_0^{1/6}$, leading to a variance of order $\sigma^2(\varepsilon\sqrt{a_0})^{1/3}$ and to the bound

$$\mathbb{P}\{|\mathcal{A}(\varepsilon, \sigma) - \mathcal{A}(\varepsilon, 0)| \geqslant H\} \leqslant \frac{C}{\varepsilon} \mathrm{e}^{-\kappa H^2/(\sigma^2(\varepsilon\sqrt{a_0})^{1/3})} \tag{4.14}$$

valid for H smaller than a constant times $\varepsilon\sqrt{a_0}$.

Unfortunately, this estimate cannot be extended to case 2b. The reason is that during the jump of x_t^{det} to the left-hand potential well, a zone of instability is crossed where paths are strongly dispersed. The maximal spreading of paths is of order $\sigma(\varepsilon\sqrt{a_0})^{-5/6}$, which is too large, in this regime, to allow for a precise control of the effect of nonlinear terms. This does not exclude the possibility that the bound (4.14) remains valid for larger σ.

However, the value λ^0 of the magnetic field when x_t changes sign can be described in all of regime 2. One can show that

$$\mathbb{P}\left\{|\lambda^0| < \lambda_c - L\right\} \leqslant \frac{C}{\varepsilon} \mathrm{e}^{-\kappa(|L|^{3/2} \vee \varepsilon\sqrt{a_0})/\sigma^2} \qquad \text{for} \quad -L_1(\varepsilon\sqrt{a_0})^{2/3} \leqslant L \leqslant L_0/\ell_\varepsilon \tag{4.15}$$

$$\mathbb{P}\left\{|\lambda^0| > \lambda_c + L\right\} \leqslant 3\mathrm{e}^{-\kappa L/(\sigma^2(\varepsilon\sqrt{a_0})^{2/3}\ell_\varepsilon)} \qquad \text{for} \quad L \geqslant L_2(\varepsilon\sqrt{a_0})^{2/3}. \tag{4.16}$$

Note that $|\lambda^0|$ cannot exceed $\lambda_c + a_0$. These bounds mean that the distribution of λ^0 is concentrated around the deterministic value of order $\lambda_c + (\varepsilon\sqrt{a_0})^{2/3}$, and decays faster to the right than to the left. They can be used to show that in case 2b,

$$\mathbb{P}\left\{\mathcal{A}(\varepsilon,\sigma) - \mathcal{A}(\varepsilon,0) \leqslant -H\right\} \leqslant \frac{C}{\varepsilon}\mathrm{e}^{-\kappa H^{3/2}/\sigma^2} \tag{4.17}$$

$$\mathbb{P}\left\{\mathcal{A}(\varepsilon,\sigma) - \mathcal{A}(\varepsilon,0) \geqslant +H\right\} \leqslant \frac{C}{\varepsilon}\mathrm{e}^{-\kappa(\varepsilon\sqrt{a_0})^{1/3}H/(\sigma^2\ell_\varepsilon)} \tag{4.18}$$

for $(\varepsilon\sqrt{a_0})^{2/3}\ell_\varepsilon \leqslant H \leqslant (\varepsilon\sqrt{a_0})^{1/3}\ell_\varepsilon$. Thus the probability that $\mathcal{A}(\varepsilon,\sigma) - \mathcal{A}(\varepsilon,0)$ is outside an interval of size $(\varepsilon\sqrt{a_0})^{2/3}\ell_\varepsilon$ is very small. We do not control, however, what happens inside this interval.

In case 3, the large-noise regime, most sample paths are driven over the potential barrier as soon as the magnetic field reaches a value of order $\lambda_c - \sigma^{4/3}$. Rather than comparing $\mathcal{A}(\varepsilon,\sigma)$ to its deterministic value, we should compare it to a reference area $\hat{\mathcal{A}}$ given by

$$\tfrac{1}{2}\hat{\mathcal{A}} = \int_{-1/4}^{t_1} x_s^{\mathrm{det},+}(-\lambda'(s))\,\mathrm{d}s + \int_{t_1}^{1/4} x_s^{\mathrm{det},-}(-\lambda'(s))\,\mathrm{d}s \tag{4.19}$$

where $x_s^{\mathrm{det},\pm}$ are solutions tracking, respectively, the right and left potential wells, and $t_1 \asymp -\sigma^{2/3}$ is the typical jump time. Checking that $\hat{\mathcal{A}} - \mathcal{A}_0$ scales as $-\sigma^{4/3}$ is straightforward.

The probability of deviations of $\mathcal{A}(\varepsilon,\sigma)$ from $\hat{\mathcal{A}}$ can be estimated by bounding separately the integrals between $-1/4$ and t_1 and between t_1 and $1/4$. The results differ slightly in two regimes. If $a_0 \leqslant \varepsilon$ or $\sigma > a_0^{3/4}$, then

$$\mathbb{P}\{\mathcal{A}(\varepsilon,\sigma) - \hat{\mathcal{A}} \leqslant -H\} \leqslant \frac{C}{\varepsilon}\mathrm{e}^{-\kappa H^{3/2}/\sigma^2} + \frac{3}{2}\mathrm{e}^{-\kappa\sigma^{4/3}/(\varepsilon\ell_\varepsilon)} \tag{4.20}$$

$$\mathbb{P}\{\mathcal{A}(\varepsilon,\sigma) - \hat{\mathcal{A}} \geqslant +H\} \leqslant \frac{C}{\varepsilon}\mathrm{e}^{-\kappa H/(\sigma^2\ell_\varepsilon)} + \frac{3}{2}\mathrm{e}^{-\kappa H/(\varepsilon\ell_\varepsilon)} \tag{4.21}$$

holds for some $C, \kappa > 0$ and all H up to a constant times $\sigma^{2/3}\ell_\varepsilon$. If $a_0 \geqslant \varepsilon$ and $\sigma \leqslant a_0^{3/4}$, two exponents are modified:

$$\mathbb{P}\{\mathcal{A}(\varepsilon,\sigma) - \hat{\mathcal{A}} \leqslant -H\} \leqslant \frac{C}{\varepsilon}\mathrm{e}^{-\kappa H^{3/2}/\sigma^2} + \frac{3}{2}\mathrm{e}^{-\kappa\sigma^2/(\varepsilon\sqrt{a_0}\ell_\varepsilon)} \tag{4.22}$$

$$\mathbb{P}\{\mathcal{A}(\varepsilon,\sigma) - \hat{\mathcal{A}} \geqslant +H\} \leqslant \frac{C}{\varepsilon}\mathrm{e}^{-\kappa H/(\sigma^2\ell_\varepsilon)} + \frac{3}{2}\mathrm{e}^{-\kappa\sigma^{2/3}H/(\varepsilon\sqrt{a_0}\ell_\varepsilon)}. \tag{4.23}$$

Again, the distribution decays faster to the right of $\hat{\mathcal{A}}$ than to the left, guaranteeing that $\mathcal{A}(\varepsilon,\sigma)$ is likely to be smaller than the static hysteresis area $\mathcal{A}_0$. The second term on the right-hand sides of (4.20) and (4.22) does not depend on H: it bounds the probability that the paths do not cross the potential barrier, and enclose an area close to zero. The situation is opposite to case 1: the distribution of the hysteresis area has a large peak near $\hat{\mathcal{A}}$ and a small peak near 0.

The qualitative behaviour of the distribution of $\mathcal{A}(\varepsilon,\sigma)$ is sketched in figure 9. When crossing the boundary between regimes 1 and 3, the probability of paths crossing the potential barrier increases, so that the peak near $\mathcal{A}(\varepsilon,0)$ shrinks while the peak near $\mathcal{A}_0$ grows. When approaching the transition line between regimes 2 and 3, the distribution of the area $\mathcal{A}(\varepsilon,\sigma)$ becomes more spread out and more symmetric, before concentrating again around the large-amplitude deterministic value of $\mathcal{A}_0 + \mathcal{O}((\varepsilon\sqrt{a_0})^{2/3})$.

5. Bifurcation delay

In the previous section, we had to deal in particular with the slow passage through a saddle–node bifurcation. This section is devoted to the slow passage through a (symmetric) pitchfork bifurcation.

We consider again the Ginzburg–Landau potential (3.1), but this time with $\lambda \equiv 0$, and a parameter $\mu(t)$ increasing monotonously through zero. As μ changes from negative to positive, the potential transforms from a single-well to a double-well potential, a scenario known as *spontaneous symmetry breaking*. In fact, the symmetry of the potential is not broken, but the symmetry of the state may be. Solutions tracking initially the potential well at $x = 0$ will choose between one of the new potential wells, but which one of the wells is chosen, and at what time, depends strongly on the noise present in the system.

5.1. Dynamic pitchfork bifurcation

In the deterministic case $\sigma = 0$, the equation of motion reads

$$\varepsilon \frac{\mathrm{d}x_t}{\mathrm{d}t} = \mu(t)x_t - x_t^3. \tag{5.1}$$

Its solution x_t^{det} with initial condition $x_{t_0}^{\mathrm{det}} = x_0 > 0$ can be written in the form

$$x_t^{\mathrm{det}} = c(x_0, t)\mathrm{e}^{\alpha(t,t_0)/\varepsilon} \qquad \alpha(t, t_0) = \int_{t_0}^{t} \mu(s)\,\mathrm{d}s \tag{5.2}$$

where the function $c(x_0, t)$ is found by substitution into (5.1). Its exact expression is of no importance here, it is sufficient to know that $0 < c(x_0, t) \leqslant x_0$ for all t.

Assume that $\mu(t)$ is negative for $t < 0$ and positive for $t > 0$. If we start at a time $t_0 < 0$, the solution (5.2) will be attracted exponentially fast by the stable origin. The function $\alpha(t, t_0)$ is negative and decreasing for $t_0 < t < 0$, which implies in particular that x_0^{det} is exponentially small. For $t > 0$, the function $\alpha(t, t_0)$ is increasing, but it remains negative for some time. As a consequence, x_t^{det} remains close to the saddle up to the first time $t = \Pi(t_0)$ for which $\alpha(t, t_0)$ reaches 0 again (if such a time exists). Shortly after time $\Pi(t_0)$, the solution will jump to the potential well at $+\sqrt{\mu(t)}$, unless x_0 is exponentially small. $\Pi(t_0)$ is called *bifurcation delay*, and depends only on μ and t_0. For instance, if $\mu(t) = t$, then $\alpha(t, t_0) = \frac{1}{2}(t^2 - t_0^2)$ and $\Pi(t_0) = |t_0|$.

The existence of a bifurcation delay may have undesired consequences. Assume for instance that we want to determine the bifurcation diagram of $\dot{x} = \mu x - x^3$ experimentally. Instead of measuring the asymptotic value of x_t for many different values of μ, which is time-consuming (especially near $\mu = 0$ where x_t decays only like $1/\sqrt{t}$), one may be tempted to vary μ slowly during the experiment. This, however, will fail to reveal part of the stable equilibrium branches because of the bifurcation delay. A similar phenomenon exists for the Hopf bifurcation [Ne1, Ne2].

The delay is due to the fact that x_t approaches the origin exponentially closely. Noise of sufficient intensity will help driving the particle away from the saddle, and should therefore reduce the bifurcation delay. The obvious question is thus: How does the delay depend on the noise intensity?

In the case of a quadratic potential $V(x, t) = -\frac{1}{2}\mu(t)x^2$, this question has been investigated and compared with experiments, with the result that the delay behaves as $\sqrt{|\log\sigma|}$ [TM, SMC, SHA]. Similar results were obtained in [JL]. Our techniques allow us to derive rigorous bounds for the nonlinear equation.

5.2. Effect of additive noise

We consider the nonlinear SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon}[\mu(t)x_t - x_t^3]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t. \tag{5.3}$$

The results presented here are a particular case of those obtained in [BG1], which apply to more general nonlinearities. We assume that $\mu(t) = t + \mathcal{O}(t^2)$ is monotonously increasing on an interval $[t_0, T]$ or $[t_0, \infty)$, $t_0 < 0$. We denote by x_t^{det} and x_t the solutions of the deterministic and stochastic equations with given initial condition $x_0 > 0$.

We already know that x_t^{det} decreases exponentially fast for $t < 0$. Results from section 2 show that paths are concentrated in a neighbourhood of order σ of x_t^{det} on any time interval $[t_0, t_1]$ bounded away from zero, so that we only need to worry about what happens after time t_1, when x^{det} is already exponentially small.

The dispersion of paths will be controlled by the variance-like function

$$\bar{v}(t) = \bar{v}_0 \mathrm{e}^{2\alpha(t,t_1)/\varepsilon} + \frac{\sigma^2}{\varepsilon}\int_{t_1}^{t} \mathrm{e}^{2\alpha(t,s)/\varepsilon}\,\mathrm{d}s \tag{5.4}$$

where $\bar{v}_0$ is a positive constant. One can show that this function grows as $\sigma^2/|\mu(t)|$ for $t \leqslant -\sqrt{\varepsilon}$, and remains of order $\sigma^2/\sqrt{\varepsilon}$ up to time $\sqrt{\varepsilon}$. Only after time $\sqrt{\varepsilon}$, $\bar{v}(t)$ grows exponentially fast. In analogy with (2.16), we define a strip

$$\mathcal{B}(h) = \left\{(x,t) : t_1 \leqslant t \leqslant \sqrt{\varepsilon}, |x - x_t^{\mathrm{det}}| < h\sqrt{\bar{v}(t)}\right\}. \tag{5.5}$$

In order to describe the behaviour for $t \geqslant \sqrt{\varepsilon}$, we further introduce the domain

$$\mathcal{D}(\varrho) = \left\{(x,t) : \sqrt{\varepsilon} \leqslant t \leqslant T, |x| \leqslant \sqrt{(1-\varrho)\mu(t)}\right\}, \tag{5.6}$$

where ϱ is a parameter in $[0, 2/3)$. Note that $\mathcal{D}(2/3)$ contains those points in space–time where the potential is concave, while $\mathcal{D}(0)$ contains the points located between the bottoms of the wells.

Theorem 5.1 ([BG1, theorems 2.10–2.12]).

- *There is a constant $h_0 > 0$ such that for all $h \leqslant h_0\sqrt{\varepsilon}/\sigma$, the first-exit time $\tau_{\mathcal{B}(h)}$ of x_t from $\mathcal{B}(h)$ satisfies*

$$\mathbb{P}^{t_1,x_{t_1}}\left\{\tau_{\mathcal{B}(h)} < \sqrt{\varepsilon}\right\} \leqslant C_\varepsilon \mathrm{e}^{-\kappa h^2} \tag{5.7}$$

where

$$C_\varepsilon = \frac{|\alpha(\sqrt{\varepsilon}, t_1)| + \mathcal{O}(\varepsilon)}{\varepsilon^2} \qquad \text{and} \qquad \kappa = \frac{1}{2} - \mathcal{O}(\sqrt{\varepsilon}) - \mathcal{O}\left(\frac{\sigma^2 h^2}{\varepsilon}\right). \tag{5.8}$$

- *Assume that $\sigma|\log\sigma|^{3/2} = \mathcal{O}(\sqrt{\varepsilon})$. Then for any $\varrho \in (0, 2/3)$, the first-exit time $\tau_{\mathcal{D}(\varrho)}$ of x_t from $\mathcal{D}(\varrho)$ satisfies*

$$\mathbb{P}^{\sqrt{\varepsilon},x_{\sqrt{\varepsilon}}}\left\{\tau_{\mathcal{D}(\varrho)} \geqslant t\right\} \leqslant C(t,\varepsilon)\frac{|\log\sigma|}{\sigma}\frac{\mathrm{e}^{-\varrho\alpha(t,\sqrt{\varepsilon})/\varepsilon}}{\sqrt{1-\mathrm{e}^{-2\varrho\alpha(t,\sqrt{\varepsilon})/\varepsilon}}} \tag{5.9}$$

where

$$C(t,\varepsilon) = \mathit{const}\;\mu(t)\left(1 + \frac{\alpha(t,\sqrt{\varepsilon})}{\varepsilon}\right). \tag{5.10}$$

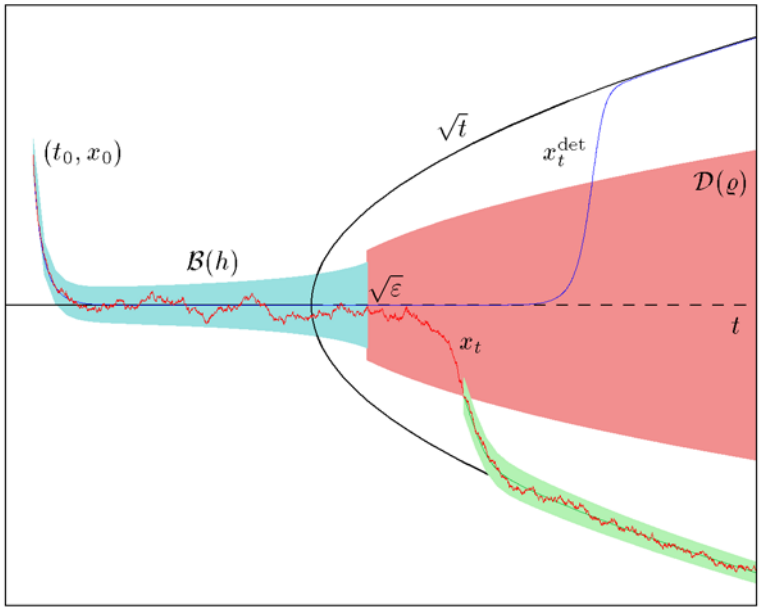

Figure 10. A sample path of the SDE (5.3) with $\mu(t) = t$, for $\varepsilon = 0.01$ and $\sigma = 0.015$. A deterministic solution is shown for comparison. Up to time $\sqrt{\varepsilon}$, the path remains in the set $\mathcal{B}(h)$ centred at x_t^{det}, shown here for $h = 3$. It then leaves the set $\mathcal{D}(\varrho)$ (here $\varrho = 2/3$) after a time of order $\sqrt{\varepsilon|\log\sigma|}$, after which it remains in a neighbourhood of the deterministic solution starting at the same time on the boundary of $\mathcal{D}(\varrho)$.

- *Assume x_t leaves $\mathcal{D}(\varrho)$ (with $1/2 < \varrho < 2/3$) through its upper (lower) boundary. Let $x_t^{\mathrm{det},\tau}$ be the deterministic solution starting at time $\tau = \tau_{\mathcal{D}(\varrho)}$ on the upper (lower) boundary of $\mathcal{D}(\varrho)$. Then $x_t^{\mathrm{det},\tau}$ approaches the equilibrium branch at $\sqrt{\mu(t)}$ (resp., $-\sqrt{\mu(t)}$) as $\varepsilon/\mu(t)^{3/2} + \sqrt{\mu(\tau)}\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}$, where $\eta = 2 - 3\varrho$. Moreover, x_t is likely to stay in a strip centred at $x_t^{\mathrm{det},\tau}$, with width of order $\sigma/\sqrt{\mu(t)}$, at least up to times of order* 1.

The bound (5.7), which is proved in a similar way as theorem 2.2, shows that paths are unlikely to leave the strip $\mathcal{B}(h)$ if $1 \ll h \leqslant h_0\sqrt{\varepsilon}/\sigma$. If σ is smaller than $\sqrt{\varepsilon}$, paths remain concentrated in a neighbourhood of the origin up to time $\sqrt{\varepsilon}$, with a typical spreading growing as $\sigma/\sqrt{|\mu(t)|}$ for $t \leqslant -\sqrt{\varepsilon}$, and remaining of order $\sigma/\varepsilon^{1/4}$ for $|t| \leqslant \sqrt{\varepsilon}$. This is again a dynamical effect: although there is a saddle at the origin for positive times, its curvature is so small that paths do not have time to escape before $t = \sqrt{\varepsilon}$ (see figure 10).

Relation (5.9) yields an upper bound on the typical time needed to leave $\mathcal{D}(\varrho)$, and thus enter a region where the potential is convex. Since $\alpha(t, \sqrt{\varepsilon})$ grows as $\frac{1}{2}t^2$ for small t, the probability not to leave $\mathcal{D}(\varrho)$ before time t becomes small as soon as

$$t \gg \sqrt{\frac{2}{\varrho}\varepsilon|\log\sigma|}. \tag{5.11}$$

The last part of the theorem implies that another time span of the same order is needed for paths to concentrate again, around an adiabatic solution tracking the bottom of the well (at a distance of order $\varepsilon/\mu(t)^{3/2}$). One can thus say that the typical bifurcation delay time of the dynamical pitchfork bifurcation with noise is of order $\sqrt{\varepsilon|\log\sigma|}$.

As a consequence, we can distinguish three parameter regimes (figure 11):

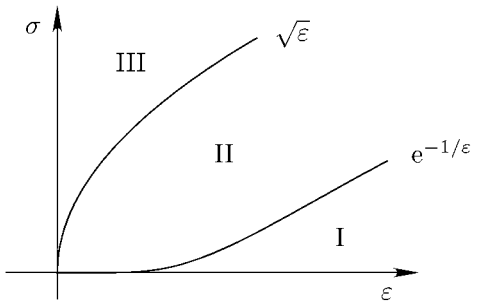

Figure 11. Depending on the value of noise intensity σ and drift velocity ε, the random bifurcation delay has one of three qualitatively different behaviours. If σ is exponentially small in $1/\varepsilon$, the typical delay is macroscopic. For σ larger than $\sqrt{\varepsilon}$, there is no such delay, and paths make large excursions near the bifurcation point. In the intermediate regime, the typical delay is of order $\sqrt{\varepsilon|\log\sigma|}$.

(I) *Exponentially small noise:* $\sigma \leqslant \mathrm{e}^{-K/\varepsilon}$ for some $K > 0$.
At time $\sqrt{\varepsilon}$, the spreading of paths is still exponentially small. In fact, one can extend relation (5.7) to all times for which $\alpha(t, \sqrt{\varepsilon}) < K$. If K is larger than $\alpha(\Pi(t_0), 0)$, where $\Pi(t_0)$ is the deterministic bifurcation delay, most paths will track the deterministic solution and follow it into the right-hand potential well.

(II) *Moderate noise:* $\mathrm{e}^{-1/\varepsilon^p} \leqslant \sigma \ll \sqrt{\varepsilon}$ for some $p < 1$.
The bifurcation delay lies between $\sqrt{\varepsilon}$ and a constant times $\sqrt{\varepsilon|\log\sigma|} \leqslant \varepsilon^{(1-p)/2}$ with high probability. One can thus speak of a 'microscopic' bifurcation delay.

(III) *Large noise:* $\sigma \geqslant \sqrt{\varepsilon}$.
The spreading of paths grows as $\sigma/\sqrt{|\mu(t)|}$ at least up to time $-\sigma$. As t approaches the bifurcation time 0, the bottom of the potential well becomes so flat that the paths are no longer localized near the origin and may switch wells several times before eventually settling for a well. So for large noise intensities, the concept of bifurcation delay should be replaced by *two* variables, namely the first-exit time from a suitably chosen neighbourhood of the saddle and the time when the potential wells become attractive enough to counteract the diffusion.

One can also estimate the probability to reach the right-hand potential well rather than the left-hand potential well. Loosely speaking, if x_s reaches 0 before time t, it has equal probability to choose either potential well. It follows that

$$\mathbb{P}^{t_0,x_0}\{x_t \geqslant 0\} = \tfrac{1}{2} + \tfrac{1}{2}\mathbb{P}^{t_0,x_0}\{x_s > 0 \ \forall s \in [t_0, t]\}\,. \tag{5.12}$$

One can show that for $t = 0$ the second term on the right-hand side is of order

$$\frac{x_0\varepsilon^{1/4}}{\sigma}\mathrm{e}^{-|\alpha(0,t_0)|/\varepsilon} \tag{5.13}$$

and for larger t this term will be even smaller. Thus in case II, paths will choose one potential well or the other with a probability exponentially close to $1/2$.

In the particular case where $\mu(t) > \text{const}\ t$ for all $t > 0$, the height of the potential barrier grows without bound. This implies that once x_t has chosen a potential well, its probability *ever* to cross the saddle again is of order $\mathrm{e}^{-\text{const}/\sigma^2}$.

The existence of three parameter regimes has some interesting consequences on the experimental determination of a bifurcation diagram. Assume we want to determine the stable equilibrium branches by sweeping the parameter with speed ε. Regime II is the most

favourable: in regime I, part of the stable branches cannot be seen due to the bifurcation delay, while in regime III, noise will blur the bifurcation diagram. For a given noise intensity σ, the sweeping rate ε should thus satisfy

$$\sigma^2 \ll \varepsilon \ll \left(1/|\log\sigma|\right)^{1/p} \tag{5.14}$$

in order to produce a good image of the stable equilibria. As long as $\sigma^2 \ll 1/|\log\sigma|^{1/p}$, artificially increasing the noise level allows one to work with higher sweeping rates, but of course the image will be more and more blurred.

On the other hand, the relation between noise and delay can be used to measure the intensity of noise present in the system. If a bifurcation delay is observed for a sweeping rate ε_0, repeating the experiment with slower and slower sweeping rates should ultimately suppress the delay. If this happens for $\varepsilon = \varepsilon_1$, then the noise intensity is of order $\mathrm{e}^{-\mathrm{const}/\varepsilon_1}$.

6. Generalizations

6.1. Multidimensional systems

We now return to n-dimensional equations such as (1.11). As in section 2, we start by examining the linear equation

$$\mathrm{d}y_t = \frac{1}{\varepsilon}A(t)y_t\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}G(t)\,\mathrm{d}W_t \qquad y_0 = 0 \tag{6.1}$$

obtained, for instance, by linearizing the equation around a given deterministic solution (in that case, the matrices A and G may depend on ε). If the drift term derives from a potential, then A is necessarily symmetric, but we will not impose such a restriction here. We will assume, however, that we are in the stable case, that is, all eigenvalues of $A(t)$ have real parts smaller than some constant $-a_0 < 0$.

The solution of (6.1) can be written as

$$y_t = \frac{\sigma}{\sqrt{\varepsilon}}\int_0^t U(t,s)G(s)\,\mathrm{d}W_s \tag{6.2}$$

where $U(t,s) = U(t)U(s)^{-1}$ is the propagator of the deterministic equation $\varepsilon\dot{y} = A(t)y$, and we denote by $U(t)$ its principal solution, i.e. $\varepsilon\dot{U}(t) = A(t)U(t)$ and $U(0) = \mathbb{1}$. The random variable y_t has a Gaussian distribution, with zero expectation and covariance matrix

$$\mathrm{Cov}\{y_t\} = \frac{\sigma^2}{\varepsilon}\int_0^t U(t,s)G(s)G(s)^T U(t,s)^T\,\mathrm{d}s. \tag{6.3}$$

In order to determine the asymptotic behaviour of the covariance matrix, we first note that it is a particular solution of the ODE

$$\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}X(t) = A(t)X(t) + X(t)A(t)^T + \sigma^2 G(t)G(t)^T. \tag{6.4}$$

The general solution of (6.4) can be written as

$$X(t) = \overline{X}(t) + U(t)[X(0) - \overline{X}(0)]U(t)^T \tag{6.5}$$

where $\overline{X}(t)$ is a particular solution admitting an asymptotic expansion (cf [Wa])

$$\overline{X}(t) = \overline{X}_0(t) + \varepsilon\overline{X}_1(t) + \varepsilon^2\overline{X}_2(t) + \cdots. \tag{6.6}$$

The terms of this expansion satisfy the Liapunov equations

$$A(t)\overline{X}_0(t) + \overline{X}_0(t)A(t)^T = -\sigma^2 G(t)G(t)^T \tag{6.7}$$

$$A(t)\overline{X}_n(t) + \overline{X}_n(t)A(t)^T = \frac{\mathrm{d}}{\mathrm{d}t}\overline{X}_{n-1}(t) \qquad \forall n \geqslant 1. \tag{6.8}$$

The solution of (6.7) admits the integral representation (see for instance [Bel])

$$\overline{X}_0(t) = \sigma^2 \int_0^\infty \mathrm{e}^{A(t)s} G(t) G(t)^T \mathrm{e}^{A(t)^T s} \, \mathrm{d}s \tag{6.9}$$

and similarly for the solutions of (6.8). Note that $\overline{X}_0$ corresponds to the asymptotic covariance matrix of (6.2) if A and G are constant in time.

At any fixed time t, the distribution of y_t is concentrated in an ellipsoid of the form $\langle y, \mathrm{Cov}\{y_t\}^{-1} y\rangle \leqslant$ const, provided the (Gaussian) distribution of y_t is nondegenerate. Assume for the moment that $G(t)G(t)^T$ is uniformly positive definite. Then $\mathrm{Cov}\{y_t\}^{-1}$ exists and so does $\overline{X}(s)^{-1}$. In addition, we find that $\sigma^2 \overline{X}(s)^{-1}$ is bounded in norm. The following result, which generalizes proposition 2.1 from the one-dimensional case, shows that the paths $\{y_t\}_{t \geqslant 0}$ are concentrated in sets $\{(t, y)\colon \langle y, \overline{X}(t)^{-1} y\rangle \leqslant \text{ const}\}$. The proof will be given in [BG4].

Theorem 6.1. *Assume that $\sigma^2 \overline{X}(s)^{-1}$ is bounded in norm. Then for all $t > 0$, $h > 0$ and any $\kappa \in (0, 1/2)$,*

$$\mathbb{P}^{0,0} \left\{ \sup_{0 \leqslant s \leqslant t} \langle y_s, \overline{X}(s)^{-1} y_s \rangle \geqslant h^2 \right\} \leqslant C_n(t, \varepsilon) \mathrm{e}^{-\kappa h^2 (1 - \mathcal{O}(\varepsilon))} \tag{6.10}$$

where

$$C_n(t, \varepsilon) = \left(\frac{t}{\varepsilon^2} + 1 \right) \left(\frac{1}{1 - 2\kappa} \right)^{n/2}. \tag{6.11}$$

The exponential growth of the prefactor as a function of the dimension implies that the estimate (6.10) is only useful for $h > \sqrt{n}$. This dependence, however, is to be expected as the tails of n-dimensional (standard) Gaussians only show their typical decay outside a ball of radius proportional to $\sqrt{n}$.

Under the condition that $\sigma^{-2} \overline{X}(s)$ and $\sigma^2 \overline{X}(s)^{-1}$ are bounded in norm, one can show that a bound similar to (6.10) holds if y_s obeys a nonlinear perturbation of equation (6.1), for all h smaller than a constant times σ^{-1}.

6.2. Coloured noise

White noise has no time-correlations, and is thus appropriate to model the random influence of a 'fast' environment on a 'slow' system, if the relaxation time of the fast system is negligible. If the relaxation time is short but not negligible, one has to use coloured noise instead of white noise.

The simplest (still Markovian) model of coloured noise is the Ornstein–Uhlenbeck process

$$Z_t = \sigma \int_0^t \mathrm{e}^{-\gamma(t-s)} \, \mathrm{d}W_s \tag{6.12}$$

whose autocorrelation function $\mathbb{E}\{Z_s Z_t\} = (\sigma^2/2\gamma)\mathrm{e}^{-\gamma(t-s)}[1 - \mathrm{e}^{-2\gamma s}]$, $s < t$, decays exponentially in $t - s$.

Let us examine the influence of such coloured noise on a slowly time-dependent, one-dimensional, linear system, described by the SDE

$$\mathrm{d}x_t = a(\varepsilon t) x_t \, \mathrm{d}t + g(\varepsilon t) \, \mathrm{d}Z_t. \tag{6.13}$$

We assume that $a(\varepsilon t)$ is negative, $g(\varepsilon t)$ is positive, and that both functions are uniformly bounded away from zero. Using the differential representation $\mathrm{d}Z_t = -\gamma Z_t \, \mathrm{d}t + \sigma \, \mathrm{d}W_t$ of the

Ornstein–Uhlenbeck process (6.12), and scaling time by a factor ε, we obtain that the random variable $y_t = (x_t, Z_t)$ obeys a two-dimensional linear SDE of the form (6.1), with

$$A(t) = \begin{pmatrix} a(t) & -\gamma g(t) \\ 0 & -\gamma \end{pmatrix} \qquad \text{and} \qquad G(t) = \begin{pmatrix} g(t) \\ 1 \end{pmatrix}. \tag{6.14}$$

For $\gamma \to 0$, we recover the white-noise case, while $\sigma \to 0$ corresponds to the deterministic limit.

Theorem 6.1 shows that paths are concentrated in a tube $\langle y_s, \overline{X}(s)^{-1} y_s \rangle \leqslant$ const, with covariance matrix $\overline{X}(s)$ admitting the asymptotic series (6.6). The leading term $\overline{X}_0(s)$ is found from (6.9) to be

$$\overline{X}_0(s) = \sigma^2 \begin{pmatrix} \frac{g(s)^2}{2(\gamma+|a(s)|)} & \frac{g(s)}{2(\gamma+|a(s)|)} \\ \frac{g(s)}{2(\gamma+|a(s)|)} & \frac{1}{2\gamma} \end{pmatrix}. \tag{6.15}$$

In particular, the paths $\{x_s\}_{s\geqslant 0}$ are concentrated in a strip of width $\sigma g(s)/\sqrt{2(\gamma + |a(s)|)}$. The effect of coloured noise of the form (6.12) is thus to narrow the distribution of the paths, in the same way as if the curvature were increased by an amount γ.

This result allows us to make some predictions on the effect of noise colour on SR. We have seen in section 3.2 that in the standard case of the asymmetric double-well potential with additive periodic forcing and white noise, paths switch between potential wells when σ exceeds the threshold $\sigma_c = (a_0 \vee \varepsilon)^{3/4}$. This value was obtained by comparing the typical spreading of paths to the minimal distance between the deterministic solution and the saddle. Repeating this argument in the case of coloured noise given by (6.12), we thus expect the threshold noise intensity to be determined by

$$\sigma_c^2 = (a_0 \vee \varepsilon)(\gamma \vee (a_0 \vee \varepsilon)^{1/2}). \tag{6.16}$$

This shows in particular that shorter correlation times of the noise require larger noise intensities to enable transitions between the wells.

A similar stabilizing effect of noise colour can be expected when sweeping the bifurcation parameter through a pitchfork bifurcation as discussed in section 5. The bifurcation delay will lie between the macroscopic delay of the deterministic case and the microscopic delay of the white-noise case. The shorter the correlation time of the noise, the larger will the delay be.

Acknowledgment

We are grateful to Anton Bovier for helpful comments on a preliminary version of the manuscript. NB thanks the WIAS for kind hospitality.

References

[Ar1] Arnold L 1998 *Random Dynamical Systems* (Berlin: Springer)

[Ar2] Arnold L 2001 Hasselmann's program revisited: the analysis of stochasticity in deterministic climate models *Stochastic Climate Models (Progress in Probability* vol 49) ed P Imkeller and J-S von Storch (Boston: Birkhäuser) pp 141–58

[Az] Azencott R 1985 Petites perturbations aléatoires des systèmes dynamiques: développements asymptotiques *Bull. Sci. Math.* **109** 253–308

[Bel] Bellman R 1960 *Introduction to Matrix Analysis* (New York: McGraw-Hill)

[BPSV] Benzi R, Parisi G, Sutera A and Vulpiani A 1983 A theory of stochastic resonance in climatic change *SIAM J. Appl. Math.* **43** 565–78

[BSV] Benzi R, Sutera A and Vulpiani A 1981 The mechanism of stochastic resonance *J. Phys. A: Math. Gen.* **14** L453–7

[BG1] Berglund N and Gentz B 2001 Pathwise description of dynamic pitchfork bifurcations with additive noise *Probab. Theory Related Fields* DOI 10.1007/s004400100174

[BG2] Berglund N and Gentz B 2000 A sample-paths approach to noise-induced synchronization: stochastic resonance in a double-well potential *Ann. Appl. Prob.* at press
(Berglund N and Gentz B 2000 *Preprint* arXiv:math.PR/0012267)

[BG3] Berglund N and Gentz B 2002 The effect of additive noise on dynamical hysteresis *Nonlinearity* **15** at press
(Berglund N and Gentz B 2001 *Preprint* arXiv:math.DS/0107199)

[BG4] Berglund N and Gentz B Geometric singular perturbation theory for stochastic differential equations, in preparation

[BK] Berglund N and Kunz H 1999 Memory effects and scaling laws in slowly driven systems *J. Phys. A: Math. Gen.* **32** 15–39

[Ce] Cessi P 1994 A simple box model of stochastically forced thermohaline flow *J. Phys. Oceanogr.* **24** 1911–20

[CF1] Crauel H and Flandoli F 1994 Attractors for random dynamical systems *Probab. Theory Related Fields* **100** 365–93

[CF2] Crauel H and Flandoli F 1998 Additive noise destroys a pitchfork bifurcation *J. Dyn. Differ. Eqns* **10** 259–74

[Day1] Day M V 1983 On the exponential exit law in the small parameter exit problem *Stochastics* **8** 297–323

[Day2] Day M V 1995 On the exit law from saddle points *Stoch. Process. Appl.* **60** 287–311

[DT] Dhar D and Thomas P B 1992 Hysteresis and self-organized criticality in the $O(N)$ model in the limit $N \to \infty$ *J. Phys. A: Math. Gen.* **25** 4967–84

[D&] Dykman M I *et al* 1992 Stochastic resonance for periodically modulated noise intensity *Phys. Rev.* A **46** R1713–16

[ET] Eckmann J-P and Thomas L E 1982 Remarks on stochastic resonance *J. Phys. A: Math. Gen.* **15** L261–6

[FJ] Fleming W H and James M R 1992 Asymptotic series and exit time probabilities *Ann. Probab.* **20** 1369–84

[Fox] Fox R F 1989 Stochastic resonance in a double well *Phys. Rev.* A **39** 4148–53

[Fr1] Freidlin M I 2000 Quasi-deterministic approximation, metastability and stochastic resonance *Physica* D **137** 333–52

[Fr2] Freidlin M I 2001 On stable oscillations and equilibriums induced by small noise *J. Stat. Phys.* **103** 283–300

[FW] Freidlin M I and Wentzell A D 1984 *Random Perturbations of Dynamical Systems* (New York: Springer)

[GHM] Gammaitoni L, Hänggi P and Marchesoni F 1998 Stochastic resonance *Rev. Mod. Phys.* **70** 223–87

[Gr] Gradšteĭn I S 1953 Applications of AM Lyapunov's theory of stability to the theory of differential equations with small coefficients in the derivatives *Mat. Sbornik N.S.* **32** 263–86

[HN] Hasegawa H and Nakagomi T 1979 Semiclassical laser theory in the stochastic and thermodynamic frameworks *J. Stat. Phys.* **21** 191–214

[Ha] Hasselmann K 1976 Stochastic climate models. Part I. Theory *Tellus* **28** 473–85

[HL] Hepp K and Lieb E H 1975 The laser: a reversible quantum dynamical system with irreversible classical macroscopic motion *Dynamical Systems, Theory and Applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974) (Lecture Notes in Phys. vol 38)* (Berlin: Springer) pp 178–207

[JL] Jansons K M and Lythe G D 1998 Stochastic calculus: application to dynamic bifurcations and threshold crossings *J. Stat. Phys.* **90** 227–51

[JGRM] Jung P, Gray G, Roy R and Mandel P 1990 Scaling law for dynamical hysteresis *Phys. Rev. Lett.* **65** 1873–6

[JH1] Jung P and Hänggi P 1989 Stochastic nonlinear dynamics modulated by external periodic forces *Europhys. Lett.* **8** 505–10

[JH2] Jung P and Hänggi P 1991 Amplification of small signals via stochastic resonance *Phys. Rev.* A **44** 8032–42

[Ki] Kifer Y 1981 The exit problem for small random perturbations of dynamical systems with a hyperbolic fixed point *Isr. J. Math.* **40** 74–96

[Lo] Longtin A 1993 Stochastic resonance in neuron models *J. Stat. Phys.* **70** 309–27

[MS1] Maier R S and Stein D L 1997 Limiting exit location distributions in the stochastic exit problem *SIAM J. Appl. Math.* **57** 752–90

[MS2] Maier R S and Stein D L 2001 Noise-activated escape from a sloshing potential well *Phys. Rev. Lett.* **86** 3942–5

[Mar] Martin Ph A 1977 On the stochastic dynamics of Ising models *J. Stat. Phys.* **16** 149–68

[McNW] McNamara B and Wiesenfeld K 1989 Theory of stochastic resonance *Phys. Rev.* A **39** 4854–69

[Mo] Monahan A H 2001 Stabilisation of climate regimes by noise in a simple model of the thermohaline circulation *Preprint*

[MW] Moss F and Wiesenfeld K 1995 The benefits of background noise *Sci. Am.* **273** 50–3

[N&] Neiman A *et al* 1999 Stochastic synchronization of electroreceptors in the paddlefish http://neurodyn.umsl.edu/~neiman/synchronization/

[NSAS] Neiman A, Silchenko A, Anishchenko V and Schimansky-Geier L 1998 Stochastic resonance: noise-enhanced phase coherence *Phys. Rev.* E **58** 7118–25

[Ne1] Neishtadt A I 1987 Persistence of stability loss for dynamical bifurcations I *Differ. Eqns* **23** 1385–91
[Ne2] Neishtadt A I 1988 Persistence of stability loss for dynamical bifurcations II *Differ. Eqns* **24** 171–6
[Ra] Rahmstorf S 1995 Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle *Nature* **378** 145–9
[RKP] Rao M, Krishnamurthy H K and Pandit R 1990 Magnetic hysteresis in two model spin systems *Phys. Rev.* B **42** 856–84
[Ri] Risken H 1966 Correlation function of the amplitude and of the intensity fluctuation for a laser model near threshold *Z. Phys.* **191** 302–12
[Schm] Schmalfuß B 1989 Invariant attracting sets of nonlinear stochastic differential equations *Math. Res.* **54** 217–28
[SHD] Schütte Ch, Huisinga W and Deuflhard P 1999 Transfer operator approach to conformational dynamics in biomolecular systems *ZIB Preprint* SC-99-36
[SNA] Shulgin B, Neiman A and Anishchenko V 1995 Mean switching frequency locking in stochastic bistable systems driven by a periodic force *Phys. Rev. Lett.* **75** 4157–60
[SRN] Sides S W, Rikvold P A and Novotny M A 1998 Stochastic hysteresis and resonance in a kinetic Ising system *Phys. Rev.* E **57** 6512–33
[SD] Somoza A M and Desai R C 1993 Kinetics of systems with continuous symmetry under the effect of an external field *Phys. Rev. Lett.* **70** 3279–82
[SMC] Stocks N G, Manella R and McClintock P V E 1989 Influence of random fluctuations on delayed bifurcations: the case of additive white noise *Phys. Rev.* A **40** 5361–9
[St] Stommel H 1961 Thermohaline convection with two stable regimes of flow *Tellus* **13** 224–30
[SHA] Swift J B, Hohenberg P C and Ahlers G 1991 Stochastic Landau equation with time-dependent drift *Phys. Rev.* A **43** 6572–80
[Ta] Talkner P 1999 Stochastic resonance in the semiadiabatic limit *New J. Phys.* **1** 4.1–4.25
[Ti] Tihonov A N 1952 Systems of differential equations containing small parameters in the derivatives *Mat. Sbornik N.S.* **31** 575–86
[TO] Tomé T and de Oliveira M J 1990 Dynamic phase transition in the kinetic Ising model under a time-dependent oscillating field *Phys. Rev.* A **41** 4251–4
[TM] Torrent M C and San Miguel M 1988 Stochastic-dynamics characterization of delayed laser threshold instability with swept control parameter *Phys. Rev.* A **38** 245–51
[Tu] Tuckwell H C 1989 *Stochastic Processes in the Neurosciences* (Philadelphia: SIAM)
[Wa] Wasow W 1976 *Asymptotic Expansions for Ordinary Differential Equations* (Reprint of the 1965 edn) (Huntington, NY: Krieger)
[WM] Wiesenfeld K and Moss F 1995 Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs *Nature* **373** 33–6
[ZZ] Zhong F and Zhang J 1995 Renormalization group theory of hysteresis *Phys. Rev. Lett.* **75** 2027–30