Metastability in simple climate models: Pathwise analysis of slowly driven Langevin equations

Nils Berglund and Barbara Gentz

Abstract

We consider simple stochastic climate models, described by slowly time-dependent Langevin equations. We show that when the noise intensity is not too large, these systems can spend substantial amounts of time in metastable equilibrium, instead of adiabatically following the stationary distribution of the frozen system. This behaviour can be characterized by describing the location of typical paths, and bounding the probability of atypical paths. We illustrate this approach by giving a quantitative description of phenomena associated with bistability, for three famous examples of simple climate models: Stochastic resonance in an energy balance model describing the Ice Ages; hysteresis in a box model for the Atlantic thermohaline circulation; and bifurcation delay in the case of the Lorenz model for Rayleigh–Bénard convection.

Date. November 4, 2001. *Final.* August 14, 2002.

2000 *Mathematical Subject Classification.* 37H20 (primary), 60H10, 34E15, 82C31 (secondary).

Keywords and phrases. Stochastic resonance, dynamical hysteresis, bifurcation delay, double-well potential, first-exit time, scaling laws, Lorenz model, thermohaline circulation, white noise, coloured noise.

1 Introduction

One of the main difficulties of realistic climate models is that they involve a huge number of interacting degrees of freedom, on a wide range of time and length scales. In order to be able to control these models analytically, or at least numerically, it is necessary to simplify them by eliminating the less relevant degrees of freedom (e. g. high-frequency or short-wavelength modes). A possible way to do this is to average the equations of motion over all fast degrees of freedom, a rather drastic approximation. As proposed by Hasselmann [20] (see also [2]), a more realistic approximation is obtained by modeling the effect of fast degrees of freedom by noise.

In a number of cases, it is appropriate to distinguish between three rather than two time scales: Fast degrees of freedom (e. g. the "weather"), which are modeled by a stochastic process; intermediate "dominant modes" (e. g. the average temperature of the atmosphere) whose dynamics we want to predict; and slow degrees of freedom (e. g. the mean insolation depending on the eccentricity of the Earth's orbit), which evolve on very long time scales of several centuries or millennia, and can be viewed as an external forcing. Such a system can often be modeled by a slowly time-dependent Langevin equation

$$\mathrm{d}x_t = f(x_t, \varepsilon t)\,\mathrm{d}t + \sigma G(\varepsilon t)\,\mathrm{d}W_t. \tag{1.1}$$

On the mathematical level, we view (1.1) as a *stochastic differential equation* (SDE), in which W_t is a standard (vector-valued) Wiener process, describing white noise, and G

is a matrix. The adiabatic parameter ε and the noise intensity σ are small parameters. Note that historically, the name "Langevin equation" is associated with the particular case where x_t is the velocity of a Brownian particle, but nowadays it is often used in a wider sense, so that we will call SDEs such as (1.1) Langevin equations.

Our aim in this paper is to describe the effect of the noise term on the dynamics of (1.1), assuming the dynamics without noise is known. For this purpose, we will concentrate on bistable systems, which frequently occur in simple climate models: For instance, in models for the major Ice Ages, where the two possible stable equilibria correspond to warm and cold climate [4], or in models of the Atlantic thermohaline circulation [34, 30]. Noise may enable transitions between the two stable states, which would be impossible in the deterministic case, and our main concern will be to quantify this effect.

The method used to study the stochastic differential equation (1.1) will depend on the time scale we are interested in. Let us first illustrate this on a static one-dimensional example, namely the overdamped motion in a symmetric double-well potential:

$$\mathrm{d}x_t = -\frac{\partial}{\partial x}V(x_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad V(x) = \frac{1}{4}bx^4 - \frac{1}{2}ax^2, \tag{1.2}$$

where a and b are positive constants. The potential has two wells at $\pm\sqrt{a/b}$, separated by a barrier of height $H = a^2/(4b)$. A first possibility to analyse this equation is to compute the probability density $p(x,t)$ of x_t. It obeys the Fokker–Planck equation

$$\frac{\partial}{\partial t}p(x,t) = \frac{\partial}{\partial x}\left[\frac{\partial V}{\partial x}(x)p(x,t)\right] + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}p(x,t), \tag{1.3}$$

which admits in particular the stationary solution

$$p_0(x) = \frac{1}{N}\,\mathrm{e}^{-2V(x)/\sigma^2}, \tag{1.4}$$

where N is the normalization. The potential $V(x)$ being even, at equilibrium there is equal probability to find x_t in either potential well. Moreover, for weak noise it is unlikely to observe x_t anywhere else than in a neighbourhood of order σ of one of the wells.

Assume now that the initial distribution x_0 is concentrated at the bottom $\sqrt{a/b}$ of the right-hand potential well. Then it may take quite a long time for the density to approach its asymptotic value (1.4). A possible way to investigate this problem relies on spectral theory. Denote the right-hand side of (1.3) as $\mathcal{L}p(x,t)$, where $\mathcal{L}$ is a linear differential operator. The stationary density (1.4) is an eigenfunction of $\mathcal{L}$ with eigenvalue 0. We may assume that $\mathcal{L}$ has eigenvalues $\dots < \lambda_k < \dots < \lambda_2 < \lambda_1 < 0$, cf. [21, Section 6.7]. Decomposing $p(x,t)$ on a basis of eigenfunctions of $\mathcal{L}$, we see that p approaches the stationary solution in a characteristic time of order $1/|\lambda_1|$.

There exists, however, a much more precise description of the process x_t than by its probability density, which gives only an instantaneous picture. Given an initial condition, the state of the system at time t is determined by the realization of the noise (i. e., the *path* of the Brownian motion) up to that time t. We indicate this by writing $W_t(\omega)$ for the Brownian motion and $x_t(\omega)$ for the solution of the SDE (1.1), where ω customarily stands for the particular realization of the noise. There are essentially two ways to look at the stochastic process x_t. On the one hand, at a fixed time t, the state $x_t(\omega)$ of the process is a random variable characterized by its density $p(x,t)$. On the other hand, for a given ω, that is, for a particular realization of the noise, we can consider $x_t(\omega)$ as a function of time t which we call a *sample path*. It corresponds to what an experimentalist would call a time series. In the sequel, we will adopt this sample-path point of view. It is known that almost every sample path of the SDE (1.1) is continuous (although not differentiable).

Instead of computing the time needed for $p(x,t)$ to relax to $p_0(x)$, we can consider the time when the sample path first crosses the saddle. (One could as well consider the first time the bottom of the left-hand well is reached). We denote this time by τ, and since the sample path depends on the realization of the Brownian motion, so does τ. Thus τ is a random variable, and we write $\tau(\omega)$ whenever we want to stress this fact. Formally, we can define

$$\tau(\omega) = \inf\bigl\{t > 0 \colon x_t(\omega) < 0\bigr\}, \tag{1.5}$$

which simply means that $\tau(\omega)$ is the infimum of all times at which $x_t(\omega)$ has reached the left-hand side of the saddle.

The small-noise behaviour of the average value of the random time τ at which the process crosses the barrier has already been studied by Arrhenius in 1889, who obtained the correct behaviour up to subexponential corrections, and by Eyring and Kramers in the 1930s who derived subexponential corrections, depending on the curvature of the potential at its extrema. Not taking the subexponential corrections into account, the average value of the time of the first barrier crossing behaves like

$$T_{\text{Kramers}} = \mathrm{e}^{2H/\sigma^2}\,. \tag{1.6}$$

We will follow the convention of calling this time "Kramers' time".

A mathematical theory allowing 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 [18]. In specific situations, more precise results are available, for instance subexponential corrections to the asymptotic expression (1.6), see [3, 15]. Even the limiting behaviour of the distribution of the first-exit time from a neighbourhood of a unique stable equilibrium point has been obtained [14]. Asymptotically, this first-exit time is exponentially distributed, with expectation behaving like Kramers' time T_{Kramers}. The first-exit time from a neighbourhood of a saddle has been considered by Kifer in the seminal paper [25].

If the noise intensity σ is small (compared to the square root of the barrier height), then the time needed to overcome the potential barrier is extremely long, and the time required to relax to the stationary distribution $p_0(x)$ is even longer. In fact, on time scales shorter than Kramers' time, solutions of (1.2) starting in one potential well will hardly feel the second potential well. As we will see in Section 2, x_t is well approximated by an Ornstein–Uhlenbeck process, describing the overdamped motion of a particle in a potential of constant curvature $c = 2a$. The Ornstein–Uhlenbeck process relaxes to a *stationary* Gaussian process with variance $\sigma^2/(2c)$ in a characteristic time

$$T_{\text{relax}} = \frac{1}{c}. \tag{1.7}$$

Thus for $0 \leqslant t \ll T_{\text{relax}}$, the behaviour of x_t is transient; for $T_{\text{relax}} \ll t \ll T_{\text{Kramers}}$, x_t is close to a stationary Ornstein–Uhlenbeck process with variance $\sigma^2/(2c)$; and only for $t \gg T_{\text{Kramers}}$ will the distribution of x_t approach the bimodal stationary solution (1.4). This phenomenon, where a process seems stationary for a long time before ultimately relaxing to a new (possibly stationary) state, is known as *metastability.* It is all the more remarkable in an asymmetric double-well potential: then a process starting at the bottom of the shallow well will first relax to a metastable distribution concentrated in the shallow well, which is radically different from the stationary distribution having most of its mass concentrated in the deeper well.

A different approach to the SDE (1.1), based on the concept of random attractors (see [12, 31, 1]), gives complementary information on the long-time regime. In particular, in [13] it is proved that for arbitrarily weak noise, paths of (1.2) with different initial conditions but same realization of noise almost surely converge to a random point. The time needed for this convergence, however, diverges rapidly in the limit $\sigma \to 0$, because paths starting in different potential wells are unlikely to overcome the potential barrier and therefore cannot start approaching each other before Kramers' time.

We now turn to situations in which the potential varies slowly in time. For simplicity, we will consider the family of Ginzburg–Landau potentials

$$V(x;\lambda,\mu) = \frac{1}{4}x^4 - \frac{1}{2}\mu x^2 - \lambda x, \tag{1.8}$$

and let either λ or μ vary in time, with low speed ε. For instance, λ or μ may depend periodically on time, with low frequency ε. The potential V has two wells if $27\lambda^2 < 4\mu^3$ and one well if $27\lambda^2 > 4\mu^3$, and when λ or μ are varied, the number of wells may change. Crossing one of the curves $27\lambda^2 = 4\mu^3$, $\mu > 0$, corresponds to a saddle–node bifurcation, and crossing the point $\lambda = \mu = 0$ corresponds to a pitchfork bifurcation.

The slow time-dependence introduces a new time scale $T_{\text{forcing}} = 1/\varepsilon$. Since curvature and barrier height are no longer constant, we replace the definitions (1.6) and (1.7) by

$$T_{\text{Kramers}}^{(\max)} = \mathrm{e}^{2H_{\max}/\sigma^2} \qquad \text{and} \qquad T_{\text{relax}}^{(\min)} = \frac{1}{c_{\max}}, \tag{1.9}$$

where $H_{\max}$ denotes the maximal barrier height during the time interval under consideration, and $c_{\max}$ denotes the maximal curvature at the bottom of a potential well. Here we are interested in the regime

$$T_{\text{relax}}^{(\min)} \ll T_{\text{forcing}} \ll T_{\text{Kramers}}^{(\max)}, \tag{1.10}$$

which means that the process has time to reach a metastable "equilibrium", but does not feel the possibly bimodal stationary distribution associated with the potential frozen at a fixed time t. Mathematically, we thus assume that $\varepsilon \ll c_{\max}$ and $\sigma^2 \ll 2H_{\max}/|\log\varepsilon|$. We allow, however, the *minimal* curvature and barrier height to become small, or even to vanish.

For time-dependent potentials, the Fokker–Planck equation (1.3) is even harder to solve (and in fact, it does not admit a stationary solution). Moreover, random attractors are not straightforward to define in this time-dependent setting. We believe that the dynamics on time scales shorter than $T_{\text{Kramers}}^{(\max)}$ is discussed best via an understanding of "typical" paths. The idea is to show that the vast majority of paths remain concentrated in small space–time sets, whose shape and size depend on the potential and the noise intensity. These sets are typically located in a neighbourhood of the potential wells, but under some conditions paths may also switch potential wells. There are thus two problems to solve: first characterize the sets in which typical paths live, and then estimate the probability of atypical paths. It turns out that these properties have universal characteristics, depending only on qualitative properties of the potential, especially its bifurcation points.

We start, in Section 2, by discussing the simplest situation, which occurs when the initial condition of the process lies in the basin of attraction of a stable equilibrium branch. For sufficiently small noise intensity, the majority of paths remain concentrated for a long time in a neighbourhood of the equilibrium branch. We determine the shape of this neighbourhood and outline how coloured noise can decrease the spreading of paths.

Section 3 is devoted to the phenomenon of stochastic resonance. We first recall the energy-budget model introduced in [4] to give a possible explanation for the close-to-periodic appearance of the major Ice Ages. This model is equivalent to the overdamped motion of a particle in a modulated double-well potential, where the driving amplitude is too small to allow for transitions between wells in the absence of noise. Turning to the description of typical paths, we find a threshold value for the noise intensity below which the paths remain in one well, while above threshold, they switch back and forth between wells twice per period. The switching events occur close to the instants of minimal barrier height. Several important quantities have a power-law dependence on the small parameters, in particular the critical noise intensity, the width of transition windows, and the exponent controlling the exponential decay of the probability of atypical paths.

In Section 4, we start by discussing a variant [11] of Stommel's box model [34] of the Atlantic thermohaline circulation. Assuming slow changes in the typical weather, this model also reduces to the motion in a modulated double-well potential, where the modulation depends on the freshwater flux. If the amplitude of the modulation exceeds a threshold, the potential barrier vanishes twice per period, so that the deterministic motion displays hysteresis. Additive noise influences the shape of hysteresis cycles, and may even create macroscopic cycles for subthreshold modulation amplitude. We characterize the distribution of the random freshwater flux causing the system to switch from one stable state to the other one.

In Section 5, we consider the Lorenz model for Rayleigh–Bénard convection with slowly increasing heating. In the deterministic case, convection rolls appear only some time after the steady state looses stability in a pitchfork bifurcation. This bifurcation delay is significantly decreased by additive noise, as soon as its intensity is not exponentially small.

Finally, Section 6 contains some concluding remarks.

2 Near stable equilibria

Let us start by investigating Equation (1.1) in the one-dimensional case, i. e., when x_s and W_s are scalar. We write it as

$$\mathrm{d}x_s = f(x_s, \varepsilon s)\,\mathrm{d}s + \sigma g(\varepsilon s)\,\mathrm{d}W_s. \tag{2.1}$$

Here we used s to denote time, because we are interested in the dynamics on the time scale $T_{\text{forcing}} = 1/\varepsilon$, and reserve the letter t to denote the so-called *slow time* $t = \varepsilon s$. With a slight abuse of notation we will write x_t for $x_{\varepsilon s}$. Note that in the trivial case $f \equiv 1$ and $g \equiv 0$, we have $x_t = s = t/\varepsilon$, while in the case $f \equiv 0$ and $g \equiv 1$, the variance of x_t grows like $\sigma^2 s = \sigma^2 t/\varepsilon$. After rescaling time, in the general case the equation becomes

$$\mathrm{d}x_t = \frac{1}{\varepsilon} f(x_t, t)\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}} g(t)\,\mathrm{d}W_t. \tag{2.2}$$

In this section, we will consider the dynamics near a stable equilibrium branch of f, i. e., a curve $x^\star(t)$ such that

$$f(x^\star(t), t) = 0 \qquad \text{and} \qquad a^\star(t) = \frac{\partial f}{\partial x}(x^\star(t), t) \leqslant -a_0 \tag{2.3}$$

for all t, where a_0 is a positive constant. In the one-dimensional case, f always derives from a potential V. The first relation in (2.3) indicates that $x^\star(t)$ is a critical point of the potential V, while the second relation states that this critical point is actually the bottom of a well, with the curvature $-a^\star(t)$.

In the deterministic case $\sigma = 0$, solutions of (2.2) track the equilibrium branch $x^\star(t)$ adiabatically. In fact, Tihonov's Theorem [36, 19] asserts that for $\sigma = 0$, (2.2) admits a particular solution $\bar{x}_t^{\det}$ with an asymptotic expansion of the form

$$\bar{x}_t^{\det} = x^\star(t) + \varepsilon \frac{\dot{x}^\star(t)}{a^\star(t)} + \mathcal{O}(\varepsilon^2). \tag{2.4}$$

Since $a^\star(t)$ is negative, $\bar{x}_t^{\det}$ lies a little bit to the left of $x^\star(t)$ if $x^\star(t)$ moves to the right, and vice versa. The *adiabatic solution* $\bar{x}_t^{\det}$ attracts nearby solutions exponentially fast in t/ε.

Consider now the SDE (2.2) with positive noise intensity. For the sake of brevity, we assume that g is positive and bounded away from zero. In a nutshell, our main result can be formulated as follows: Up to Kramer's time, paths starting near $\bar{x}_0^{\det}$ are concentrated in a neighbourhood of order $\sigma g(t)/\sqrt{|a^\star(t)|}$ of the deterministic solution with the same initial condition, as shown in Figure 1. Larger noise intensities and smaller curvatures thus lead to a larger spreading of paths. This result holds as long as the spreading is smaller than the distance between $x^\star(t)$ and the nearest unstable equilibrium (i. e., the nearest saddle of the potential).

To make this claim mathematically precise, we need a few definitions. For simplicity, we discuss first the particular case when the process starts on the adiabatic solution, i.e., $x_0 = \bar{x}_0^{\det}$. We denote the curvature at the adiabatic solution $\bar{x}_t^{\det}$ and two related integrals by

$$a(t) = \frac{\partial f}{\partial x}(\bar{x}_t^{\det}, t), \qquad \alpha(t,s) = \int_s^t a(u)\,\mathrm{d}u \qquad \text{and} \qquad \alpha(t) = \alpha(t,0). \tag{2.5}$$

Note that by (2.4), $a(t) = a^\star(t) + \mathcal{O}(\varepsilon)$ is negative for sufficiently small ε. The main idea is that $x_t - \bar{x}_t^{\det}$ is well approximated by a generalized Ornstein–Uhlenbeck process, with time-dependent damping $a(t)/\varepsilon$ and diffusion coefficient $\sigma g(t)/\sqrt{\varepsilon}$. This process is obtained by linearizing the SDE (2.2) around $\bar{x}_t^{\det}$, namely by considering the SDE

$$\mathrm{d}(x_t^0 - \bar{x}_t^{\det}) = \frac{1}{\varepsilon} a(t)(x_t^0 - \bar{x}_t^{\det})\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}} g(t)\,\mathrm{d}W_t. \tag{2.6}$$

The process $x_t^0 - \bar{x}_t^{\det}$ is Gaussian with variance

$$v(t) = \frac{\sigma^2}{\varepsilon} \int_0^t \mathrm{e}^{2\alpha(t,s)/\varepsilon} g(s)^2\,\mathrm{d}s. \tag{2.7}$$

The function $v(t)$ solves the ordinary differential equation (ODE) $\varepsilon\dot{v} = 2a(t)v + \sigma^2 g(t)^2$. In analogy with (2.4), this equation also admits a particular solution $\bar{v}(t)$ satisfying

$$\bar{v}(t) = \frac{\sigma^2}{2|a(t)|}\bigl[g(t)^2 + \mathcal{O}(\varepsilon)\bigr]. \tag{2.8}$$

The actual variance is related to $\bar{v}(t)$ by $v(t) = \bar{v}(t) - \bar{v}(0)\,\mathrm{e}^{2\alpha(t)/\varepsilon}$. Since $\alpha(t) \leqslant -a_0 t$ for $t \geqslant 0$, $v(t)$ approaches $\bar{v}(t)$ exponentially fast.

The density of the linear approximation x_t^0 is concentrated, at any time $t \geqslant 0$, in an interval of width proportional to $\sqrt{v(t)}$, centred in the adiabatic solution $\bar{x}_t^{\det}$. We claim

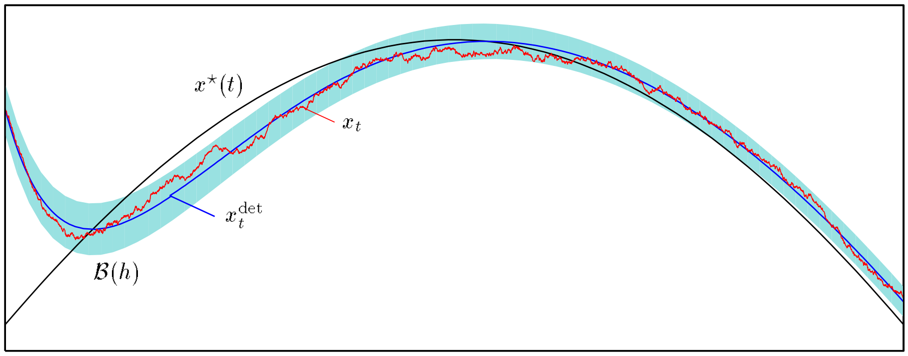

FIGURE 1. A sample path of the SDE (2.2), for $f(x,t) = a^\star(t)(x - x^\star(t))$ deriving from a quadratic single-well potential, and $g(t) \equiv 1$. The potential well is located at $x^\star(t) = \sin(2\pi t)$, and has curvature $-a^\star(t) = 4 - 2\sin(4\pi t)$. Parameter values are $\varepsilon = 0.04$ and $\sigma = 0.025$. After a short transient motion, the deterministic solution x_t^{det} tracks $x^\star(t)$ at a distance of order ε. The path x_t is likely to stay in the shaded set $\mathcal{B}(h)$ (shown here for $h = 3$), which is centred at x_t^{det} and has time-dependent width of order $h\sigma/\sqrt{|a^\star(t)|}$.

that similar properties hold for the *whole sample paths* of the nonlinear equation (2.3), where the interval of width proportional to $v(t)$ is to be replaced by a strip consisting of the union of such intervals. This strip can be defined by

$$\mathcal{B}(h) = \bigl\{(x,t)\colon |x - \bar{x}_t^{\text{det}}| < h\sqrt{\bar{v}(t)}\bigr\}, \tag{2.9}$$

where $h > 0$ is a real parameter we may still choose. For $h = 1$, the strip $\mathcal{B}(1)$ is centred in the adiabatic solution $\bar{x}_t^{\text{det}}$ tracking the bottom of the potential well, and has time-dependent width $\sigma g(t)/\sqrt{2|a^\star(t)|}[1 + \mathcal{O}(\varepsilon)]$. The parameter h simply allows to shrink or stretch the width of that strip without changing its centre or shape. To lowest order in ε and σ, $\mathcal{B}(h)$ coincides with the points in the potential well for which $V(x,t) - V(\bar{x}_t^{\text{det}}, t)$ is smaller than $(\frac{1}{2}h\sigma g(t))^2$.

The main result is that for $h \gg 1$, paths $\{x_s\}_{s\geqslant 0}$ are unlikely to leave the strip $\mathcal{B}(h)$ before Kramers' time. The first time at which a path leaves the strip $\mathcal{B}(h)$ depends, of course, on the realization of the Brownian motion. It is thus a random variable, denoted by

$$\tau_{\mathcal{B}(h)} = \inf\bigl\{t > 0\colon (x_t, t) \notin \mathcal{B}(h)\bigr\}, \tag{2.10}$$

and called the *first-exit time* of x_t from the strip. This first-exit time is unlikely to be smaller than T_{Kramers}. Indeed, one can prove the following estimate (see [7, Theorem 2.4] and [5, Theorem 2.2]). There is a constant $h_0 > 0$ such that the probability of observing a sample path x_t leaving the strip $\mathcal{B}(h)$ before time t, satisfies

$$\mathbb{P}\bigl\{\tau_{\mathcal{B}(h)} < t\bigr\} \leqslant C(t,\varepsilon)\,\mathrm{e}^{-\kappa h^2} \tag{2.11}$$

for all $t > 0$ and all $h \leqslant h_0/\sigma$. Here

$$C(t,\varepsilon) = \frac{|\alpha(t)|}{\varepsilon^2} + 2 \qquad \text{and} \qquad \kappa = \frac{1}{2} - \mathcal{O}(\varepsilon) - \mathcal{O}(\sigma h). \tag{2.12}$$

The exponential term $\mathrm{e}^{-\kappa h^2}$ in (2.11) is independent of time, and becomes small as soon as $h \gg 1$. The constant h_0 depends on f and is the smaller the smaller a_0 is: The flatter the well, the more restrictive the condition $h \leqslant h_0/\sigma$ becomes. The prefactor $C(t,\varepsilon)$, which grows as time increases (and provides only an upper estimate for the correct prefactor), only leads to subexponential corrections on the time scale T_{forcing}. Some time dependence of the prefactor is to be expected, as it reflects the fact that occasionally a path will make an unusually large excursion, and the longer we wait the more excursions we will observe. The prefactor also depends on ε. A factor $1/\varepsilon$ is due to the fact that we are working on the time scale T_{forcing}, while the actual factor of $1/\varepsilon^2$ allows us to obtain the best possible exponent. Choosing κ slightly smaller allows to replace ε^2 by the more natural ε in the definition of $C(t,\varepsilon)$. Thus we find that paths are unlikely to leave the strip $\mathcal{B}(h)$ before time t, provided $h_0^2/\sigma^2 \geqslant h^2 \gg \log C(t,\varepsilon)$.

Let us give a numerical example for the function $f(x,t)$ considered in Figure 1. It is well-known that the 95%-confidence interval for a centred Gaussian random variable is the symmetric interval around zero, of width 2×1.96 its standard deviation. Hence, at any fixed time $t > 0$, with a probability of 0.95 we will find x_t in the strip $\mathcal{B}(1.96)$. For the *whole sample path* to be contained in $\mathcal{B}(h)$ up to time t with the same probability, we need to take a larger h. Estimate (2.11) gives the condition

$$h^2 = \frac{1}{\kappa}\log\biggl(\frac{C(t,\varepsilon)}{0.05}\biggr). \tag{2.13}$$

For $t = 1/2$, $\varepsilon = 0.04$ and $-a^\star(u) = 4 - 2\sin(4\pi u)$ as in Figure 1, we find $C(t,\varepsilon) = 1252[1+\mathcal{O}(\varepsilon)]$, leading to $h = 4.5$. In fact, we expect the bound (2.12) to be too pessimistic, and that (2.11) should already hold for $C(t,\varepsilon)$ close to $|\alpha(t)|/\varepsilon = 50$; in this case, $h = 3.7$ would correspond to a 95%-confidence interval.

The same results hold if x_t does not start on the adiabatic solution $\bar{x}_t^{\text{det}}$, but in some deterministic x_0 sufficiently close to it. Then $\bar{x}_t^{\text{det}}$ has to be replaced in (2.5) and (2.9) by the solution x_t^{det} of the deterministic equation with initial condition x_0. We still have that $a(t)$ is negative (and bounded away from zero), but note that (2.8) may not hold for very small t, when x_t^{det} has not yet approached $\bar{x}_t^{\text{det}}$. These results also apply if the noise coefficient g in (2.2) depends on x: in that case, g in Equations (2.6)–(2.8) has to be evaluated at the particular deterministic solution under consideration, cf. [9].

If the potential V grows at least quadratically for large $|x|$, one can deduce from (2.11) that the moments of $|x_t - x_t^{\text{det}}|$ are bounded by those of a centred Gaussian distribution with variance of order $\bar{v}(t)$, for times small compared to Kramers' time [5, Corollary 2.4], even if V has other potential wells than the one at $x^\star(t)$. Assume for instance that V has two potential wells, with the shallower one at $x^\star(t)$. Then the system is in metastable "equilibrium" for an exponentially long time span during which the existence of the deeper well is not felt.

Similar statements are valid in the multidimensional case (in which f does not necessarily derive from a potential). Let $x^\star(t)$ be an equilibrium branch of f, and denote by $A^\star(t)$ the Jacobian matrix of f at $x^\star(t)$. We assume that the eigenvalues of $A^\star(t)$ have real parts smaller than some negative constant $-a_0$ for all times, so that $x^\star(t)$ is asymptotically stable. In the deterministic case $\sigma = 0$, Tihonov's theorem shows the existence of an adiabatic solution

$$\bar{x}_t^{\text{det}} = x^\star(t) + \varepsilon A^\star(t)^{-1}\dot{x}^\star(t) + \mathcal{O}(\varepsilon^2), \tag{2.14}$$

which attracts nearby orbits exponentially fast. Let $A(t)$ be the Jacobian matrix of f at $\bar{x}_t^{\text{det}}$. It satisfies $A(t) = A^\star(t) + \mathcal{O}(\varepsilon)$. The solution of the SDE (1.1) linearized at $\bar{x}_t^{\text{det}}$ has a Gaussian distribution, with covariance matrix

$$X(t) = \frac{\sigma^2}{\varepsilon}\int_0^t U(t,s)G(s)G(s)^T U(t,s)^T \,\mathrm{d}s, \tag{2.15}$$

where $U(t,s)$ is the fundamental solution of $\varepsilon\dot{y} = A(t)y$ with initial condition $U(s,s) = \mathbb{1}$. Here and in the sequel we use the notation $G(s)^T$ to denote the transposed matrix of $G(s)$, and similarly for other matrices. To keep the presentation simple, we will assume that the smallest eigenvalue of $G(s)G(s)^T$ is bounded away from zero and the largest one is bounded above. Note that $X(t)$ obeys the ODE $\varepsilon\dot{X} = AX + XA^T + \sigma^2 GG^T$, and approaches exponentially fast a matrix $\overline{X}(t)$ which satisfies

$$\overline{X}(t) = \overline{X}_0(t) + \mathcal{O}(\varepsilon). \tag{2.16}$$

The matrix $\overline{X}_0(t)$ satisfies the Liapunov equation

$$(-A^\star)\overline{X}_0 + \overline{X}_0(-A^\star)^T = \sigma^2 GG^T, \tag{2.17}$$

which admits a unique solution because all eigenvalues of $A^\star$ have negative real parts. Relation (2.17) can be interpreted as describing the energy balance between fluctuations (on the right-hand side), and dissipation by damping (on the left-hand side).

The solution of the SDE obtained by linearizing (1.1) at $\bar{x}_t^{\text{det}}$ is a Gaussian process, and at any fixed time t, its probability density is constant on sets of the form

$$(x - x_t^{\text{det}})^T \overline{X}(t)^{-1}(x - x_t^{\text{det}}) = \mathit{const.} \tag{2.18}$$

Also note that for all $h > 0$, the sets

$$\bigl\{x\colon (x - x_t^{\text{det}})^T \overline{X}(t)^{-1}(x - x_t^{\text{det}}) < h^2\bigr\} \tag{2.19}$$

are ellipsoïds, each consisting of a union of level sets of the probability density. The rôle of the "strip" (2.9) in the one-dimensional case is now played by the "tube"

$$\mathcal{B}(h) = \bigl\{(x,t)\colon (x - x_t^{\text{det}})^T \overline{X}(t)^{-1}(x - x_t^{\text{det}}) < h^2\bigr\}, \tag{2.20}$$

which is centred at the deterministic solution $\bar{x}_t^{\text{det}}$. At any time t, the cross-section of the tube is the corresponding ellipsoïd (2.19). The shape of the tube takes into account that the typical spreading of the process is not necessarily symmetric with respect to rotations. Assume for the moment that the drift coefficient derives from a potential. Then it is quite obvious that the typical spreading of paths will be the larger in a particular direction the flatter the well is in that direction.

One can again show that the sample paths are unlikely to leave the tube $\mathcal{B}(h)$ before Kramers' time. Indeed, (2.11) generalizes to the following statement (see [5, Theorem 6.1] for a discussion; the proof is given in [9]): There is a constant $h_0 > 0$ such that for all $h \leqslant h_0/\sigma$ and all $\kappa \in (0, 1/2)$, the probability of observing a sample path leaving the tube before time t satisfies

$$\mathbb{P}\bigl\{\tau_{\mathcal{B}(h)} < t\bigr\} \leqslant C(t,\varepsilon)\,\mathrm{e}^{-\kappa h^2(1-\mathcal{O}(\varepsilon)-\mathcal{O}(\sigma h))}, \tag{2.21}$$

where

$$C(t,\varepsilon) = \Bigl(\frac{t}{\varepsilon^2} + 1\Bigr)\Bigl(\frac{1}{1-2\kappa}\Bigr)^{n/2}, \tag{2.22}$$

n being the dimension of x. Paths are thus concentrated, up to a given time t, in a tube $\mathcal{B}(h)$, which have an ellipsoïdal cross-section defined by $\overline{X}(t)$. Again the parameter h must satisfy $h_0^2/\sigma^2 \geqslant h^2 \gg \log C(t,\varepsilon)$.

This result can be used, in particular, to understand the effect of coloured noise. Assume for instance that the one-dimensional system

$$\mathrm{d}x_s = f(x_s, \varepsilon s)\,\mathrm{d}s + g(\varepsilon s)\,\mathrm{d}Z_s \tag{2.23}$$

is not driven by white noise, but by an Ornstein–Uhlenbeck process Z_s obeying the SDE

$$\mathrm{d}Z_s = -\gamma Z_s\,\mathrm{d}s + \sigma\,\mathrm{d}W_s. \tag{2.24}$$

The equations (2.23) and (2.24) can be rewritten, on the time scale $1/\varepsilon$, as a two-dimensional system of the form (1.1) for (x_t, Z_t). We assume that f has a stable equilibrium branch $x^\star(t)$ with linearization $a^\star(t) \leqslant -a_0 < 0$. To leading order in ε, the asymptotic covariance matrix (2.16) is given by

$$\overline{X}_0(t) = \sigma^2 \begin{pmatrix} \dfrac{g(t)^2}{2(\gamma+|a^\star(t)|)} & \dfrac{g(t)}{2(\gamma+|a^\star(t)|)} \\ \dfrac{g(t)}{2(\gamma+|a^\star(t)|)} & \dfrac{1}{2\gamma} \end{pmatrix}. \tag{2.25}$$

The conditions on GG^T mentioned above can be relaxed (cf. [5, Theorem 6.1]), so that (2.21) is applicable. We find in particular that the path $\{x_t\}_{t\geqslant 0}$ is concentrated in a strip of width proportional to $\sigma g(t)/\sqrt{\gamma+|a^\star(t)|}$, centred around x_t^{det}. Hence larger "noise colour" γ yields a smaller spreading of the paths, in the same way as if the curvature of the potential were increased by γ.

3 Stochastic resonance

In the previous section, we have seen that on a certain time scale, paths typically remain in metastable equilibrium. With overwhelming probability, they are concentrated in a strip of order $\sigma g(t)/\sqrt{|a^\star(t)|}$ near the bottom of a potential well with curvature $|a^\star(t)|$. This roughly holds as long as the strip does not extend to the nearest saddle of the potential. New phenomena may occur when this hypothesis is violated, either because the noise coefficient $\sigma g(t)$ becomes too large, or because the curvature or the distance to the saddle become too small. Then paths may overcome the potential barrier and reach another potential well. This mechanism has various interesting consequences, one of them being the effect called stochastic resonance.

Stochastic resonance (SR) was initially introduced as a possible explanation for the close-to-periodic appearance of the major Ice Ages [4]. While this explanation remains controversial, SR has been detected in several other physical and biological systems, see for instance [29, 39] for a review.

The original model in [4] is based on an energy balance of the Earth in integrated form. The evolution of the mean surface temperature T is described by the differential equation

$$c\frac{\mathrm{d}T}{\mathrm{d}s} = Q(1+A\cos\omega s)(1-\alpha(T)) - E(T). \tag{3.1}$$

Here the term $R_{\mathrm{in}} = Q(1+A\cos\omega s)$ is the incoming solar radiation, where Q denotes the solar constant, and the periodic term models the effect of the Earth's varying orbital eccentricity. The amplitude A of this modulation is very small, of the order 5×10^{-4}, while its period $2\pi/\omega$ equals 92 000 years. The outgoing radiation $R_{\text{out}} = \alpha(T)R_{\text{in}} + E(T)$ depends on the albedo $\alpha(T)$ of the Earth and its emissivity. c denotes the heat capacity.

To account for the existence of two stable climate states (warm climate and Ice Age), the right-hand side of (3.1) should have two stable and one unstable equilibrium points. The authors of [4] postulate that

$$\gamma(T) = \frac{Q}{E(T)}(1-\alpha(T)) - 1 = \beta\Big(1-\frac{T}{T_1}\Big)\Big(1-\frac{T}{T_2}\Big)\Big(1-\frac{T}{T_3}\Big), \tag{3.2}$$

where $T_1 = 278.6$ K and $T_3 = 288.6$ K are the representative temperatures of the two stable states, and $T_2 = 283.3$ K represents the unstable state. Since $E(T) \sim T^4$ varies little on this range, the problem can be further simplified by neglecting the T-dependence of $E(T) \simeq \langle E\rangle$. Equation (3.1) becomes

$$\frac{\mathrm{d}T}{\mathrm{d}s} = \frac{\langle E\rangle}{c}\Big[\beta\Big(1-\frac{T}{T_1}\Big)\Big(1-\frac{T}{T_2}\Big)\Big(1-\frac{T}{T_3}\Big)(1+A\cos\omega s) + A\cos\omega s\Big]. \tag{3.3}$$

The parameter β is related to the relaxation time $\tau \simeq 8$ years of the system via

$$\frac{1}{\tau} = \frac{\langle E\rangle}{c}\beta\frac{1}{T_3}\Big(1-\frac{T_3}{T_1}\Big)\Big(1-\frac{T_3}{T_2}\Big). \tag{3.4}$$

Let us now transform this system to a dimensionless form. We do this in two steps: First we scale time by a factor $\omega/2\pi$, so that in the new variables, the system has period 1. Then we introduce the variable $x = (T-T_2)/\Delta T$, where $\Delta T = (T_3 - T_1)/2 = 5$ K. The resulting system is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{\varepsilon}\big[-x(x-x_1)(x-x_3)(1+A\cos 2\pi t) + K\cos 2\pi t\big], \tag{3.5}$$

where $x_1 = (T_1 - T_2)/\Delta T \simeq -0.94$ and $x_3 = (T_3 - T_2)/\Delta T \simeq 1.06$. The adiabatic parameter ε is given by

$$\varepsilon = \frac{\omega\tau}{2\pi}\,\frac{2(T_3-T_2)}{\Delta T} \simeq 1.8\times 10^{-4}. \tag{3.6}$$

This confirms that we are in the adiabatic regime. Using the value $\langle E\rangle/c = 8.77 \times 10^{-3}/4000\ \mathrm{Ks}^{-1}$ from [4], we find an effective driving amplitude

$$K = \frac{A}{\beta}\frac{T_1T_2T_3}{(\Delta T)^3} \simeq 0.12. \tag{3.7}$$

The term in brackets in (3.5) derives from a double-well potential, which is almost of the Ginzburg–Landau type (1.8). If we set, for simplicity, $x_1 = -1$ and $x_3 = 1$, and neglect the term $A\cos 2\pi t$, then we obtain indeed a force deriving from the potential (1.8), with $\mu = 1$ and $\lambda = K\cos 2\pi t$. This potential has two wells if and only if $|\lambda| < \lambda_{\mathrm{c}} = 2/3\sqrt{3} \simeq 0.38$, and thus the amplitude K of the forcing is too small to enable transitions between the potential wells. Note, however, that although A is very small, the effective driving amplitude K is not negligible compared to the threshold value λ_{c}.

The main new idea in [4] is that if one models the effect of the "weather" by an additive noise term, then transitions between potential wells not only become possible but, due to the periodic forcing, these transitions will be more likely at some times than at others, so that the evolution of T can be close to periodic. We will illustrate this on the model SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon}\big[x_t - x_t^3 + K\cos 2\pi t\big]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t. \tag{3.8}$$

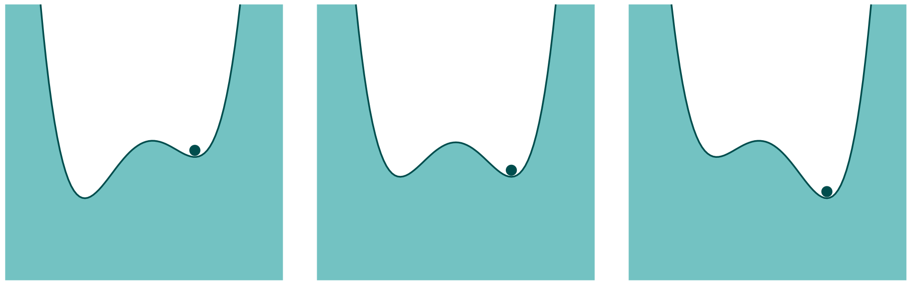

FIGURE 2. The potential $V(x,t) = \frac{1}{4}x^4 - \frac{1}{2}x^2 - K\cos(2\pi t)x$, from which derives the drift term in (3.8). For $\cos(2\pi t) = 0$, the potential is symmetric (middle), for integer times, the left-hand well approaches the saddle (right), while for half-integer times, the right-hand well approaches the saddle (left). If the amplitude K is smaller than the threshold λ_c, there is always a potential barrier, which an overdamped particle cannot overcome in the deterministic case. Sufficiently strong noise, however, helps the particle to switch from the shallower to the deeper well. This effect is the stronger the lower the barrier is, so that switching typically occurs close to the instants of minimal barrier height.

However, the results in [8] apply to a more general class of periodically forced double-well potentials, including (3.5).

Various characterizations of the effect of noise on the dynamics of (3.8), and various measures of periodicity have been proposed. A widespread approach uses the signal-to-noise ratio, a property of the power spectrum of x_t, which shows peaks near multiples of the driving frequency [16, 27, 24]. For small driving amplitudes K, the signal-to-noise ratio behaves like $e^{-H/\sigma^2}/\sigma^4$, where H is the height of the potential barrier in the absence of periodic driving (i. e., for $K = 0$). The signal's "periodicity" is thus optimal for $\sigma^2 = H/2$. A different approach is used in [17], where the L^p-distance between sample paths and a deterministic, periodic limiting function is shown to converge to zero in probability as $\sigma \to 0$. This result requires ε to be of order e^{-2H/σ^2}, which implies exponentially long forcing periods.

We examine here a different regime, in which the forcing amplitude K is not necessarily a small parameter, but may approach λ_c. In this way, transitions become possible for values of ε which are not exponentially small (as in (3.6)). The potential barrier is lowest at integer and half-integer times. At integer times, the left-hand well approaches the saddle, while at half-integer times, the right-hand well approaches the saddle, cf. Figure 2.

The minimal values $H_{\min}$, $c_{\min}$ and $\delta_{\min}$ of the barrier height, the curvature at the bottom of the wells, and the distance between the bottom of one of the wells and the saddle can be expressed as functions of a parameter $a_0 = \lambda_c - K$. For small a_0, they behave like $H_{\min}(a_0) \asymp a_0^{3/2}$, $c_{\min}(a_0) \asymp a_0^{1/2}$ and $\delta_{\min}(a_0) \asymp a_0^{1/2}$ (meaning $c_- a_0^{3/2} \leqslant H_{\min}(a_0) \leqslant c_+ a_0^{3/2}$ for some positive constants $c_\pm$ independent of a_0, and so on).

Intuitively, our results from Section 2 indicate that the maximal spreading of paths is of order $\sigma/c_{\min}(a_0)^{1/2}$, provided this value is smaller than $\delta_{\min}(a_0)$, i. e., provided $\sigma \ll a_0^{3/4}$. Assume for instance that we start at time 1/4 (when the potential is symmetric) near the right-hand potential well. We call *transition probability* the probability P_{trans} of having reached the left-hand potential well by time 3/4, after passing through the configuration with the shallowest right-hand well. Extrapolating (2.11) with h of the order $\delta_{\min} c_{\min}^{1/2}/\sigma \asymp H_{\min}^{1/2}/\sigma$, we find

$$P_{\text{trans}} \leqslant \frac{const}{\varepsilon^2} e^{-const\, a_0^{3/2}/\sigma^2} = \frac{const}{\varepsilon^2} e^{-const\, H_{\min}/\sigma^2} \qquad \text{for } \sigma \leqslant a_0^{3/4}. \tag{3.9}$$

Note the similarity with Kramers' time for the potential frozen at the moment of minimal barrier height.

A bound of this form can indeed be proved, but (3.9) turns out to be a little bit too pessimistic for very small a_0. This is a rather subtle dynamical effect, related to the behaviour of the deterministic system. Recall that the strip $\mathcal{B}(h)$ in (2.9) is defined via the linearization at the adiabatic solution $\bar{x}_t^{\text{det}}$, not at the bottom $x^\star(t)$ of the potential well. This distinction is irrelevant as long as the minimal curvature remains of order one, but *not* when it is a small parameter. In that case, the asymptotic expansion (2.4) does not necessarily converge. Using methods from singular perturbation theory [10], one can show that $\bar{x}_t^{\text{det}}$ never approaches the saddle closer than a distance of order $\sqrt{\varepsilon}$, so that the curvature at $\bar{x}_t^{\text{det}}$ never becomes smaller than a quantity of order $\sqrt{\varepsilon}$, even if $a_0 < \varepsilon$. As a consequence, for $a_0 < \varepsilon$, the system behaves as if there were an effective potential barrier of height $\varepsilon^{3/2}$.

In fact, one can prove the following bound (see [8, Theorem 2.6] and [5, Theorem 3.1]): There exist constants $C, \kappa > 0$ such that the transition probability satisfies

$$P_{\text{trans}} \leqslant \frac{C}{\varepsilon}\, \mathrm{e}^{-\kappa\sigma_{\text{c}}^2/\sigma^2} \qquad \text{for } \sigma \leqslant \sigma_{\text{c}} = (a_0 \vee \varepsilon)^{3/4}, \tag{3.10}$$

where $a \vee b$ denotes the maximum of two real numbers a and b. In addition, paths remain concentrated in a strip $\mathcal{B}(h)$ of the form (2.9). Examining the behaviour of the integral (2.7), one can show that the width of $\mathcal{B}(h)$ behaves, near $t = 1/2$, like $h\sigma/(|t - 1/2|^{1/2} \vee \sigma_{\text{c}}^{1/3})$. The various exponents entering these relations do not depend on the details of the potential, but only on some qualitative properties of the "avoided bifurcation", and can be deduced geometrically from a Newton polygon [10].

What happens when σ exceeds the threshold value σ_{c}? Away from half-integer times, the right-hand well may still be sufficiently deep to confine the paths. However, there are time intervals near half-integer t during which it becomes possible to overcome the barrier. Near $t = 1/2$, the curvature $c(t)$ at $\bar{x}_t^{\text{det}}$ and the distance between $\bar{x}_t^{\text{det}}$ and the saddle both behave like $|t - 1/2| \vee \sigma_{\text{c}}^{2/3}$. Transitions thus become possible for $|t - 1/2| \leqslant \sigma^{2/3}$.

During this time interval, the process x_t makes a certain number of attempts to overcome the barrier. If the saddle is reached, x_t has roughly equal probability to fall back into the right-hand well, in which case it will make further attempts to cross the barrier, or to fall into the deeper left-hand well, where it is likely to stay during the next half-period. One can show that the typical time for each excursion is of order $\varepsilon/c(t)$. Although the different attempts are not independent, the probability *not* to reach the left-hand well during the transition window $|t - 1/2| \leqslant \sigma^{2/3}$ behaves roughly like $(1/2)^N$, where N is the maximal number of possible excursions.

These arguments can be used to show (see [8, Theorem 2.7] and [5, Theorem 3.1]) that there exist constants $C, \kappa > 0$ such that the transition probability satisfies

$$P_{\text{trans}} \geqslant 1 - C\, \mathrm{e}^{-\kappa\sigma^{4/3}/(\varepsilon|\log\sigma|)} \qquad \text{for } \sigma \geqslant \sigma_{\text{c}}. \tag{3.11}$$

The factor $\sigma^{4/3}$ is proportional to the integral of the curvature $c(t)$ over the transition window, and the factor $|\log\sigma|$ takes into account the time needed to travel from the saddle to the left-hand well. Amplification by SR is thus optimal for noise intensities just above the threshold σ_{c}, because stronger noise intensities will gradually blur the signal.

In the large-noise regime $\sigma \geqslant \sigma_{\text{c}}$, the vast majority of paths stay in a strip switching back and forth between potential wells each time the barrier height becomes minimal, as

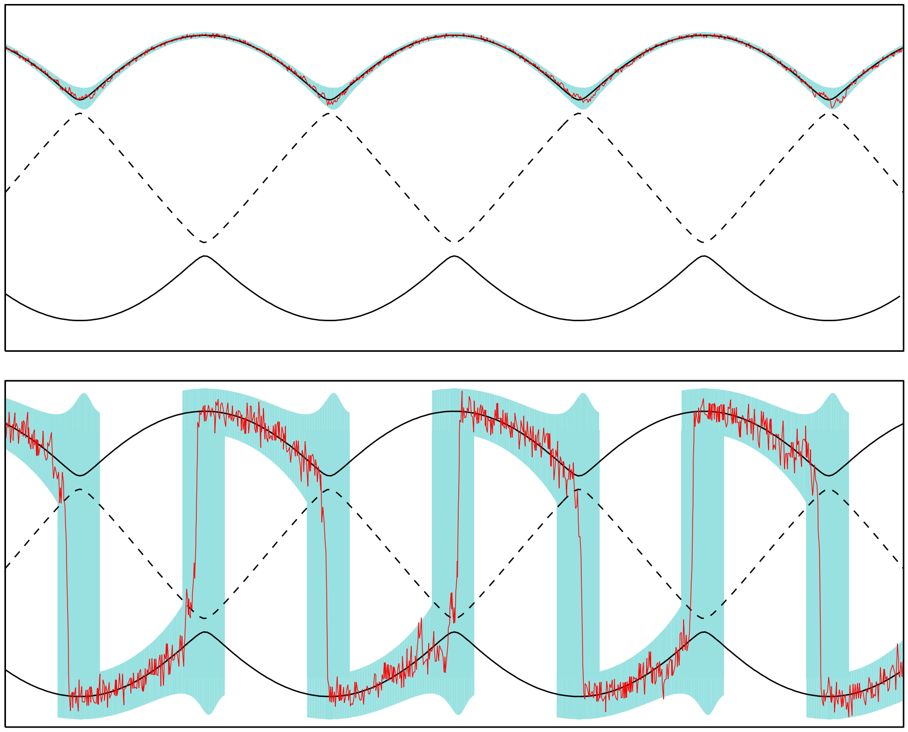

FIGURE 3. Sample paths of the SDE (3.8) for $\varepsilon = a_0 = 0.005$, and $\sigma = 0.02$ (upper picture) and $\sigma = 0.14$ (lower picture). Full curves represent the location of potential wells, the broken curve represents the saddle. For weak noise, the path x_t is likely to stay in the shaded set $\mathcal{B}(h)$, centred at the deterministic solution tracking the right-hand well. The maximal width of $\mathcal{B}(h)$ is of order $h\sigma/(a_0 \vee \varepsilon)^{1/4}$ and is reached at half-integer times. For strong noise, typical paths stay in the shaded set which switches back and forth between the wells at integer and half-integer times. The width of the vertical strips is of order $\sigma^{2/3}$. The "bumps" are due to the fact that one of the wells becomes very flat during the transition window so that paths might also make excursions away from the saddle.

shown in Figure 3. If a path fails to switch wells during a transition window, then with high probability it is confined to the "wrong well" for one period.

Paths spend approximately half the time (for $1/4 < t < 1/2$, $3/4 < t < 1$, and so on) in metastable equilibrium in the shallower potential well. This differs from the quasistatic picture, when the driving period is larger than the maximal Kramers time, and paths spend most of the time in the deeper potential well with occasional excursions to the shallower one.

While the details of the transition process depend on the potential, the exponents in (3.10) and (3.11) depend only on qualitative properties of the avoided bifurcation. Other exponents arise, for instance, if V is a symmetric potential with modulated barrier height of the form (1.8) with $\lambda = 0$ and $\mu(t) = a_0 + 1 - \cos 2\pi t$, cf. [5, Theorem 3.2]. Here an additional feature can be observed: For sufficiently strong noise, the process is likely to reach the saddle during a certain transition window, but due to symmetry, it has about equal probability to be in either of the wells when transitions become unlikely again. Observing the process for several periods, we see that near the instants of minimal barrier height, the process chooses randomly between potential wells, with probability exponentially close to $1/2$ for choosing either.

One can also consider the effect of coloured noise on SR. If the system is driven by an Ornstein–Uhlenbeck process with damping γ, the typical spreading of paths will be smaller, making transitions more difficult. One can show that transitions only become likely above a threshold noise intensity σ_c, given by

$$\sigma_c^2 = (a_0 \vee \varepsilon)\big(\gamma \vee (a_0 \vee \varepsilon)^{1/2}\big), \tag{3.12}$$

where again $a \vee b$ denotes the maximum of the real numbers a, b. If $\gamma < (a_0 \vee \varepsilon)^{1/2}$, we recover the white-noise result, but for larger γ, the threshold grows linearly with γ, namely like $(a_0 \vee \varepsilon)\gamma$.

It is, of course, not easy to decide whether the observed periodicity in the appearance of Ice Ages can be explained by a simple, one-dimensional SDE of the form (3.8). Our results show, however, that in order to match the observations, the noise intensity should lie in a relatively narrow interval. Too weak noise will not allow regular transitions between stable states, while too strong noise increases the width of the transition windows so much that although switching does occur, no periodicity can be observed.

4 Hysteresis

The model (3.1) of the glacial cycle is not the only important bistable system in climate physics. Another well-known example is a model of the Atlantic thermohaline circulation. At present time, the Gulf Stream transports enormous amounts of heat from the Tropics as far north as the Barents Sea, causing the current mild climate in Western Europe. It is believed, however, that this has not always been the case in the past, and that during long time spans, the thermohaline circulation was locked in a stable state with far less heat transported to the North (see for instance [30]).

A simple model for oceanic circulation showing bistability is Stommel's box model [34], where the ocean is represented by two boxes, a low-latitude box with temperature T_1 and salinity S_1, and a high-latitude box with temperature T_2 and salinity S_2. Here we will follow the presentation in [11], where the intrinsic dynamics of salinity and of temperature are not modeled in the same way. The differences $\Delta T = T_1 - T_2$ and $\Delta S = S_1 - S_2$ are assumed to evolve in time s according to the equations

$$\frac{\mathrm{d}}{\mathrm{d}s}\Delta T = -\frac{1}{\tau_r}(\Delta T - \theta) - Q(\Delta\rho)\Delta T \tag{4.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}s}\Delta S = \frac{F(s)}{H}S_0 - Q(\Delta\rho)\Delta S. \tag{4.2}$$

Here τ_r is the relaxation time of ΔT to its reference value θ, S_0 is a reference salinity, and H is the depth of the model ocean. $F(s)$ is the freshwater flux, modeling imbalances between evaporation (which dominates at low latitudes) and precipitation (which dominates at high latitudes). The dynamics of ΔT and ΔS are coupled via the density difference $\Delta\rho$, approximated by the linearized equation of state

$$\Delta\rho = \alpha_S \Delta S - \alpha_T \Delta T, \tag{4.3}$$

which induces an exchange of mass $Q(\Delta\rho)$ between the boxes. We will use here Cessi's model [11] for Q,

$$Q(\Delta\rho) = \frac{1}{\tau_d} + \frac{q}{V}\Delta\rho^2, \tag{4.4}$$

where τ_d is the diffusion time scale, q the Poiseuille transport coefficient and V the volume of the box. Stommel uses a different relation, with $\Delta\rho^2$ replaced by $|\Delta\rho|$, but we will not make this choice here because it leads to a singularity (and thus adds some technical difficulties which do not lie at the heart of our discussion here).

Using the dimensionless variables $y = \alpha_S \Delta S/(\alpha_T \theta)$, $z = \Delta T/\theta$ and rescaling time by a factor τ_d, (4.1) and (4.2) can be rewritten as

$$\begin{aligned} \dot{y} &= p(s) - y\bigl[1 + \eta^2(y-z)^2\bigr] \\ \varepsilon_0 \dot{z} &= -(z-1) - \varepsilon_0 z\bigl[1 + \eta^2(y-z)^2\bigr], \end{aligned} \tag{4.5}$$

where $\varepsilon_0 = \tau_\mathrm{r}/\tau_\mathrm{d}$, $\eta^2 = \tau_\mathrm{d}(\alpha_T\theta)^2 q/V$, and $p(s)$ is proportional to the freshwater flux $F(s)$, with a factor $\alpha_S S_0 \tau_\mathrm{d}/(\alpha_T \theta H)$. Cessi uses the estimates $\eta^2 \simeq 7.5$, $\tau_\mathrm{r} \simeq 25$ days and $\tau_\mathrm{d} \simeq 219$ years. This yields $\varepsilon_0 \simeq 3 \times 10^{-4}$, implying that (4.5) is a slow–fast system. Tihonov's theorem [36] allows us to reduce the dynamics to the attracting slow manifold $z = 1 + \mathcal{O}(\varepsilon_0)$. To leading order, we thus find

$$\dot{y} = -y\bigl[1 + \eta^2(y-1)^2\bigr] + p(s). \tag{4.6}$$

Stochasticity shows up in this model through the weather-dependent term $p(s)$. To model long-scale variations in the typical weather, we will assume that $p(s)$ can be represented as the sum of a periodic term $\bar{p}(s)$ and white noise, where the period $1/\varepsilon$ of $\bar{p}(s)$ is much longer than the diffusion time, which equals 1. (Recall that we have already rescaled time by a factor of τ_d.) We thus obtain the SDE

$$\mathrm{d}y_s = f(y_s, s)\,\mathrm{d}s + \sigma_0\,\mathrm{d}W_s, \qquad \text{where} \qquad f(y,s) = -y\bigl[1 + \eta^2(y-1)^2\bigr] + \bar{p}(s). \tag{4.7}$$

Here we have been a bit sloppy. Actually, one should consider the effect of noise on the two-dimensional system (4.5) before passing to the reduced dynamics (4.6), see [9] for estimates on the typical deviation from the slow manifold due to noise and related results.

Note that f has an inflection point at $y = 2/3$, and that

$$\eta f\Bigl(\frac{2}{3} + \frac{x}{\eta}, s\Bigr) = \eta\Bigl[\bar{p}(s) - \frac{2}{3} - \frac{2}{27}\eta^2\Bigr] + \Bigl[\frac{1}{3}\eta^2 - 1\Bigr]x - x^3, \tag{4.8}$$

which derives from the Ginzburg–Landau potential (1.8) with parameters $\mu = (\eta^2/3 - 1)$ and $\lambda(s) = \eta(\bar{p}(s) - 2/3 - 2\eta^2/27)$. As we already know, the potential has two wells if and only if $\lambda^2 < \lambda_\mathrm{c}^2 = 4\mu^3/27$, which means, for $\eta^2 = 7.5$, that $\bar{p} \in [0.96, 1.48]$. The double-well potential is symmetric for $\bar{p} = \bar{p}_0 = 2/3 + 2\eta^2/27 \simeq 1.22$.

For a deterministic forcing given by $\lambda(s) = K\cos 2\pi\varepsilon s$, the SDE for $x = \eta(y - 2/3)$ becomes, on the time scale $1/\varepsilon$,

$$\mathrm{d}x_t = \frac{1}{\varepsilon}\bigl[\mu x - x^3 + K\cos 2\pi t\bigr]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \tag{4.9}$$

where $\sigma = \sigma_0\eta$. This SDE is of the same form as (3.8). While in Section 3, we assumed $K < \lambda_\mathrm{c}$, we will now allow K to exceed λ_c, so that the difference $a_0 = K - \lambda_\mathrm{c}$ may change sign. (Note that in Section 3, a_0 had the opposite sign.)

In the deterministic case $\sigma = 0$, Equation (4.9) has also been used to model a laser [23], and a similar equation describes the dynamics of a mean-field Curie–Weiss ferromagnet [37]. In the limit of infinitely slow forcing, solutions always remain in the same potential well if $K < \lambda_\mathrm{c}$. If $K > \lambda_\mathrm{c}$, however, the well tracked by x_t disappears in a saddle–node

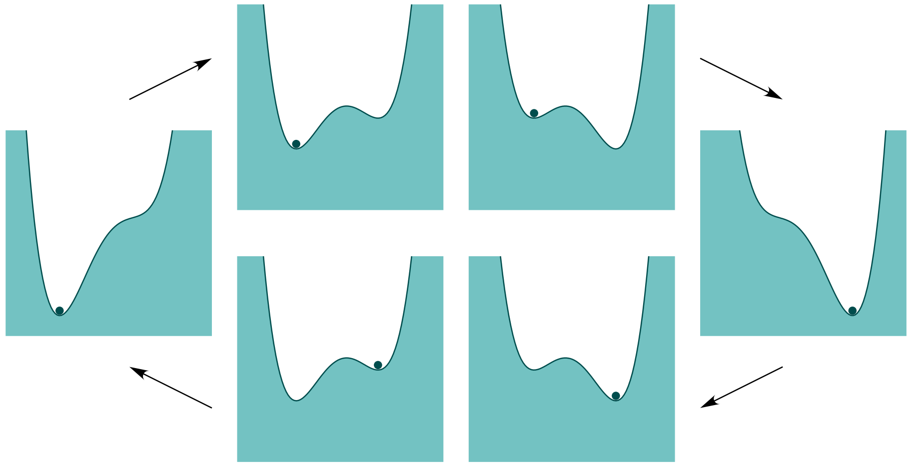

FIGURE 4. The potential $V(x,t) = \frac{1}{4}x^4 - \frac{1}{2}x^2 - \lambda(t)x$, with $\lambda(t) = K\cos(2\pi t)$, when K exceeds the threshold λ_c. In the deterministic case, with $\varepsilon \ll 1$, the overdamped particle jumps to a new well whenever $|\lambda(t)|$ becomes larger than λ_c, leading to hysteresis. Larger values of ε increase the size of hysteresis cycles, but additive noise of sufficient intensity decreases the size of typical cycles, because it advances transitions to the deeper well.

bifurcation when $|\lambda(t)|$ crosses λ_c from below, causing x_t to jump to the other well, which leads to hysteresis, see Figure 4.

For positive ε, the system does not react immediately to changes in the potential, so that the hysteresis cycles are deformed. One can show [23, 10] that

- For $K \leqslant \lambda_c + \mathcal{O}(\varepsilon)$, x_t always tracks the same potential well, at a distance at most of order $\varepsilon/\sqrt{|a_0|}$ if $a_0 \leqslant -\varepsilon$, and of order $\sqrt{\varepsilon}$ if $|a_0|$ is of order ε.
- For $K \geqslant \lambda_c + \mathcal{O}(\varepsilon)$, x_t is attracted by a hysteresis cycle, which is larger than the static hysteresis cycle; in particular, x_t crosses the λ-axis when $\lambda(t) = K\cos 2\pi t$ reaches a value λ^0 which satisfies the scaling relation

$$|\lambda^0| - \lambda_c \asymp \varepsilon^{2/3} a_0^{1/3}, \qquad \text{with } a_0 = K - \lambda_c. \tag{4.10}$$

Additive noise will also influence the shape of hysteresis cycles, because it can kick the state over the potential barrier, as has been noted in [28] in the context of the thermohaline circulation. For positive noise intensities σ, the value λ^0 at which x_t crosses the λ-axis, becomes a random variable. Assume for instance that we start at time $t_0 = 1/4$ in the right-hand potential well. We define the random crossing time and field, respectively, by

$$\tau^0(\omega) = \inf\Big\{t \in \Big[\frac{1}{4}, \frac{3}{4}\Big] : x_t(\omega) < 0\Big\}, \qquad \lambda^0(\omega) = \lambda(\tau^0(\omega)). \tag{4.11}$$

We thus have $\tau^0 \in [\frac{1}{4}, \frac{3}{4}]$ and $\lambda^0 \in [-K, K]$, whenever there is a crossing during the time interval $[\frac{1}{4}, \frac{3}{4}]$. We will indicate the parameter-dependence by $\lambda^0 = \lambda^0(\varepsilon, \sigma)$, keeping in mind that this random variable also depends on a_0 and μ. In the deterministic case, there is no crossing if $K \leqslant \lambda_c + \mathcal{O}(\varepsilon)$, and $\lambda^0(\varepsilon, 0)$ satisfies (4.10) if $K \geqslant \lambda_c + \mathcal{O}(\varepsilon)$.

As we know from the previous section, for $K < \lambda_c$, there is an amplitude-dependent threshold noise level σ_c such that during one period, x_t is unlikely to cross the potential barrier for $\sigma \ll \sigma_c$, while it is likely to cross it for $\sigma \gg \sigma_c$. In fact, in the latter case, there is a large probability to cross the barrier a time of order $\sigma^{2/3}$ before the instant $t = 1/2$ of minimal barrier height, when λ is of order $\lambda_c - \sigma^{4/3}$. In that case, the hysteresis cycle

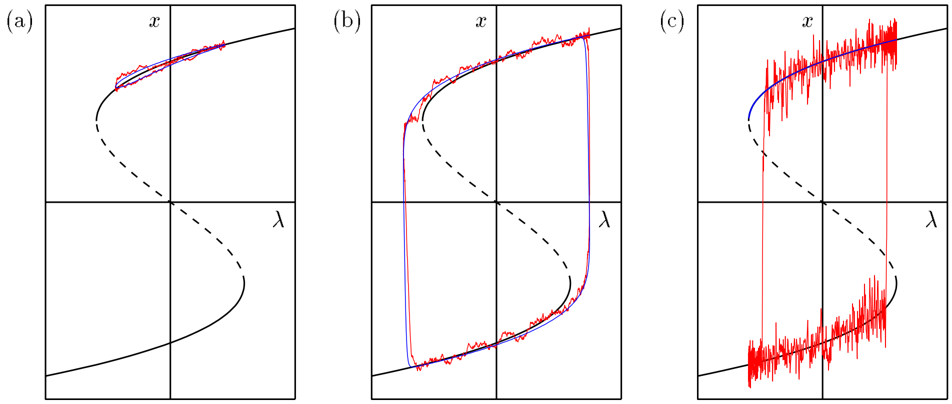

FIGURE 5. Typical random hysteresis "cycles" in the three parameter regimes. (a) Case I: Driving amplitude K and noise intensity σ are too small to allow the path to switch potential wells. (b) Case II: For large amplitude but weak noise, the path tracks the deterministic hysteresis cycle, which is larger than the static one. (c) Case III: For sufficiently strong noise, the path can overcome the potential barrier, so that typical hysteresis cycles are smaller than the static one.

will be *smaller* than the static cycle. A similar distinction between a small-noise and a large-noise regime exists for large-amplitude forcing.

It turns out that the distribution of the random variable λ^0 can be of three different types, depending on the values of the parameters (cf. Figure 5 and Figure 6):

- **Case I – Small-amplitude regime:** $a_0 \leqslant \textit{const}\, \varepsilon$ and $\sigma \leqslant (|a_0| \vee \varepsilon)^{3/4}$.
  Then x_t is unlikely to cross the potential barrier, and there are constants $C, \kappa > 0$ such that the probability to observe a crossing during the time interval $[\frac{1}{4}, \frac{3}{4}]$ is bounded above by (see [6, Theorem 2.3])
  $$\frac{C}{\varepsilon}\, \mathrm{e}^{-\kappa(|a_0|\vee\varepsilon)^{3/2}/\sigma^2}\,. \tag{4.12}$$
  The probability to observe a "macroscopic" hysteresis cycle is very small, as most paths are concentrated in a small neighbourhood of the bottom of the right-hand potential well (Figure 5a).
- **Case II – Large-amplitude regime:** $a_0 \geqslant \textit{const}\, \varepsilon$ and $\sigma \leqslant (\varepsilon\sqrt{a_0})^{1/2}$.
  This regime is actually the most difficult to study, since the deterministic solution jumps when $|\lambda(t)| - \lambda_{\mathrm{c}} \asymp (\varepsilon\sqrt{a_0})^{2/3}$, and crosses a zone of instability before reaching the left-hand potential well. One can show, however, that $|\lambda^0|$ is concentrated in an interval of length of order $(\varepsilon\sqrt{a_0})^{2/3}$ around the deterministic value [6, Theorem 2.4]. More precisely, there are constants $C, \kappa > 0$ such that
  $$\mathbb{P}\bigl\{|\lambda^0| < \lambda_{\mathrm{c}} - L\bigr\} \leqslant \frac{C}{\varepsilon}\, \mathrm{e}^{-\kappa L^{3/2}/\sigma^2} \tag{4.13}$$
  for $(\varepsilon\sqrt{a_0})^{2/3} \leqslant L \leqslant L_0/|\log(\varepsilon\sqrt{a_0})|$, and
  $$\mathbb{P}\bigl\{|\lambda^0| < \lambda_{\mathrm{c}} + L_1(\varepsilon\sqrt{a_0})^{2/3}\bigr\} \leqslant \frac{C}{\varepsilon}\, \mathrm{e}^{-\kappa\varepsilon\sqrt{a_0}/\sigma^2}, \tag{4.14}$$
  where the constants $L_0, L_1 > 0$ are independent of the small parameters. Hence it is unlikely to observe a substantially smaller absolute value of the crossing field $|\lambda^0|$

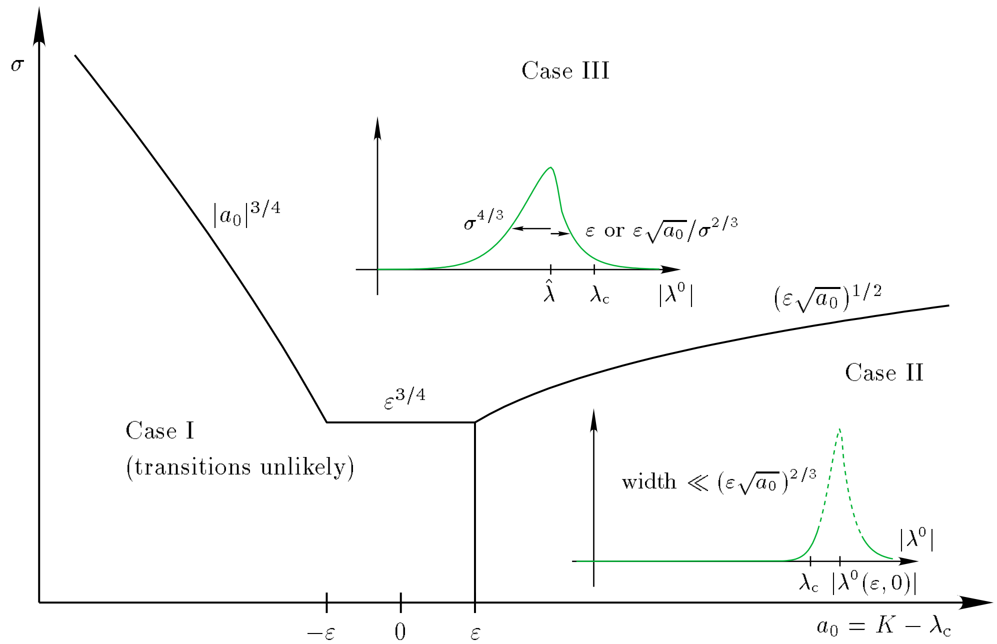

FIGURE 6. The three hysteresis regimes, shown in the plane driving-amplitude–noise-intensity, for fixed driving frequency. The insets sketch the distribution of the random value λ^0 of the forcing $\lambda(t)$ when x_t changes sign for the first time. In Case I, such transitions are unlikely. In Case II, $|\lambda^0|$ is concentrated in an interval $[\lambda_c + L_1(\varepsilon\sqrt{a_0})^{2/3}, \lambda_c + L_2(\varepsilon\sqrt{a_0})^{2/3}]$ containing the deterministic value $|\lambda^0(\varepsilon, 0)|$. The broken curve indicates that we do not control the distribution inside this interval. In Case III, $|\lambda^0|$ is concentrated around a value $\hat\lambda$ which is smaller than λ_c by an amount of order $\sigma^{4/3}$. The distribution decays faster to the right, with a width of order ε (actually, $\varepsilon|\log\sigma|$) if $a_0 \leqslant \varepsilon$ or $\sigma \geqslant |a_0|^{3/4}$, and of order $\varepsilon\sqrt{a_0}/\sigma^{2/3}$ if $a_0 \geqslant \varepsilon$ and $\sigma \leqslant |a_0|^{3/4}$.

than the deterministic one, provided $\sigma \ll (\varepsilon\sqrt{a_0})^{1/2}$. On the other hand, there is a constant $L_2 > L_1$ such that

$$\mathbb{P}\bigl\{|\lambda^0| > \lambda_c + L\bigr\} \leqslant 3\,\mathrm{e}^{-\kappa L/(\sigma^2(\varepsilon\sqrt{a_0})^{2/3}|\log(\varepsilon\sqrt{a_0})|)} \tag{4.15}$$

for all $L \geqslant L_2(\varepsilon\sqrt{a_0})^{2/3}$. As a consequence, the vast majority of hysteresis cycles will look very similar to the deterministic ones, which are slightly *larger* than the static hysteresis cycle (Figure 5b).

- **Case III – Large-noise regime:** Either $a_0 \leqslant \varepsilon$ and $\sigma \geqslant (|a_0| \vee \varepsilon)^{3/4}$ or $a_0 \geqslant \varepsilon$ and $\sigma \geqslant (\varepsilon\sqrt{a_0})^{1/2}$.
In this case, the noise is sufficiently strong to drive x_t over the potential barrier, with large probability, some time before the barrier is lowest or vanishes, leading to a *smaller* hysteresis cycle than in the deterministic case (Figure 5c). It turns out that the absolute value $|\lambda^0|$ of the crossing field is always concentrated around a (deterministic) value $\hat\lambda$ satisfying $\lambda_c - \hat\lambda \asymp \sigma^{4/3}$. It follows from [6, Proposition 5.1] that

$$\mathbb{P}\bigl\{|\lambda^0| < \hat\lambda - L\bigr\} \leqslant \frac{C}{\varepsilon}\,\mathrm{e}^{-\kappa L^{3/2}/\sigma^2} + \frac{3}{2}\,\mathrm{e}^{-\kappa\sigma^{4/3}/(\varepsilon|\log\sigma|)} \tag{4.16}$$

for $0 \leqslant L \leqslant \hat\lambda$ and

$$\mathbb{P}\bigl\{|\lambda^0| > \hat\lambda + L\bigr\} \leqslant \frac{3}{2}\,\mathrm{e}^{-\kappa L/(\varepsilon|\log\sigma|)} \tag{4.17}$$

for positive L up to $K - \hat{\lambda}$ if $a_0 \leqslant \varepsilon$. If $a_0 \geqslant \varepsilon$, the same bound holds for $L \leqslant \lambda_c - \hat{\lambda}$, while the behaviour for larger L is described by (4.15). The estimates (4.16) and (4.17) hold if $a_0 \leqslant \varepsilon$ or $\sigma > a_0^{3/4}$. In the other case, two exponents are modified: $\sigma^{4/3}/(\varepsilon|\log\sigma|)$ is replaced by $\sigma^2/(\varepsilon\sqrt{a_0}|\log\sigma|)$, and $L/(\varepsilon|\log\sigma|)$ is replaced by $\sigma^{2/3}L/(\varepsilon\sqrt{a_0}|\log\sigma|)$.

Note that in all cases, the distribution of the crossing field λ^0 decays faster to the right than to the left of $\hat{\lambda}$, and it is unlikely to observe λ^0 larger than λ_c, except when approaching the lower boundary of Region III.

In some physical applications, for instance in ferromagnets, the area enclosed by hysteresis cycles represents the energy dissipation per period. The distribution of the random hysteresis area can also be described, and bounds on its expectation and variance can be obtained. We refer to [6] and [5, Section 4] for details.

For Stommel's box model, the above properties have two important consequences. First, noise can drive the system from one stable equilibrium to the other *before* the potential barrier between them disappears, so that a smaller deviation from the mean freshwater flux than expected from the deterministic analysis can switch the system's state. Second, this early switching to the other state is likely only if the noise intensity exceeds a threshold value (which is lowest when the amplitude K is close to λ_c). Still, the system spends roughly half of the time per period in metastable equilibrium in the shallower well.

5 Delay

Convective motions in the atmosphere can be simulated in a laboratory experiment known as Rayleigh–Bénard convection. A fluid contained between two horizontal plates is heated from below. For low heating, the fluid remains at rest. Above a threshold, stationary convection rolls develop. With increasing energy supply, the angular velocity of the rolls becomes time-dependent, first periodically, and then, after a sequence of bifurcations depending on the geometry of the set-up, chaotic. For still stronger heating, the convection rolls are destroyed and the dynamics becomes turbulent.

Lorenz' famous model [26] uses a three-modes Galerkin approximation of the hydrodynamic equations. The amplitudes of these modes obey the ODEs

$$\begin{aligned} \dot{X} &= \mathrm{Pr}(Y - X) \\ \dot{Y} &= rX - Y - XZ \\ \dot{Z} &= -bZ + XY. \end{aligned} \tag{5.1}$$

Here X measures the angular velocity of convection rolls, while Y and Z parametrize the temperature field. The Prandtl number $\mathrm{Pr} > 0$ is a characteristic of the fluid, b depends on the geometry of the container, and r is proportional to the heating.

For $0 \leqslant r \leqslant 1$, the origin $(X, Y, Z) = (0, 0, 0)$ is a global attractor of the system, corresponding to the fluid at rest. At $r = 1$, this state becomes unstable in a pitchfork bifurcation. Two new stable equilibrium branches $(\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$ are created, which correspond to convection rolls with the two possible directions of rotation. We will focus on this simplest bifurcation, ignoring all the other sequences of bifurcations ultimately leading to a strange attractor (see for instance [32]).

We are interested in the situation where $r = r(\varepsilon s)$ grows monotonously through $r(0) = 1$ with low speed ε (e. g. $r = 1 + \varepsilon s$). Near the bifurcation point, one can reduce the system

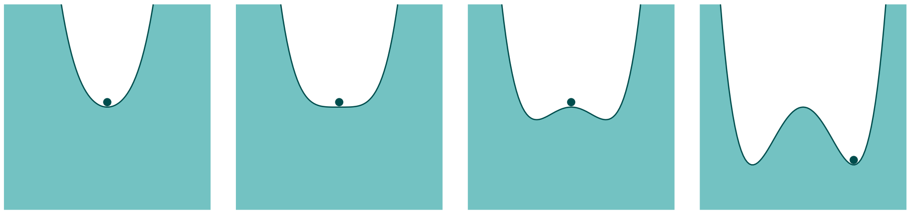

FIGURE 7. The potential $V(x,t) = \frac{1}{4}x^4 - \frac{1}{2}\mu(t)x^2$ transforms, as μ changes from negative to positive, from a single-well to a double-well potential. In the deterministic case, an overdamped particle stays close to the saddle for a macroscopic time before falling into one of the wells. Noise tends to reduce this delay.

to an invariant center manifold, on which the dynamics is governed (cf. [10]), after scaling time by a factor ε, by the one-dimensional equation

$$\varepsilon \frac{\mathrm{d}x}{\mathrm{d}t} = \mu(t)x + c(t)x^3 + \mathcal{O}(x^5). \tag{5.2}$$

Here $\mu(t) = a(t) + \mathcal{O}(\varepsilon)$, where $a(t) = \frac{1}{2}\bigl[-(\mathrm{Pr}+1) + \sqrt{(\mathrm{Pr}+1)^2 + 4\mathrm{Pr}(r(t)-1)}\bigr]$ is the largest eigenvalue of the linearization of (5.1) at 0, which has the same sign as $r(t)-1$, and $c(t)$ is negative and bounded away from zero. The right-hand side of (5.2) derives from a potential similar to the Ginzburg–Landau potential (1.8) with $\lambda = 0$, which remains symmetric while transforming from a single-well to a double-well as $\mu(t)$ becomes positive, see Figure 7.

The solution of (5.2) with initial condition $x_0 > 0$ for $t_0 < 0$ can be written in the form

$$x_t = \varphi(x_0,t)\,\mathrm{e}^{\alpha(t,t_0)/\varepsilon}, \qquad \alpha(t,t_0) = \int_{t_0}^{t} \mu(s)\,\mathrm{d}s, \tag{5.3}$$

with $0 < \varphi(x_0,t) \leqslant x_0$ for all t. Thus x_t is exponentially small if $\alpha(t,t_0)$ is negative. The important point to note is that $\alpha(t,t_0)$ can be negative even when $a(t)$ is positive. For instance, if $\mu(s) = s$, then $\alpha(t,t_0) = \frac{1}{2}(t^2 - t_0^2)$ is negative for $t_0 < t < |t_0|$. Thus x_t will remain exponentially close to the saddle at $x=0$ up to time $|t_0|$ after crossing the bifurcation point. This phenomenon is called *bifurcation delay*. It means that when r is slowly increased, convection rolls will not appear at $r=1$, as expected from the static analysis, but only for some larger value of r, which depends on the initial condition.

It is clear that the existence of a delay depends crucially on the fact that x_t can approach the saddle exponentially closely, where the repulsion is very small. Noise present in the system will help kicking x_t away from the saddle, and thus reduce the delay. The question is to determine how the delay depends on the noise intensity σ.

For brevity, we will illustrate the results in the particular case of a Ginzburg–Landau potential, with dynamics governed by the SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon}\bigl[\mu(t)x_t - x_t^3\bigr]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t. \tag{5.4}$$

The case without the term $-x_t^3$ has been analysed by several authors [38, 33, 35, 22], with the result that the typical bifurcation delay in the presence of noise behaves like $\sqrt{|\log\sigma|}$. The results in [7] cover more general nonlinearities than $-x^3$.

We assume that $\mu(t)$ is increasing, and satisfies $\mu(0) = 0$, $\mu'(0) \geqslant$ *const* > 0. For simplicity, we consider first the case where x_t starts at a time $t_0 < 0$ at the origin $x_0 = 0$.

From the results of Section 2, we expect the paths to remain concentrated, for some time, in a set whose width is related to the linearization of (5.4) around $x = 0$. We define the function

$$\bar{v}(t) = \bar{v}_0 \, \mathrm{e}^{2\alpha(t,t_0)/\varepsilon} + \frac{\sigma^2}{\varepsilon} \int_{t_0}^{t} \mathrm{e}^{2\alpha(t,s)/\varepsilon} \, \mathrm{d}s. \tag{5.5}$$

For a suitably chosen $\bar{v}_0 \asymp \sigma^2/|\mu(t_0)|$, one can show that $\bar{v}(t)$ is increasing and satisfies

$$\bar{v}(t) \asymp \begin{cases} \sigma^2/|\mu(t)| & \text{for } t_0 \leqslant t \leqslant -\sqrt{\varepsilon} \\ \sigma^2/\sqrt{\varepsilon} & \text{for } -\sqrt{\varepsilon} \leqslant t \leqslant \sqrt{\varepsilon} \\ \sigma^2 \, \mathrm{e}^{2\alpha(t)/\varepsilon} / \sqrt{\varepsilon} & \text{for } t \geqslant \sqrt{\varepsilon}. \end{cases} \tag{5.6}$$

Note that although the curvature $|\mu(t)|$ of the potential at the origin vanishes at time 0, $\bar{v}(t)$ grows slowly until time $\sqrt{\varepsilon}$ *after* the bifurcation point, and only then starts growing faster and faster.

We now introduce, as in Section 2, the strip

$$\mathcal{B}(h) = \bigl\{(x,t) \colon |x| \leqslant h\sqrt{\bar{v}(t)}\bigr\}. \tag{5.7}$$

Then one can show (see [7, Theorem 2.10]) the existence of a constant $h_0 > 0$ such that the first-exit time $\tau_{\mathcal{B}(h)}$ of x_t from the strip $\mathcal{B}(h)$ satisfies

$$\mathbb{P}\bigl\{\tau_{\mathcal{B}(h)} < t\bigr\} \leqslant C(t,\varepsilon) \, \mathrm{e}^{-\kappa h^2} \tag{5.8}$$

for all $h \leqslant h_0\sigma/\bar{v}(t)$, where

$$C(t,\varepsilon) = \frac{1}{\varepsilon^2} \int_{t_0}^{t} |\mu(s)| \, \mathrm{d}s + \mathcal{O}\biggl(\frac{1}{\varepsilon}\biggr), \qquad \text{and} \qquad \kappa = \frac{1}{2} - \mathcal{O}(\sqrt{\varepsilon}) - \mathcal{O}\biggl(\frac{h^2\bar{v}(t)^2}{\sigma^2}\biggr). \tag{5.9}$$

The paths are concentrated in the strip $\mathcal{B}(h)$, provided $h_0^2\sigma^2/\bar{v}(t)^2 \geqslant h^2 \gg \log C(t,\varepsilon)$. As a consequence, we can distinguish between three regimes, depending on noise intensity:

- **Regime I:** $\sigma \leqslant \mathrm{e}^{-K/\varepsilon}$ for some $K > 0$.
  The paths are concentrated near $x = 0$ at least as long as $2\alpha(t) \ll K$. This implies that there is still a macroscopic bifurcation delay.
- **Regime II:** $\mathrm{e}^{-1/\varepsilon^p} \leqslant \sigma \ll \sqrt{\varepsilon}$ for some $p < 1$.
  The paths are concentrated near $x = 0$ at least up to time $\sqrt{\varepsilon}$, with a typical spreading growing like $\sigma/\sqrt{|\mu(t)|}$ for $t \leqslant -\sqrt{\varepsilon}$, and remaining of order $\sigma/\varepsilon^{1/4}$ for $|t| \leqslant \sqrt{\varepsilon}$.
- **Regime III:** $\sigma \geqslant \sqrt{\varepsilon}$.
  The paths are concentrated near $x = 0$ at least up to time $-\sigma$, with a typical spreading growing like $\sigma/\sqrt{|\mu(t)|}$. Near $t = 0$, the potential becomes too flat to counteract the diffusion, and as t grows further, paths keep switching back and forth between the wells, before ultimately settling for a well.

Similar results hold if x_t starts, at $t_0 < 0$, away from $x = 0$, say in $x_0 > 0$. Then the strip $\mathcal{B}(h)$ is centred at the deterministic solution $x_t^{\det}$ (with the same initial condition), which jumps to the right-hand well when $\alpha(t,t_0)$ becomes positive, see Figure 8. In Regime I, with K sufficiently large, the majority of paths follow $x_t^{\det}$ into the right-hand potential well.

It remains to understand the behaviour after time $\sqrt{\varepsilon}$ in Regime II. To this end, we introduce the set

$$\mathcal{D}(\varrho) = \bigl\{(x,t) \colon t \geqslant \sqrt{\varepsilon}, |x| \leqslant \sqrt{(1-\varrho)\mu(t)}\,\bigr\}, \tag{5.10}$$

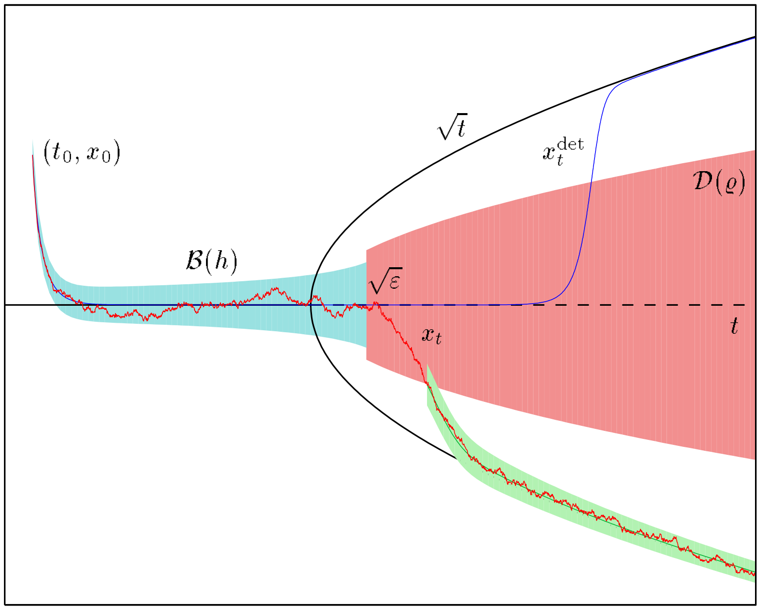

FIGURE 8. A sample path x_t of the SDE (5.4) with $\mu(t) = t$, for $\varepsilon = 0.01$ and $\sigma = 0.015$. The deterministic solution x_t^{det}, starting in $x_0 > 0$ at time t_0, jumps to the right-hand well, located at $x^\star(t) = \sqrt{t}$, at time $|t_0|$. Typical paths stay in the set $\mathcal{B}(h)$, whose width increases like $h\sigma/(\sqrt{|t|} \vee \varepsilon^{1/4})$, until time $\sqrt{\varepsilon}$ after the bifurcation. They leave the domain $\mathcal{D}(\varrho)$ (shown for $\varrho = 2/3$) at a random time $\tau = \tau_{\mathcal{D}(\varrho)}$, which is typically of order $\sqrt{\varepsilon|\log\sigma|}$. After leaving $\mathcal{D}(\varrho)$, each path is likely to stay in a strip of width of order $h\sigma/\sqrt{t}$, centred at a deterministic solution approaching either $+x^\star(t)$ or $-x^\star(t)$.

depending on a parameter $\varrho \in [0, 2/3)$, see Figure 8. The set $\mathcal{D}(0)$ contains the points lying between the two stable equilibrium branches $\pm\sqrt{\mu(t)}$. One can show (see [7, Theorem 2.11]) that if $\varrho \in (0, 2/3)$ and $\sigma|\log\sigma|^{3/2} = \mathcal{O}(\sqrt{\varepsilon})$, then the first-exit time $\tau_{\mathcal{D}(\varrho)}$ of x_t from $\mathcal{D}(\varrho)$ satisfies

$$\mathbb{P}\{\tau_{\mathcal{D}(\varrho)} \geqslant t\} \leqslant C(t,\varepsilon)\frac{|\log\sigma|}{\sigma}\frac{\mathrm{e}^{-\varrho\alpha(t,\sqrt{\varepsilon})/\varepsilon}}{\sqrt{1-\mathrm{e}^{-2\varrho\alpha(t,\sqrt{\varepsilon})/\varepsilon}}}, \tag{5.11}$$

where

$$C(t,\varepsilon) = const\ \mu(t)\biggl(1 + \frac{\alpha(t,\sqrt{\varepsilon})}{\varepsilon}\biggr). \tag{5.12}$$

The estimate (5.11) shows that paths are unlikely to stay in $\mathcal{D}(\varrho)$ as soon as t satisfies $\varrho\alpha(t,\sqrt{\varepsilon}) \gg \varepsilon|\log\sigma|$. Since α is quadratic in t, most paths will have left $\mathcal{D}(\varrho)$ for

$$t \gg \sqrt{\varepsilon|\log\sigma|}. \tag{5.13}$$

Once x_t has left $\mathcal{D}(\varrho)$, one can further show that it is likely to track a deterministic solution which approaches the bottom of one of the potential wells. Assume for instance that x_t leaves $\mathcal{D}(\varrho)$ through the upper boundary, at a random time $\tau = \tau_{\mathcal{D}(\varrho)}$. Then, for $1/2 < \varrho < 2/3$, [7, Theorem 2.12] shows that the deterministic solution $x_t^{\text{det},\tau}$, starting in the same point on the upper boundary at time τ, approaches the bottom of the well at $\sqrt{\mu(t)}$ like $\varepsilon/\mu(t)^{3/2} + \sqrt{\mu(\tau)}\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}$, where $\eta = 2 - 3\varrho$. The path x_t is likely to stay in a strip of width $\sigma/\sqrt{\mu(t)}$ around $x^{\text{det},\tau}$. Thus after another time span of length (5.13), most paths will have concentrated near the bottom of a potential well again.

We note that different kinds of metastability play a rôle here. First, paths remain concentrated for some time near the *unstable* saddle. Second, they will concentrate again near one of the potential wells after some time. Some paths will choose the left-hand well and others the right-hand well (with probability exponentially close to $1/2$ in Regime II), but all the paths which choose a given potential well are unlikely to cross the barrier again. In fact, one can show that if $\mu(t)$ grows at least linearly, then the probability *ever* to cross the saddle again after settling for a well, is of order $\mathrm{e}^{-const/\sigma^2}$. If we start the system at a *positive* t_0 in one of the wells, the distribution will never approach a symmetric bimodal one.

In the case of the Rayleigh–Bénard convection with slowly growing heat supply $r(\varepsilon s)$ and additive noise, these results mean that exponentially weak noise will not prevent the delayed appearance of convection rolls. For moderate noise intensity, rolls will appear after a delay of order $\sqrt{|\log \sigma|/\varepsilon}$, which is considerably shorter than the delay in the deterministic case which is of order $1/\varepsilon$. The direction of rotation is unlikely to change after another time span of that order. For strong noise, convection rolls may appear early, but their angular velocity will fluctuate around zero until a time of order σ/ε after the bifurcation before settling for a sign, and even then occasional changes of rotation direction are possible.

6 Concluding remarks

For climate models described by slowly time-dependent Langevin equations of the form (1.1), it is important to understand in which way the noise term affects the predictability of the system.

We showed that in certain cases, namely when the noise-free system admits stable equilibrium curves, the addition of weak noise will perturb most trajectories only locally. Although the randomness prevents us from exactly predicting the behaviour of individual solutions, the vast majority of sample paths remain, for exponentially long time spans, in strips or tubes shadowing the deterministic trajectory (defined in Equations (2.9) and (2.20)). We emphasize that the tracked equilibrium branch is selected by the initial condition, not by the noise, even if this equilibrium branch happens not to be the system's most stable one: this is the phenomenon of metastability.

Transitions between equilibrium branches become visible on subexponential timescales only in the vicinity of bifurcation points. As a rule of thumb, transitions become likely once the strips or tubes containing typical sample paths start to overlap neighbouring unstable equilibrium branches; in this case another, more stable equilibrium can be selected. Depending on the kind of time-dependence, phenomena such as hysteresis and close-to-periodic oscillations can be observed.

It is worth noting that many qualitative properties of the transition process (transition probability, size of transition windows) exhibit universal power-law behaviour, with exponents depending only on the type of bifurcation or avoided bifurcation (e. g. saddle–node or pitchfork). It is thus of fundamental importance to obtain a good understanding of the bifurcation diagram of the associated deterministic system before investigating the effect of additive noise.

Although most of the results presented here apply to simple one-dimensional systems, the method can be generalized to multidimensional, fully coupled slow–fast systems with state-dependent noise [9]. While the case of stable slow manifolds is well understood, the possibility of more complicated motion in many dimensions (starting with periodic orbits) still offers interesting challenges for further investigation.

Acknowledgements

We thank the organisers for the invitation to Chorin and the opportunity to present our results during the *Second Workshop on Stochastic Climate Models*. We enjoyed stimulating discussions in a pleasant atmosphere. In addition, we thank the anonymous climatologist for her or his detailed and encouraging referee report which led to further improvements in the presentation.

References

[1] L. Arnold. *Random Dynamical Systems*. Springer-Verlag, Berlin, 1998.

[2] L. Arnold. Hasselmann's program revisited: The analysis of stochasticity in deterministic climate models. In P. Imkeller and J.-S. von Storch, editors, *Stochastic Climate Models*, volume 49 of *Progress in Probability*, pages 141–158, Boston, 2001. Birkhäuser.

[3] R. Azencott. Petites perturbations aléatoires des systèmes dynamiques: développements asymptotiques. *Bull. Sci. Math. (2)*, 109:253–308, 1985.

[4] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. A theory of stochastic resonance in climatic change. *SIAM J. Appl. Math.*, 43(3):565–578, 1983.

[5] N. Berglund and B. Gentz. Beyond the Fokker–Planck equation: Pathwise control of noisy bistable systems. *J. Phys. A*, 35(9):2057–2091, 2002.

[6] N. Berglund and B. Gentz. The effect of additive noise on dynamical hysteresis. *Nonlinearity*, 15(3):605–632, 2002.

[7] N. Berglund and B. Gentz. Pathwise description of dynamic pitchfork bifurcations with additive noise. *Probab. Theory Related Fields*, 122(3):341–388, 2002.

[8] N. Berglund and B. Gentz. A sample-paths approach to noise-induced synchronization: Stochastic resonance in a double-well potential. *Ann. Appl. Probab.*, 12:1419–1470, 2002.

[9] N. Berglund and B. Gentz. Geometric singular perturbation theory for stochastic differential equations. *J. Differential Equations*, 191:1–54, 2003.

[10] N. Berglund and H. Kunz. Memory effects and scaling laws in slowly driven systems. *J. Phys. A*, 32(1):15–39, 1999.

[11] P. Cessi. A simple box model of stochastically forced thermohaline flow. *J. Phys. Oceanogr.*, 24:1911–1920, 1994.

[12] H. Crauel and F. Flandoli. Attractors for random dynamical systems. *Probab. Theory Related Fields*, 100(3):365–393, 1994.

[13] H. Crauel and F. Flandoli. Additive noise destroys a pitchfork bifurcation. *J. Dynam. Differential Equations*, 10(2):259–274, 1998.

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

[15] W. H. Fleming and M. R. James. Asymptotic series and exit time probabilities. *Ann. Probab.*, 20(3):1369–1384, 1992.

[16] R. F. Fox. Stochastic resonance in a double well. *Phys. Rev. A*, 39:4148–4153, 1989.

[17] M. I. Freidlin. Quasi-deterministic approximation, metastability and stochastic resonance. *Physica D*, 137:333–352, 2000.

[18] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag, New York, second edition, 1998.

[19] I. S. Gradšteĭn. Application of A. M. Lyapunov's theory of stability to the theory of differential equations with small coefficients in the derivatives. *Mat. Sbornik N. S.*, 32(74):263–286, 1953.

[20] K. Hasselmann. Stochastic climate models. Part I. Theory. *Tellus*, 28:473–485, 1976.

[21] W. Horsthemke and R. Lefever. *Noise-induced transitions*. Springer-Verlag, Berlin, 1984.

[22] K. M. Jansons and G. D. Lythe. Stochastic calculus: application to dynamic bifurcations and threshold crossings. *J. Statist. Phys.*, 90(1–2):227–251, 1998.

[23] P. Jung, G. Gray, R. Roy, and P. Mandel. Scaling law for dynamical hysteresis. *Phys. Rev. Letters*, 65:1873–1876, 1990.

[24] P. Jung and P. Hänggi. Stochastic nonlinear dynamics modulated by external periodic forces. *Europhys. Letters*, 8:505–510, 1989.

[25] Y. Kifer. The exit problem for small random perturbations of dynamical systems with a hyperbolic fixed point. *Israel J. Math.*, 40(1):74–96, 1981.

[26] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sciences*, 20:130–141, 1963.

[27] B. McNamara and K. Wiesenfeld. Theory of stochastic resonance. *Phys. Rev. A*, 39:4854–4869, 1989.

[28] A. H. Monahan. Stabilisation of climate regimes by noise in a simple model of the thermohaline circulation. *Journal of Physical Oceanography*, 32:2072–2085, 2002.

[29] F. Moss and K. Wiesenfeld. The benefits of background noise. *Scientific American*, 273:50–53, 1995.

[30] S. Rahmstorf. Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. *Nature*, 378:145–149, 1995.

[31] B. Schmalfuß. Invariant attracting sets of nonlinear stochastic differential equations. In H. Langer and V. Nollau, editors, *Markov processes and control theory (Gaußig, 1988)*, volume 54 of *Math. Res.*, pages 217–228, Berlin, 1989. Akademie-Verlag.

[32] C. Sparrow. *The Lorenz Equations: Bifurcations, Chaos and Strange Attractors*. Springer-Verlag, New York, 1982.

[33] N. G. Stocks, R. Manella, and P. V. E. McClintock. Influence of random fluctuations on delayed bifurcations: The case of additive white noise. *Phys. Rev. A*, 40:5361–5369, 1989.

[34] H. Stommel. Thermohaline convection with two stable regimes of flow. *Tellus*, 13:224–230, 1961.

[35] J. B. Swift, P. C. Hohenberg, and G. Ahlers. Stochastic Landau equation with time-dependent drift. *Phys. Rev. A*, 43:6572–6580, 1991.

[36] A. N. Tihonov. Systems of differential equations containing small parameters in the derivatives. *Mat. Sbornik N. S.*, 31:575–586, 1952.

[37] T. Tomé and M. J. de Oliveira. Dynamic phase transition in the kinetic Ising model under a time-dependent oscillating field. *Phys. Rev. A*, 41:4251–4254, 1990.

[38] M. C. Torrent and M. San Miguel. Stochastic-dynamics characterization of delayed laser threshold instability with swept control parameter. *Phys. Rev. A*, 38:245–251, 1988.

[39] K. Wiesenfeld and F. Moss. Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs. *Nature*, 373:33–36, 1995.

Nils Berglund
Department of Mathematics, ETH Zürich
ETH Zentrum, 8092 Zürich, Switzerland

Current address
FRUMAM, CPT–CNRS Luminy
Case 907, 13288 Marseille Cedex 9, France
and
PHYMAT, Université de Toulon
B. P. 132, 83957 La Garde Cedex, France
E-mail address: berglund@cpt.univ-mrs.fr

Barbara Gentz
Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstraße 39, 10117 Berlin, Germany
E-mail address: gentz@wias-berlin.de