Pathwise description of dynamic pitchfork bifurcations with additive noise

Nils Berglund and Barbara Gentz

Abstract

The slow drift (with speed ε) of a parameter through a pitchfork bifurcation point, known as the dynamic pitchfork bifurcation, is characterized by a significant delay of the transition from the unstable to the stable state. We describe the effect of an additive noise, of intensity σ, by giving precise estimates on the behaviour of the individual paths. We show that until time $\sqrt{\varepsilon}$ after the bifurcation, the paths are concentrated in a region of size $\sigma/\varepsilon^{1/4}$ around the bifurcating equilibrium. With high probability, they leave a neighbourhood of this equilibrium during a time interval $[\sqrt{\varepsilon}, c\sqrt{\varepsilon|\log\sigma|}\,]$, after which they are likely to stay close to the corresponding deterministic solution. We derive exponentially small upper bounds for the probability of the sets of exceptional paths, with explicit values for the exponents.

Date. August 4, 2000. *Revised.* April 19, 2001.

2000 *Mathematics Subject Classification.* 37H20, 60H10 (primary), 34E15, 93E03 (secondary).

Keywords and phrases. Dynamic bifurcation, pitchfork bifurcation, additive noise, bifurcation delay, singular perturbations, stochastic differential equations, dynamical systems, pathwise description, concentration of measure.

1 Introduction

Physical systems are often described by ordinary differential equations (ODEs) of the form

$$\frac{\mathrm{d}x}{\mathrm{d}s} = f(x,\lambda), \tag{1.1}$$

where x is the state of the system, λ a parameter, and s denotes time. The model (1.1) may however be too crude, since it neglects all kinds of perturbations acting on the system. We are interested here in the combined effect of two perturbations: a slow drift of the parameter, and an additive noise.

A slowly drifting parameter $\lambda = \varepsilon s$, (with $\varepsilon \ll 1$), may model the deterministic change in time of some exterior influence, such as the climate acting on an ecosystem or a magnetic field acting on a ferromagnet. Obviously, nontrivial dynamics can only be expected when λ is allowed to vary by an amount of order 1, and thus the system has to be considered on the time scale ε^{-1}. This is usually done by introducing the *slow time* $t = \varepsilon s$, which transforms (1.1) into the singularly perturbed equation

$$\varepsilon\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t). \tag{1.2}$$

It is known that solutions of this system tend to stay close to stable equilibrium branches of f [Gr, Ti], see Fig. 1a. New, and sometimes surprising phenomena occur when such an

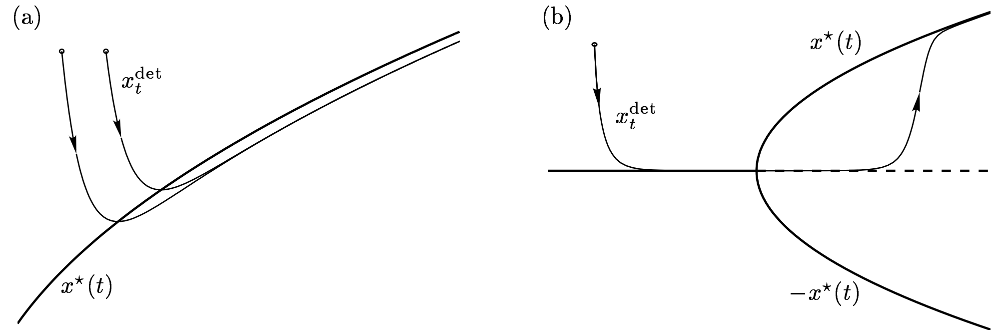

FIGURE 1. Solutions of the slowly time-dependent equation (1.2) represented in the (t, x)-plane. (a) Stable case: A stable equilibrium branch $x^\star(t)$ attracts nearby solutions x_t^{det}. Two solutions with different initial conditions are shown. They converge exponentially fast to each other, as well as to a neighbourhood of order ε of $x^\star(t)$. (b) Pitchfork bifurcation: The stable equilibrium $x = 0$ becomes unstable at $t = 0$ (broken line) and expels two stable equilibrium branches $\pm x^\star(t)$. A solution x_t^{det} is shown, which is attracted by $x = 0$, and stays close to the origin for a finite time after the bifurcation. This phenomenon is known as bifurcation delay.

equilibrium branch undergoes a bifurcation. These phenomena are usually called *dynamic bifurcations* [Ben][1]. In the case of the Hopf bifurcation, when the equilibrium gets unstable while expelling a stable periodic orbit, the bifurcation is substantially delayed: solutions of (1.2) track the unstable equilibrium (for a *non-vanishing* time interval in the limit $\varepsilon \to 0$) before jumping to the limit cycle [Sh, Ne]. A similar phenomenon exists for the dynamic pitchfork bifurcation of an equilibrium without drift, the simplest example being $f(x, t) = tx - x^3$ (Fig. 1b). The delay has been observed experimentally, for instance, in lasers [ME] and in a damped rotating pendulum [BK].

These phenomena have the advantage of providing a genuinely dynamic point of view for the concept of a bifurcation. Although one often says that a bifurcation diagram (representing the asymptotic states of the system as a function of the parameter) is obtained by varying the control parameter λ, the impatient experimentalist taking this literally may have the surprise to discover unstable stationary states of the system (s)he investigates. The asymptotic state of the system (1.1) with slowly varying parameter $\lambda(\varepsilon s) = \lambda(t)$ may depend not only on the initial condition (x_0, t_0), but also on the history of variation of the parameter $\{\lambda(t)\}_{t \geqslant t_0}$.

The perturbation of (1.1) by additive noise can be modeled by a stochastic differential equation (SDE) of the form

$$\mathrm{d}x_s = f(x_s, \lambda)\,\mathrm{d}s + \sigma\,\mathrm{d}W_s, \tag{1.3}$$

where W_s denotes the standard Wiener process, and σ measures the noise intensity. A widespread approach is to analyse the probability density of x_s, which satisfies the Fokker–Planck equation. In particular, if $-f$ can be written as the gradient of a *potential function* F, then there is a unique stationary density $p(x, \lambda) = \mathrm{e}^{-F(x,\lambda)/\sigma^2}/N$, where N is the normalization. This formula shows that for small noise intensity, the stationary density is sharply peaked around stable equilibria of f.

[1]Unfortunately, the term "dynamical bifurcation" is used in a different sense in the context of random dynamical systems, namely to describe a bifurcation of the family of invariant measures as opposed to a "phenomenological bifurcation", see for instance [Ar].

That method has, however, two major limitations. The first one is that the Fokker–Planck equation is difficult to solve, except in the linear and in the gradient case. The second limitation is more serious: the density gives no information on correlations in time, and even when the density is strongly localized, individual paths can perform large excursions. This is why other approaches are important. A classical one is based on the computation of first exit times from the neighbourhood of stable equilibria [FW, FJ].

The effect of bifurcations has been studied more recently by methods based on the concept of random attractors [CF94, Schm, Ar]. In particular, Crauel and Flandoli showed that according to their definition, "Additive noise destroys a pitchfork bifurcation" [CF98]. The physical interpretation of random attractors is, however, not straightforward, and alternative characterizations of stochastic bifurcations are desirable. In the same way a slowly varying parameter helps our understanding of bifurcations in the deterministic case, it can provide a new point of view in the case of random dynamical systems.

Let us consider the combined effect of a slowly drifting parameter and additive noise on the ODE (1.1). We will focus on the case of a pitchfork bifurcation, where the questions *How does the additive noise affect the bifurcation delay?* and *Where does the path go after crossing the bifurcation point?* are of major physical interest. The situation of the drift term f in (1.3) depending explicitly on time is considerably more difficult than the autonomous case, and thus much less understood. One can expect, however, that a slow time dependence makes the problem accessible to perturbation theory, and that one may take advantage of techniques developed to study singularly perturbed equations such as (1.2). With $\lambda = \varepsilon s$, Equation (1.3) becomes

$$\mathrm{d}x_s = f(x_s, \varepsilon s)\,\mathrm{d}s + \sigma\,\mathrm{d}W_s. \tag{1.4}$$

If we introduce again the slow time $t = \varepsilon s$, the Brownian motion is rescaled, resulting in the SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon} f(x_t, t)\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t. \tag{1.5}$$

Our analysis of (1.5) is restricted to one-dimensional x. The noise intensity σ should be considered as a function of ε. Indeed, since we now consider the equation on the time scale ε^{-1}, a constant noise intensity would lead to an infinite spreading of trajectories as $\varepsilon \to 0$. In the case of the pitchfork bifurcation, we will need to assume that $\sigma \ll \sqrt{\varepsilon}$.

Various particular cases of equation (1.5) have been studied before, from a mathematically non-rigorous point of view. In the linear case $f(x,\lambda) = \lambda x$, the distribution of first exit times was investigated and compared with experiments in [TM, SMC, SHA], while [JL] derived a formula for the last crossing of zero. In the case $f(x,\lambda) = \lambda x - x^3$, [Ga] studied the dependence of the delay on ε and σ numerically, while [Ku] considered the associated Fokker–Planck equation, the solution of which she approximated by a Gaussian ansatz.

In the present work, we analyse (1.5) for a general class of odd functions $f(x,\lambda)$ undergoing a pitchfork bifurcation. We use a different approach, based on a precise control of the *whole paths* $\{x_s\}_{t_0 \leqslant s \leqslant t}$ of the process. The results thus contain much more information than the probability density. It also turns out that the technique we use allows to deal with nonlinearities in quite a natural way. Our results can be summarized in the following way (see Fig. 2):

- Solutions of the deterministic equation (1.2) starting near a stable equilibrium branch of f are known to reach a neighbourhood of order ε of that branch in a time of order $\varepsilon|\log\varepsilon|$. We show that the paths of the SDE (1.5) with the same initial condition

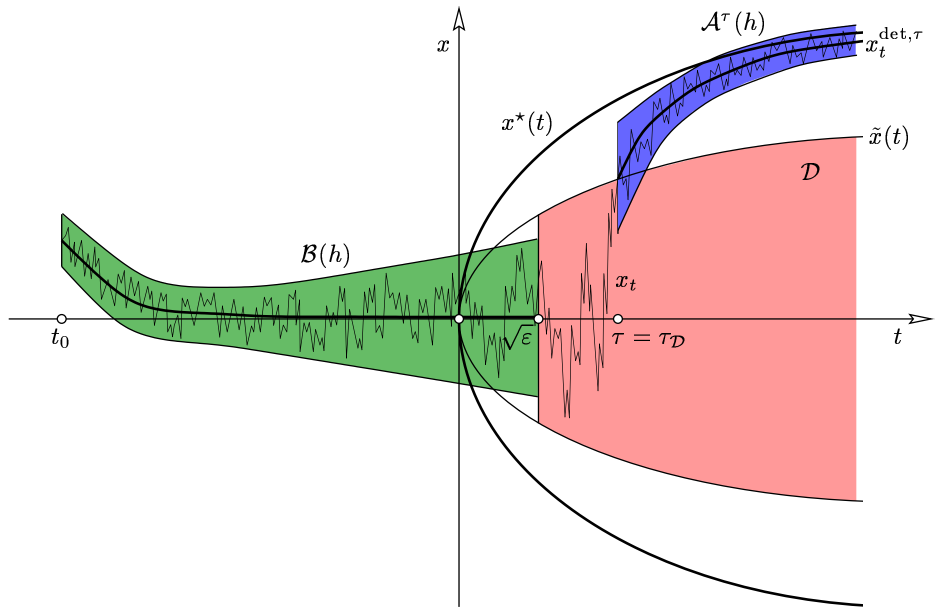

FIGURE 2. A typical path x_t of the stochastic differential equation (1.5) near a pitchfork bifurcation. We prove that with probability exponentially close to 1, the path has the following behaviour. For $t_0 \leqslant t \leqslant \sqrt{\varepsilon}$, it stays in a strip $\mathcal{B}(h)$ constructed around the deterministic solution with the same initial condition. After $t = \sqrt{\varepsilon}$, it leaves the domain $\mathcal{D}$ at a random time $\tau = \tau_{\mathcal{D}}$, which is typically of the order $\sqrt{\varepsilon|\log\sigma|}$. Then it stays (up to times of order 1 at least) in a strip $\mathcal{A}^\tau(h)$ constructed around the deterministic solution $x_t^{\det,\tau}$ starting at time τ on the boundary of $\mathcal{D}$. The widths of $\mathcal{B}(h)$ and $\mathcal{A}^\tau(h)$ are proportional to a parameter h satisfying $\sigma \ll h \ll \sqrt{\varepsilon}$.

are typically concentrated in a neighbourhood of order σ of the deterministic solution (Theorem 2.4).

- A particular solution of the deterministic equation (1.2) is known to exist in a neighbourhood of order ε of each unstable equilibrium branch of f. Paths that start in a neighbourhood of order σ of this solution are likely to leave that neighbourhood in a time of order $\varepsilon|\log\varepsilon|$ (Theorem 2.6).
- When a pitchfork bifurcation occurs at $x = 0$, $t = 0$, the typical paths are concentrated in a neighbourhood of order $\sigma/\varepsilon^{1/4}$ of the deterministic solution with the same initial condition up to time $\sqrt{\varepsilon}$ (Theorem 2.10).
- After the bifurcation point, the paths are likely to leave a neighbourhood of order $\sqrt{t}$ of the unstable equilibrium before a time of order $\sqrt{\varepsilon|\log\sigma|}$ (Theorem 2.11).
- Once they have left this neighbourhood, the paths remain with high probability in a region of size $\sigma/\sqrt{t}$ around the corresponding deterministic solution, which approaches a stable equilibrium branch of f like $\varepsilon/t^{3/2}$ (Theorem 2.12).

These results show that the bifurcation delay, which is observed in the dynamical system (1.2), is destroyed by additive noise as soon as the noise is not exponentially small. Do they mean that the dynamic bifurcation itself is destroyed by additive noise? This is mainly a matter of definition. On the one hand, we will see that independently of the initial condition, the probability of reaching the upper, rather than the lower branch emerging from the bifurcation point, is close to $\frac{1}{2}$. The asymptotic state is thus selected by the noise, and not by the initial condition. Hence, the bifurcation is destroyed in the sense of [CF98]. On the other hand, individual paths are concentrated near the stable equilibrium branches of f, which means that the bifurcation diagram will be made visible by the noise, much more so than in the deterministic case. So we do observe a qualitative change in behaviour when λ changes its sign, which can be considered as a bifurcation.

The precise statements and a discussion of their consequences are given in Section 2. In Section 2.2, we analyse the motion near equilibrium branches away from bifurcation points. The actual pitchfork bifurcation is discussed in Section 2.3. A few consequences are derived in Section 2.4. Section 3 contains the proofs of the first two theorems on the motion near nonbifurcating equilibria, while the proofs of the last three theorems on the pitchfork bifurcation are given in Section 4.

Acknowledgements:

It's a great pleasure to thank Anton Bovier for sharing our enthusiasm. We enjoyed lively discussions and his constant interest in the progress of our work. The central ideas were developed during mutual visits in Berlin resp. Atlanta. N. B. thanks the WIAS and B. G. thanks Turgay Uzer and the School of Physics at Georgia Tech for their kind hospitality. N. B. was partially supported by the Fonds National Suisse de la Recherche Scientifique, and by the Nonlinear Control Network of the European Community, Grant ERB FMRXCT–970137.

2 Statement of results

2.1 Preliminaries

We consider nonlinear SDEs of the form

$$dx_t = \frac{1}{\varepsilon} f(x_t, t)\, dt + \frac{\sigma}{\sqrt{\varepsilon}}\, dW_t, \qquad x_{t_0} = x_0, \tag{2.1}$$

where $\{W_t\}_{t \geqslant t_0}$ is a standard Wiener process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Initial conditions x_0 are always assumed to be square-integrable with respect to $\mathbb{P}$ and independent of $\{W_t\}_{t \geqslant t_0}$. All stochastic integrals are considered as Itô integrals, but note that Itô and Stratonovich integrals agree for integrands depending only on time and ω. Without further mentioning we always assume that f satisfies the usual (local) Lipschitz and bounded-growth conditions which guarantee existence and (pathwise) uniqueness of a (strong) solution $\{x_t\}_t$ of (2.1). Under these conditions, there exists a continuous version of $\{x_t\}_t$. Therefore we may assume that the paths $\omega \mapsto x_t(\omega)$ are continuous for $\mathbb{P}$-almost all $\omega \in \Omega$.

We introduce the notation $\mathbb{P}^{t_0,x_0}$ for the law of the process $\{x_t\}_{t \geqslant t_0}$, starting in x_0 at time t_0, and use $\mathbb{E}^{t_0,x_0}$ to denote expectations with respect to $\mathbb{P}^{t_0,x_0}$. Note that the stochastic process $\{x_t\}_{t \geqslant t_0}$ is an (inhomogeneous) Markov process. We are interested in first exit times of x_t from space–time sets. Let $\mathcal{A} \subset \mathbb{R} \times [t_0, t_1]$ be Borel-measurable. Assuming that $\mathcal{A}$ contains (x_0, t_0), we define the first exit time of (x_t, t) from $\mathcal{A}$ by

$$\tau_{\mathcal{A}} = \inf\{t \in [t_0, t_1] \colon (x_t, t) \notin \mathcal{A}\}, \tag{2.2}$$

and agree to set $\tau_{\mathcal{A}}(\omega) = \infty$ for those $\omega \in \Omega$ which satisfy $(x_t(\omega), t) \in \mathcal{A}$ for all $t \in [t_0, t_1]$. For convenience, we shall call $\tau_{\mathcal{A}}$ the *first exit time of x_t from $\mathcal{A}$*. Typically, we will consider sets of the form $\mathcal{A} = \{(x, t) \in \mathbb{R} \times [t_0, t_1] \colon g_1(t) < x < g_2(t)\}$ with continuous functions $g_1 < g_2$. Note that in this case, $\tau_{\mathcal{A}}$ is a stopping time[2] with respect to the canonical filtration of $(\Omega, \mathcal{F}, \mathbb{P})$ generated by $\{x_t\}_{t \geqslant t_0}$.

Before turning to the precise statements of our results, let us introduce some notations. We shall use

- $\lceil y \rceil$ for $y \geqslant 0$ to denote the smallest integer which is greater than or equal to y, and
- $y \vee z$ and $y \wedge z$ to denote the maximum or minimum, respectively, of two real numbers y and z.
- By $g(u) = \mathcal{O}(u)$ we indicate that there exist $\delta > 0$ and $K > 0$ such that $g(u) \leqslant Ku$ for all $u \in [0, \delta]$, where δ and K of course do not depend on ε or σ. Similarly, $g(u) = \mathcal{o}(u)$ is to be understood as $\lim_{u \to 0} g(u)/u = 0$. From time to time, we write $g(u) = \mathcal{o}_T(1)$ as a shorthand for $\lim_{T \to 0} \sup_{0 \leq u \leq T} |g(u)| = 0$.

Finally, let us point out that most estimates hold for small enough ε only, and often only for $\mathbb{P}$-almost all $\omega \in \Omega$. We will stress these facts only when confusion might arise.

2.2 Nonbifurcating equilibria

We start by considering the nonlinear SDE (2.1) in the case of f admitting a nonbifurcating equilibrium branch. We will make the following assumptions.

Assumption 2.1. There exist an interval $I = [0, T]$ or $[0, \infty)$ and a constant $d > 0$ such that the following properties hold:

- there exists a function $x^\star : I \to \mathbb{R}$, called *equilibrium branch*, such that

$$f(x^\star(t), t) = 0 \qquad \forall t \in I; \tag{2.3}$$

- f is twice continuously differentiable with respect to x and t for $|x - x^\star(t)| \leqslant d$ and $t \in I$, with uniformly bounded derivatives. In particular, there exists a constant $M > 0$ such that $|\partial_{xx} f(x, t)| \leqslant 2M$ in that domain;
- the linearization of f at $x^\star(t)$, defined as

$$a(t) = \partial_x f(x^\star(t), t), \tag{2.4}$$

is bounded away from zero, that is, there exists a constant $a_0 > 0$ such that

$$|a(t)| \geqslant a_0 \quad \forall t \in I. \tag{2.5}$$

We need no additional assumption on σ in this section. However, the results are only of interest for $\sigma = \mathcal{o}_\varepsilon(1)$.

In the deterministic case $\sigma = 0$, the following result is known (see Fig. 1a):

Theorem 2.2 (Deterministic case [Ti, Gr])**.** *Consider the equation*

$$\varepsilon \frac{\mathrm{d}x_t}{\mathrm{d}t} = f(x_t, t). \tag{2.6}$$

There are constants $\varepsilon_0, c_0, c_1 > 0$, depending only on f, such that for $0 < \varepsilon \leqslant \varepsilon_0$,

- (2.6) *admits a particular solution $\widehat{x}_t^{\mathrm{det}}$ such that*

$$|\widehat{x}_t^{\mathrm{det}} - x^\star(t)| \leqslant c_1 \varepsilon \quad \forall t \in I; \tag{2.7}$$

[2] For a general Borel-measurable set $\mathcal{A}$, the first exit time $\tau_{\mathcal{A}}$ is still a stopping time with respect to the canonical filtration, completed by the null sets.

- *if* $|x_0 - x^\star(0)| \leqslant c_0$ *and* $a(t) \leqslant -a_0$ *for all* $t \in I$ *(that is, when* $x^\star(t)$ *is a stable equilibrium branch), then the solution* $x_t^{\det}$ *of* (2.6) *with initial condition* $x_0^{\det} = x_0$ *satisfies*

$$|x_t^{\det} - \widehat{x}_t^{\det}| \leqslant |x_0 - \widehat{x}_0^{\det}|\,\mathrm{e}^{-a_0 t/2\varepsilon} \qquad \forall t \in I. \tag{2.8}$$

Remark 2.3. The particular solution $\widehat{x}^{\det}$ is often called *slow solution* or *adiabatic solution* of Equation (2.6). It is not unique in general, as suggested by (2.8).

We return now to the SDE (2.1) with $\sigma > 0$. Let us first consider the *stable case*, that is, we assume that $a(t) \leqslant -a_0 < 0$ for all $t \in I$. We assume that at $t = 0$, x_t starts at some (deterministic) x_0 sufficiently close to $x^\star(0)$. Theorem 2.2 shows that the deterministic solution $x_t^{\det}$ with the same initial condition $x_0^{\det} = x_0$ reaches a neighbourhood of order ε of $x^\star(t)$ exponentially fast.

We are interested in the stochastic process $y_t = x_t - x_t^{\det}$, which describes the deviation due to noise from the deterministic solution $x_t^{\det}$. It obeys the SDE

$$\mathrm{d}y_t = \frac{1}{\varepsilon}\bigl[f(x_t^{\det} + y_t, t) - f(x_t^{\det}, t)\bigr]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad y_0 = 0. \tag{2.9}$$

We will prove that y_t remains in a neighbourhood of 0 with high probability. It is instructive to consider first the linearization of (2.9) around $y = 0$, which has the form

$$\mathrm{d}y_t^0 = \frac{1}{\varepsilon}\bar{a}(t) y_t^0\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \tag{2.10}$$

where

$$\bar{a}(t) = \partial_x f(x_t^{\det}, t) = a(t) + \mathcal{O}(\varepsilon) + \mathcal{O}\bigl(|x_0 - x^\star(0)|\,\mathrm{e}^{-a_0 t/2\varepsilon}\bigr). \tag{2.11}$$

Taking ε and $|x_0 - x^\star(0)|$ sufficiently small, we may assume that $\bar{a}(t)$ is negative and bounded away from zero. The solution of (2.10) with arbitrary initial condition y_0^0 is given by

$$y_t^0 = y_0^0\,\mathrm{e}^{\overline{\alpha}(t)/\varepsilon} + \frac{\sigma}{\sqrt{\varepsilon}}\int_0^t \mathrm{e}^{\overline{\alpha}(t,s)/\varepsilon}\,\mathrm{d}W_s, \qquad \overline{\alpha}(t,s) = \int_s^t \bar{a}(u)\,\mathrm{d}u, \tag{2.12}$$

where we write $\overline{\alpha}(t,0) = \overline{\alpha}(t)$ for brevity. Note that $\overline{\alpha}(t,s) \leqslant -\mathit{const}(t-s)$ whenever $t \geqslant s$. If y_0^0 has variance $v_0 \geqslant 0$, then y_t^0 has variance

$$v(t) = v_0\,\mathrm{e}^{2\overline{\alpha}(t)/\varepsilon} + \frac{\sigma^2}{\varepsilon}\int_0^t \mathrm{e}^{2\overline{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s. \tag{2.13}$$

Since the first term decreases exponentially fast, the initial variance v_0 is "forgotten" as soon as $\mathrm{e}^{2\overline{\alpha}(t)/\varepsilon}$ is small enough, which happens already for $t > \mathcal{O}(\varepsilon|\log\varepsilon|)$. For $y_0^0 = 0$, (2.12) implies in particular that for any $\delta > 0$,

$$\mathbb{P}^{0,0}\bigl\{|y_t^0| \geqslant \delta\bigr\} \leqslant \mathrm{e}^{-\delta^2/2v(t)}, \tag{2.14}$$

and thus the probability of finding y_t^0, at any given $t \in I$, outside a strip of width much larger than $\sqrt{2v(t)}$ is very small.

Our first main result states that the *whole path* $\{x_s\}_{0\leqslant s\leqslant t}$ of the solution of the *non-linear* equation (2.1) lies with high probability in a similar strip, centred around $x_t^{\det}$. We only need to make one concession: the width of the strip has to be bounded away from zero. Therefore, we define the strip as

$$\mathcal{B}_{\mathrm{s}}(h) = \bigl\{(x,t) \in \mathbb{R} \times I \colon |x - x_t^{\det}| < h\sqrt{\zeta(t)}\bigr\}, \tag{2.15}$$

where

$$\zeta(t) = \frac{1}{2|\bar{a}(0)|}\,\mathrm{e}^{2\bar{\alpha}(t)/\varepsilon} + \frac{1}{\varepsilon}\int_0^t \mathrm{e}^{2\bar{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s. \tag{2.16}$$

Here $\sigma^2\zeta(t)$ can be interpreted as the variance (2.13) at time t of the process (2.12) starting with initial variance $v_0 = \sigma^2/(2|\bar{a}(0)|)$. We shall show in Lemma 3.1 that

$$\zeta(t) = \frac{1}{2|a(t)|} + \mathcal{O}(\varepsilon) + \mathcal{O}\big(|x_0 - x^\star(0)|\,\mathrm{e}^{-a_0 t/2\varepsilon}\big). \tag{2.17}$$

Let $\tau_{\mathcal{B}_\mathrm{s}(h)}$ denote the first exit time of x_t from $\mathcal{B}_\mathrm{s}(h)$.

Theorem 2.4 (Stable case)**.** *There exist ε_0, d_0 and h_0, depending only on f, such that for $0 < \varepsilon \leqslant \varepsilon_0$, $h \leqslant h_0$ and $|x_0 - x^\star(0)| \leqslant d_0$,*

$$\mathbb{P}^{0,x_0}\big\{\tau_{\mathcal{B}_\mathrm{s}(h)} < t\big\} \leqslant C(t,\varepsilon)\exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\big[1 - \mathcal{O}(\varepsilon) - \mathcal{O}(h)\big]\Big\}, \tag{2.18}$$

where

$$C(t,\varepsilon) = \frac{|\bar{\alpha}(t)|}{\varepsilon^2} + 2, \tag{2.19}$$

and $\mathcal{O}(\varepsilon)$ and $\mathcal{O}(h)$ do not depend on t.

The proof, given in Section 3.1, is divided into two main steps. First, we show that an estimate of the form (2.18), but without the term $\mathcal{O}(h)$, holds for the solution of the linearized equation (2.10). Then we show that whenever $|y_s^0| < h\sqrt{\zeta(s)}$ for $0 \leqslant s \leqslant t$, the bound $|y_s| < h(1 + \mathcal{O}(h))\sqrt{\zeta(s)}$ almost surely holds for $0 \leqslant s \leqslant t$.

Remark 2.5. The result of the preceding theorem remains true when $1/2|\bar{a}(0)|$ in the definition (2.16) of $\zeta(t)$ is replaced be an arbitrary ζ_0, *provided* $\zeta_0 > 0$. The terms $\mathcal{O}(\cdot)$ may then depend on ζ_0. Note that $\zeta(t)$ and $v(t)/\sigma^2$ are both solutions of the same differential equation $\varepsilon z' = 2\bar{a}(t)z + 1$, with possibly different initial conditions. If $x_0 - x^\star(0) = \mathcal{O}(\varepsilon)$, $\zeta(t)$ is an adiabatic solution (in the sense of Theorem 2.2) of the differential equation, staying close to the equilibrium branch $z^\star = 1/|2\bar{a}(t)|$.

The estimate (2.18) has been designed for situations where $\sigma \ll 1$, and is useful for $\sigma \ll h \ll 1$. We expect the exponent to be optimal in this case, but did not attempt to optimize the prefactor $C(t,\varepsilon)$, which leads to subexponential corrections. If we assume, for instance, that $\sigma = \varepsilon^q$, $q > 0$, and take $h = \varepsilon^p$ with $0 < p < q$, (2.18) can be written as

$$\mathbb{P}^{0,x_0}\big\{\tau_{\mathcal{B}_\mathrm{s}(h)} < t\big\} \leqslant (t+\varepsilon^2)\exp\Big\{-\frac{1}{2\varepsilon^{2(q-p)}}\big[1 - \mathcal{O}(\varepsilon) - \mathcal{O}(\varepsilon^p) - \mathcal{O}(\varepsilon^{2(q-p)}|\log\varepsilon|)\big]\Big\}. \tag{2.20}$$

The t-dependence of the prefactor is to be expected. It is due to the fact that as time increases, the probability of x_t escaping from a neighbourhood of $x_t^{\det}$ also increases, but very slowly if σ is small. The estimate (2.18) shows that for a fraction γ of trajectories to leave the strip $\mathcal{B}_\mathrm{s}(h)$, we have to wait *at least* for a time t_γ given by

$$|\bar{\alpha}(t_\gamma)| = \gamma\varepsilon^2\exp\Big\{\frac{1}{2}\frac{h^2}{\sigma^2}\big[1 - \mathcal{O}(\varepsilon) - \mathcal{O}(h)\big]\Big\} - 2\varepsilon^2, \tag{2.21}$$

which is compatible with results on the autonomous case.

Let us now consider the *unstable case*, that is, we now assume that the linearization $a(t) = \partial_x f(x^\star(t),t)$ satisfies $a(t) \geqslant a_0 > 0$ for all $t \in I$. Theorem 2.2 shows the existence of a particular solution $\widehat{x}_t^{\det}$ of the deterministic equation (2.6) such that $|\widehat{x}_t^{\det} - x^\star(t)| \leqslant c_1\varepsilon$ for all $t \in I$. We define $\bar{a}(t) = \partial_x f(\widehat{x}_t^{\det}, t) = a(t) + \mathcal{O}(\varepsilon) > 0$ and $\bar{\alpha}(t) = \int_0^t \bar{a}(s)\,\mathrm{d}s$.

The linearization of (2.1) around $\widehat{x}_t^{\det}$ again admits a solution y_t^0 of the form (2.12). Unlike in the stable case, the variance (2.13) grows exponentially fast (at least with $\mathrm{e}^{a_0 t/\varepsilon}$), and thus one expects the probability of x_t remaining close to $\widehat{x}_t^{\det}$ to be small. Indeed, if $\rho \geqslant |y_0^0|$, then we have

$$\begin{aligned}\mathbb{P}^{0,y_0^0}\Bigl\{\sup_{0\leqslant s\leqslant t} |y_s^0| < \rho\Bigr\} &\leqslant \mathbb{P}^{0,y_0^0}\bigl\{|y_t^0| < \rho\bigr\} \\ &= \int_{-\rho - y_0^0\,\mathrm{e}^{\bar{\alpha}(t)/\varepsilon}}^{\rho - y_0^0\,\mathrm{e}^{\bar{\alpha}(t)/\varepsilon}} \frac{\mathrm{e}^{-x^2/2v(t)}}{\sqrt{2\pi v(t)}}\,\mathrm{d}x \leqslant \frac{2\rho}{\sqrt{2\pi v(t)}},\end{aligned} \tag{2.22}$$

which goes to zero as $\rho\sigma^{-1}\,\mathrm{e}^{-\bar{\alpha}(t)/\varepsilon}$ for $t \to \infty$. In this estimate, however, we neglect all trajectories which leave the interval $(-\rho, \rho)$ before time t and come back. We will derive a more precise estimate for the general, nonlinear case. This is the contents of the second main result of this section. Let

$$\mathcal{B}_\mathrm{u}(h) = \biggl\{(x,t) \in \mathbb{R} \times I \colon |x - \widehat{x}_t^{\det}| < \frac{h}{\sqrt{2\bar{a}(t)}}\biggr\} \tag{2.23}$$

and denote by $\tau_{\mathcal{B}_\mathrm{u}(h)}$ the first exit time of x_t from $\mathcal{B}_\mathrm{u}(h)$.

Theorem 2.6 (Unstable case)**.** *There exist ε_0 and h_0, depending only on f, such that for all $h \leqslant \sigma \wedge h_0$, all $\varepsilon \leqslant \varepsilon_0$ and all x_0 satisfying $(x_0, 0) \in \mathcal{B}_\mathrm{u}(h)$, we have*

$$\mathbb{P}^{0,x_0}\bigl\{\tau_{\mathcal{B}_\mathrm{u}(h)} \geqslant t\bigr\} \leqslant \sqrt{\mathrm{e}}\exp\Bigl\{-\kappa\frac{\sigma^2}{h^2}\frac{\bar{\alpha}(t)}{\varepsilon}\Bigr\}, \tag{2.24}$$

where $\kappa = \frac{\pi}{2\mathrm{e}}\bigl(1 - \mathcal{O}(h) - \mathcal{O}(\varepsilon)\bigr)$ does not depend on t.

The proof, given in Section 3.2, is based on a partition of the interval $[0,t]$ into small intervals, and a comparison of the nonlinear equation with its linearization on each interval.

The above result shows that x_t is unlikely to remain in $\mathcal{B}_\mathrm{u}(h)$ as soon as $t \gg \varepsilon h^2/\sigma^2$. A major limitation of (2.24) is that it requires $h \leqslant \sigma$. Obtaining an estimate for larger h is possible, but requires considerably more work. We will provide such an estimate in the more difficult, but also more interesting case of the pitchfork bifurcation, see Theorem 2.11 below.

2.3 Pitchfork bifurcation

We now consider the SDE (2.1) in the case of f undergoing a pitchfork bifurcation. We will assume the following.

Assumption 2.7. There exists a neighbourhood $\mathcal{N}_0$ of the origin $(0,0)$ such that

- f is three times continuously differentiable with respect to x and t in $\mathcal{N}_0$, and there exists a constant $M > 0$ such that $|\partial_{xxx} f(x,t)| \leqslant 6M$ for all $(x,t) \in \mathcal{N}_0$;
- $f(x,t) = -f(-x,t)$ for all $(x,t) \in \mathcal{N}_0$;
- f exhibits a supercritical pitchfork bifurcation at the origin, i.e.,

$$\partial_x f(0,0) = 0, \qquad \partial_{tx} f(0,0) > 0 \qquad \text{and} \qquad \partial_{xxx} f(0,0) < 0. \tag{2.25}$$

By rescaling x and t, we may and will assume that

$$\partial_x f(0,0) = 0, \qquad \partial_{tx} f(0,0) = 1 \qquad \text{and} \qquad \partial_{xxx} f(0,0) = -6 \tag{2.26}$$

as in the standard case $f(x,t) = tx - x^3$.

Note that the assumption that f be odd is not necessary for the existence of a pitchfork bifurcation. However, the deterministic system behaves very differently if $x = 0$ is not always an equilibrium. The most natural situation in which $f(0,t) = 0$ for all t is the one where f is odd. The proofs below can easily be extended to the case where f is not necessarily odd—provided f exhibits a supercritical pitchfork bifurcation with $x(t) \equiv 0$ being the equilibrium branch which becomes unstable at the bifurcation point.

Using (2.26), we find that the linearization of f at $x = 0$ satisfies

$$a(t) = \partial_x f(0,t) = t + \mathcal{O}(t^2). \tag{2.27}$$

A standard result of bifurcation theory [GH, IJ] states that under these assumptions, there exists a neighbourhood $\mathcal{N} \subset \mathcal{N}_0$ of $(0,0)$ in which the only solutions of $f(x,t) = 0$ are the line $x = 0$ and the curves

$$x = \pm x^\star(t), \qquad x^\star(t) = \sqrt{t}\bigl[1 + \mathcal{O}_t(1)\bigr], \qquad t \geqslant 0. \tag{2.28}$$

If $\mathcal{N}$ is small enough, the equilibrium $x = 0$ is stable for $t < 0$ and unstable for $t > 0$, while $x = \pm x^\star(t)$ are stable equilibria with linearization

$$a^\star(t) = \partial_x f(x^\star(t), t) = -2t\bigl[1 + \mathcal{O}_t(1)\bigr]. \tag{2.29}$$

The only solutions of $\partial_x f(x,t) = 0$ in $\mathcal{N}$ are the curves

$$x = \pm\bar{x}(t), \qquad \bar{x}(t) = \sqrt{t/3}\bigl[1 + \mathcal{O}_t(1)\bigr], \qquad t \geqslant 0. \tag{2.30}$$

If f is four times continuously differentiable, the terms $\mathcal{O}_t(1)$ in the last three equations can be replaced by $\mathcal{O}(t)$.

We briefly state what is known for the deterministic equation

$$\varepsilon \frac{\mathrm{d}x_t}{\mathrm{d}t} = f(x_t, t), \tag{2.31}$$

where we take an initial condition $(x_0, t_0) \in \mathcal{N}$ with $x_0 > 0$ and $t_0 < 0$, see Fig. 1b. Observe that $\alpha(t, t_0) = \int_{t_0}^t a(s)\,\mathrm{d}s$ is decreasing for $t_0 < t < 0$ and increasing for $t > 0$.

Definition 2.8. *The* bifurcation delay *is defined as*

$$\Pi(t_0) = \inf\bigl\{t > 0 \colon \alpha(t, t_0) > 0\bigr\}, \tag{2.32}$$

with the convention $\Pi(t_0) = \infty$ *if* $\alpha(t, t_0) < 0$ *for all* $t > 0$*, for which* $\alpha(t, t_0)$ *is defined.*

One easily shows that $\Pi(t_0)$ is differentiable for t_0 sufficiently close to 0, and satisfies $\lim_{t_0 \to 0-} \Pi(t_0) = 0$ and $\lim_{t_0 \to 0-} \Pi'(t_0) = -1$.

Theorem 2.9 (Deterministic case)**.** *Let* $x_t^{\det}$ *be the solution of* (2.31) *with initial condition* $x_{t_0}^{\det} = x_0$*. Then there exist constants* ε_0*,* c_0*,* c_1 *depending only on* f*, and times*

$$\begin{aligned} t_1 &= t_0 + \mathcal{O}(\varepsilon|\log\varepsilon|) \\ t_2 &= \Pi(t_1) = \Pi(t_0) - \mathcal{O}(\varepsilon|\log\varepsilon|) \\ t_3 &= \Pi(t_0) + \mathcal{O}(\varepsilon|\log\varepsilon|) \end{aligned} \tag{2.33}$$

such that, if $0 < x_0 \leqslant c_0$, $0 < \varepsilon \leqslant \varepsilon_0$ *and* $(x_t^{\mathrm{det}}, t) \in \mathcal{N}$,

$$\begin{cases} 0 < x_t^{\mathrm{det}} \leqslant c_1 \varepsilon\, \mathrm{e}^{\alpha(t,t_1)/\varepsilon} & \text{for } t_1 \leqslant t \leqslant t_2 \\ |x_t^{\mathrm{det}} - x^\star(t)| \leqslant c_1 \varepsilon & \text{for } t \geqslant t_3. \end{cases} \tag{2.34}$$

The proof is a straightforward consequence of differential inequalities, see for instance [Ber, Propositions 4.6 and 4.8].

We now consider the SDE (2.1) for $\sigma > 0$. The results in this section are only interesting for $\sigma = o(\sqrt{\varepsilon})$, while one of them (Theorem 2.11) *requires* a condition of the form $\sigma|\log\sigma|^{3/2} = \mathcal{O}(\sqrt{\varepsilon})$ (where we have not tried to optimize the exponent 3/2).

Let us fix an initial condition $(x_{t_0}, t_0) \in \mathcal{N}$ with $t_0 < 0$. For any $T \in (0, |t_0|)$, we can apply Theorem 2.4 on the interval $[t_0, -T]$ to show that $|x_{-T}|$ is likely to be at most of order $\sigma^{1-\delta} + c_1\varepsilon\, \mathrm{e}^{\alpha(-T,t_1)/\varepsilon}$ for any $\delta > 0$. We can also apply the theorem for $t > T$ to show that the curves $\pm x^\star(t)$ attract nearby trajectories. Hence there is no limitation in considering the SDE (2.1) in a domain of the form $|x| \leqslant d$, $|t| \leqslant T$ where d and T can be taken small (independently of ε and σ of course!), with an initial condition $x_{-T} = x_0$ satisfying $|x_0| \leqslant d$. From now on, we will always assume that

$$\mathcal{N} = \{(x,t) \in \mathbb{R}^2 \colon |x| \leqslant d,\ |t| \leqslant T\} \tag{2.35}$$

for some $d, T > 0$.

We first show that x_t is likely to remain small for $-T \leqslant t \leqslant \sqrt{\varepsilon}$. Actually, it turns out to be convenient to show that x_t remains close to the solution $x_0\, \mathrm{e}^{\alpha(t,-T)/\varepsilon}$ of the linearization of (2.31). We define the "variance-like" function

$$\zeta(t) = \frac{1}{2|a(-T)|}\, \mathrm{e}^{2\alpha(t,-T)/\varepsilon} + \frac{1}{\varepsilon} \int_{-T}^{t} \mathrm{e}^{2\alpha(t,s)/\varepsilon}\, \mathrm{d}s. \tag{2.36}$$

We shall show in Lemma 4.2 that for sufficiently small ε, there exist constants $c_\pm$ such that

$$\frac{c_-}{|t|} \leqslant \zeta(t) \leqslant \frac{c_+}{|t|} \qquad \text{for } -T \leqslant t \leqslant -\sqrt{\varepsilon}, \tag{2.37}$$

$$\frac{c_-}{\sqrt{\varepsilon}} \leqslant \zeta(t) \leqslant \frac{c_+}{\sqrt{\varepsilon}} \qquad \text{for } -\sqrt{\varepsilon} \leqslant t \leqslant \sqrt{\varepsilon}. \tag{2.38}$$

The function $\zeta(t)$ is used to define the strip

$$\mathcal{B}(h) = \bigl\{(x,t) \in [-d,d] \times [-T, \sqrt{\varepsilon}\,] \colon |x - x_0\, \mathrm{e}^{\alpha(t,-T)/\varepsilon}| < h\sqrt{\zeta(t)}\bigr\}. \tag{2.39}$$

Let $\tau_{\mathcal{B}(h)}$ denote the first exit time of x_t from $\mathcal{B}(h)$.

Theorem 2.10 (Behaviour for $t \leqslant \sqrt{\varepsilon}$)**.** *There exist constants* ε_0 *and* h_0*, depending only on* f*,* T *and* d*, such that for* $0 < \varepsilon \leqslant \varepsilon_0$*,* $h \leqslant h_0\sqrt{\varepsilon}$*,* $|x_0| \leqslant h/\varepsilon^{1/4}$ *and* $-T \leqslant t \leqslant \sqrt{\varepsilon}$*,*

$$\mathbb{P}^{-T,x_0}\bigl\{\tau_{\mathcal{B}(h)} < t\bigr\} \leqslant C(t,\varepsilon) \exp\biggl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\biggl[1 - r(\varepsilon) - \mathcal{O}\biggl(\frac{h^2}{\varepsilon}\biggr)\biggr]\biggr\} \tag{2.40}$$

where

$$C(t,\varepsilon) = \frac{|\alpha(t,-T)| + \mathcal{O}(\varepsilon)}{\varepsilon^2} \tag{2.41}$$

and $r(\varepsilon) = \mathcal{O}(\varepsilon)$ *for* $-T \leqslant t \leqslant -\sqrt{\varepsilon}$*, and* $r(\varepsilon) = \mathcal{O}(\sqrt{\varepsilon})$ *for* $-\sqrt{\varepsilon} \leqslant t \leqslant \sqrt{\varepsilon}$*.*

The proof (given in Section 4.2) and the interpretation of this result are very close in spirit to those of Theorem 2.4. The only difference lies in the kind of ε-dependence of the error terms. The estimate (2.40) is useful when $\sigma \ll h \ll \sqrt{\varepsilon}$, and shows that the typical spreading of paths around the deterministic solution will slowly grow until $t = \sqrt{\varepsilon}$, where it is of order $\sigma/\varepsilon^{1/4}$, see Fig. 2.

Let us now examine what happens for $t \geqslant \sqrt{\varepsilon}$. We first show that x_t is likely to leave quite soon a suitably defined region $\mathcal{D}$ containing the line $x = 0$. We define $\mathcal{D} = \mathcal{D}(\kappa)$ by

$$\mathcal{D}(\kappa) = \Bigl\{(x,t) \in [-d,d] \times [\sqrt{\varepsilon}, T] \colon \frac{1}{x} f(x,t) \geqslant \kappa a(t)\Bigr\}, \tag{2.42}$$

where $\kappa \in (0,1)$ is a free parameter. The upper and lower boundary of $\mathcal{D}(\kappa)$ are given by $\pm\tilde{x}(t)$, where $\tilde{x}(t)$ satisfies

$$\tilde{x}(t) = \sqrt{1 - \kappa - \mathcal{O}_T(1)}\,\sqrt{a(t)}. \tag{2.43}$$

Later, we will assume $\kappa \in (1/2, 2/3)$. This will simplify the analysis of the dynamics after x_t has left $\mathcal{D}(\kappa)$. The upper bound on κ implies $\bar{x}(t) \leqslant \tilde{x}(t) \leqslant x^\star(t)$, while the lower bound guarantees that the solution of the corresponding deterministic equation does not return to $\mathcal{D}(\kappa)$ once it has left this set, cf. Proposition 4.11 below.

Let $\tau_{\mathcal{D}(\kappa)}$ denote the first exit time of x_t from $\mathcal{D}(\kappa)$. Then we have the following result.

Theorem 2.11 (Escape from $\mathcal{D}(\kappa)$)**.** *Choose $\kappa \in (0, 2/3)$ and let $(x_0, t_0) \in \mathcal{D}(\kappa)$. Assume that $\sigma|\log\sigma|^{3/2} = \mathcal{O}(\sqrt{\varepsilon})$. Then for $t_0 \leqslant t \leqslant T$,*

$$\mathbb{P}^{t_0,x_0}\bigl\{\tau_{\mathcal{D}(\kappa)} \geqslant t\bigr\} \leqslant C_0\,\tilde{x}(t)\sqrt{a(t)}\,\frac{|\log\sigma|}{\sigma}\biggl(1 + \frac{\alpha(t,t_0)}{\varepsilon}\biggr)\frac{\mathrm{e}^{-\kappa\alpha(t,t_0)/\varepsilon}}{\sqrt{1 - \mathrm{e}^{-2\kappa\alpha(t,t_0)/\varepsilon}}}, \tag{2.44}$$

where $C_0 > 0$ is a (numerical) constant.

The proof of this result (given in Section 4.3) is by far the most involved of the present work. We start by estimating, in a similar way as in Theorem 2.6, the first exit time from a strip $\mathcal{S}$ of width slightly larger than $\sigma/\sqrt{a(s)}$. The probability of returning to zero after leaving $\mathcal{S}$ can be estimated; it is small but not exponentially small. However, the probability of neither leaving $\mathcal{D}(\kappa)$ nor returning to zero is exponentially small. This fact can be used to devise an iterative scheme that leads to the exponential estimate (2.44).

We point out that for any subset $\mathcal{D}' \subset \mathcal{D}(\kappa)$, we have the trivial estimate $\mathbb{P}^{t_0,x_0}\{\tau_{\mathcal{D}'} \geqslant t\} \leqslant \mathbb{P}^{t_0,x_0}\{\tau_{\mathcal{D}(\kappa)} \geqslant t\}$, and thus (2.44) also provides an upper bound for the first exit time from smaller sets.

Let us finally consider what happens after the path has left $\mathcal{D}(\kappa)$ at time $\tau = \tau_{\mathcal{D}(\kappa)}$ (with $\kappa \in (1/2, 2/3)$). First note that (2.43) immediately implies that for $\sqrt{\varepsilon} \leqslant t \leqslant T$ and $|x| \geqslant \tilde{x}(t)$,

$$\partial_x f(x,t) \leqslant \tilde{a}(t) = \partial_x f(\tilde{x}(t),t) \leqslant -\eta a(t) \qquad \text{provided } \eta \leqslant 2 - 3\kappa - \mathcal{O}_T(1). \tag{2.45}$$

Let $x_t^{\det,\tau}$ denote the solution of the deterministic equation (2.31) starting in $\tilde{x}(t)$ at time τ (the case where one starts at $-\tilde{x}(t)$ is obtained by symmetry). We shall show in Proposition 4.11 that $x_t^{\det,\tau}$ always remains between $\tilde{x}(t)$ and $x^\star(t)$, and approaches $x^\star(t)$ according to

$$x_t^{\det,\tau} = x^\star(t) - \mathcal{O}\Bigl(\frac{\varepsilon}{t^{3/2}}\Bigr) - \mathcal{O}\bigl(\sqrt{\tau}\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}\bigr). \tag{2.46}$$

Moreover, deterministic solutions starting at different times approach each other like

$$0 \leqslant x_t^{\det,\sqrt{\varepsilon}} - x_t^{\det,\tau} \leqslant \big(x_\tau^{\det,\sqrt{\varepsilon}} - \tilde{x}(\tau)\big)\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon} \qquad \forall t \in [\tau, T]. \tag{2.47}$$

The linearization of f at $x_t^{\det,\tau}$ satisfies

$$a^\tau(t) = \partial_x f(x_t^{\det,\tau}, t) = a^\star(t) + \mathcal{O}\Big(\frac{\varepsilon}{t}\Big) + \mathcal{O}\big(t\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}\big). \tag{2.48}$$

For given τ, we construct a strip $\mathcal{A}^\tau(h)$ around $x^{\det,\tau}$ of the form

$$\mathcal{A}^\tau(h) = \big\{(x,t)\colon \tau \leqslant t \leqslant T, |x - x_t^{\det,\tau}| < h\sqrt{\zeta^\tau(t)}\big\}, \tag{2.49}$$

where the function $\zeta^\tau(t)$ is defined by

$$\zeta^\tau(t) = \frac{1}{2|\tilde{a}(\tau)|}\,\mathrm{e}^{2\alpha^\tau(t,\tau)/\varepsilon} + \frac{1}{\varepsilon}\int_\tau^t \mathrm{e}^{2\alpha^\tau(t,s)/\varepsilon}\,\mathrm{d}s, \qquad \alpha^\tau(t,s) = \int_s^t a^\tau(u)\,\mathrm{d}u, \tag{2.50}$$

and satisfies

$$\zeta^\tau(t) = \frac{1}{2|a^\star(t)|} + \mathcal{O}\Big(\frac{\varepsilon}{t^3}\Big) + \mathcal{O}\Big(\frac{1}{t}\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}\Big), \tag{2.51}$$

cf. Lemma 4.12. Let $\tau_{\mathcal{A}^\tau(h)}$ denote the first exit time of x_t from $\mathcal{A}^\tau(h)$.

Theorem 2.12 (Approach to $x^\star$)**.** *Let $\kappa \in (1/2, 2/3)$. Then there exist constants ε_0 and h_0, depending only on f, T and d, such that for $0 < \varepsilon \leqslant \varepsilon_0$, $h < h_0\tau$ and $\tau \leqslant t \leqslant T$,*

$$\mathbb{P}^{\tau,\tilde{x}(\tau)}\big\{\tau_{\mathcal{A}^\tau(h)} < t\big\} \leqslant C^\tau(t,\varepsilon)\exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\Big[1 - \mathcal{O}(\varepsilon) - \mathcal{O}\Big(\frac{h}{\tau}\Big)\Big]\Big\} \tag{2.52}$$

where

$$C^\tau(t,\varepsilon) = \frac{|\alpha^\tau(t,\tau)|}{\varepsilon^2} + 2 \leqslant \frac{1}{\varepsilon^2}\bigg|\int_{\sqrt{\varepsilon}}^t a^\star(s)\,\mathrm{d}s\bigg| + 2. \tag{2.53}$$

The proof is given in Section 4.4. This result is useful for $\sigma \ll h \ll \tau$, and shows that the typical spreading of paths around $x_t^{\det,\tau}$ is of order $\sigma/\sqrt{t}$, see Fig. 2.

2.4 Discussion

Let us now examine some of the consequences of these results. First of all, they allow to characterize the influence of additive noise on the bifurcation delay. In the deterministic case, this delay is defined as the first exit time from a strip of width ε around $x = 0$, see Theorem 2.9. A possible definition of the delay in the stochastic case is thus the first exit time τ^{delay} from a similar strip. An appropriate choice for the width of the strip is $\tilde{x}(\sqrt{\varepsilon}) = \mathcal{O}(\varepsilon^{1/4})$, since such a strip will contain $\mathcal{B}(h)$ for every admissible h, and the part of the strip with $t \geqslant \sqrt{\varepsilon}$ will be contained in $\mathcal{D}(\kappa)$. Theorems 2.10 and 2.11 then imply that if $t \geqslant \sqrt{\varepsilon}$,

$$\mathbb{P}^{-T,x_0}\big\{\tau^{\text{delay}} < \sqrt{\varepsilon}\big\} \leqslant C(\sqrt{\varepsilon},\varepsilon)\,\mathrm{e}^{-\mathcal{O}(\varepsilon/\sigma^2)} \tag{2.54}$$

$$\mathbb{P}^{-T,x_0}\big\{\tau^{\text{delay}} \geqslant t\big\} \leqslant C_0\,\tilde{x}(t)\sqrt{a(t)}\,\frac{|\log\sigma|}{\sigma}\bigg(1 + \frac{\alpha(t,\sqrt{\varepsilon})}{\varepsilon}\bigg)\frac{\mathrm{e}^{-\kappa\alpha(t,\sqrt{\varepsilon})/\varepsilon}}{\sqrt{1 - \mathrm{e}^{-2\kappa\alpha(t,\sqrt{\varepsilon})/\varepsilon}}}. \tag{2.55}$$

If we choose t in such a way that $\alpha(t,\sqrt{\varepsilon}) = c\varepsilon|\log\sigma|$ for some $c > 0$, the last expression reduces to

$$\mathbb{P}^{-T,x_0}\big\{\tau^{\text{delay}} \geqslant t\big\} = \mathcal{O}\big(\sigma^{\kappa c - 1}|\log\sigma|^2\big), \tag{2.56}$$

which becomes small as soon as $c > 1/\kappa$. The bifurcation delay will thus lie with overwhelming probability in the interval

$$\big[\sqrt{\varepsilon}, \mathcal{O}\big(\sqrt{\varepsilon|\log\sigma|}\,\big)\big]. \tag{2.57}$$

Theorem 2.12 implies that for times larger than $\mathcal{O}(\sqrt{\varepsilon|\log\sigma|}\,)$, the paths are unlikely to return to zero in a time of order 1. The wildest behaviour of the paths is to be expected in the interval (2.57), because a region of instability is crossed, where $\partial_x f > 0$.

Our results on the pitchfork bifurcation require $\sigma \ll \sqrt{\varepsilon}$, while the estimate (2.57) is useful as long as σ is not exponentially small. We can thus distinguish three regimes, depending on the noise intensity:

- $\sigma \geqslant \sqrt{\varepsilon}$: A modification of Theorem 2.10 shows that for $t < -\sigma$, the typical spreading of paths is of order $\sigma/\sqrt{|t|}$. Near the bifurcation point, the process is dominated by noise, because the drift term $f \sim -x^3$ is too weak to counteract the diffusion. Depending on the global structure of f, an appreciable fraction of the paths might escape quite early from a neighbourhood of the bifurcation point. In that situation, the notion of bifurcation delay becomes meaningless.
- $e^{-1/\varepsilon^p} \leqslant \sigma \ll \sqrt{\varepsilon}$ for some $p < 1$: The bifurcation delay lies in the interval (2.57) with high probability, where $\sqrt{\varepsilon|\log\sigma|} \leqslant \varepsilon^{(1-p)/2}$ is still "microscopic".
- $\sigma \leqslant e^{-K/\varepsilon}$ for some $K > 0$: The noise is so small that the paths remain concentrated around the deterministic solution for a time interval of order 1. The typical spreading is of order $\sigma\sqrt{\zeta(t)}$, which behaves like $\sigma\, e^{\alpha(t)/\varepsilon}/\varepsilon^{1/4}$ for $t \geqslant \sqrt{\varepsilon}$, see Lemma 4.2. Thus the paths remain close to the origin until $\alpha(t) \simeq \varepsilon|\log\sigma| \geqslant K$. If $\varepsilon|\log\sigma| > \alpha(\Pi(t_0)) = |\alpha(t_0)|$, they follow the deterministic solution which makes a quick transition to $x^\star(t)$ at $t = \Pi(t_0)$.

The expression (2.57) characterizing the delay is in accordance with experimental results [TM, SMC], and with the approximate calculation of the last crossing of zero [JL]. The numerical results in [Ga], which are fitted, at $\varepsilon = 0.01$, to $\tau^{\text{delay}} \simeq \sigma^{0.105}$ for weak noise and $\tau^{\text{delay}} \simeq e^{-851\,\sigma}$ for strong noise, seem rather mysterious. Finally, the results in [Ku], who approximates the probability density by a Gaussian centered at the deterministic solution, can obviously only apply to the regime of exponentially small noise.

Note that the estimate (2.57) suggests how to choose the speed ε at which the bifurcation parameter is swept when determining a bifurcation diagram experimentally: Since we want the bifurcation delay to be microscopic, ε should not exceed a certain value depending on the noise intensity σ. In fact, repeating the experiment for different values of ε yields an estimate for $|\log\sigma|$. On the other hand, increasing artificially the noise level σ allows to work with larger sweeping rates ε, reducing the time cost of the experiment.

Another interesting question is how fast the paths concentrate near the equilibrium branches $\pm x^\star(t)$. The deterministic solutions, starting at $\tilde{x}(t_0)$ at some time $t_0 > 0$, all track $x^\star(t)$ at a distance which is asymptotically of order $\varepsilon/t^{3/2}$. Therefore, we can choose one of them, say $x_t^{\det,\sqrt{\varepsilon}}$, and measure the distance of x_t from that deterministic solution. We restrict our attention to those paths which are still in a neighbourhood of the origin at time $\sqrt{\varepsilon}$, as most paths are. We want to show that for suitably chosen $t_1 \in (\sqrt{\varepsilon}, t)$ and $\Delta \in (0,t)$, most paths will leave $\mathcal{D}(\kappa)$ until time t_1 and reach a δ-neighbourhood of $x_t^{\det,\sqrt{\varepsilon}}$

at time $\tau_{\mathcal{D}(\kappa)} + \Delta$. Let us estimate

$$\begin{aligned}
&\mathbb{P}^{\sqrt{\varepsilon},x_{\sqrt{\varepsilon}}}\biggl\{\Bigl(\tau_{\mathcal{D}(\kappa)} < t_1, \sup_{s\in[\tau_{\mathcal{D}(\kappa)}+\Delta,t]}\Bigl||x_s| - x_s^{\det,\sqrt{\varepsilon}}\Bigr| < \delta\Bigr)^{\mathrm{c}}\biggr\} \\
&\quad\leqslant \mathbb{P}^{\sqrt{\varepsilon},x_{\sqrt{\varepsilon}}}\bigl\{\tau_{\mathcal{D}(\kappa)} \geqslant t_1\bigr\} \\
&\qquad + \mathbb{E}^{\sqrt{\varepsilon},x_{\sqrt{\varepsilon}}}\biggl\{1_{\{\tau_{\mathcal{D}(\kappa)}<t_1\}}\, \mathbb{P}^{\tau_{\mathcal{D}(\kappa)},\tilde{x}(\tau_{\mathcal{D}(\kappa)})}\biggl\{\sup_{s\in[\tau_{\mathcal{D}(\kappa)}+\Delta,t]}|x_s - x_s^{\det,\sqrt{\varepsilon}}| \geqslant \delta\biggr\}\biggr\}. \qquad (2.58)
\end{aligned}$$

The first term decreases roughly like $\sigma^{-1}\,\mathrm{e}^{-\kappa\alpha(t_1,\sqrt{\varepsilon})/\varepsilon}$ and becomes small as soon as $\alpha(t_1,\sqrt{\varepsilon}) \gg \varepsilon|\log\sigma|$. The second summand is bounded above by

$$const\; \mathbb{E}^{\sqrt{\varepsilon},x_{\sqrt{\varepsilon}}}\Bigl\{1_{\{\tau_{\mathcal{D}(\kappa)}<t_1\}}\, \exp\Bigl\{-\frac{t^2}{\sigma^2}\Bigl[\delta - \mathcal{O}\bigl(\sqrt{\tau_{\mathcal{D}(\kappa)}}\,\mathrm{e}^{-\eta\alpha(\tau_{\mathcal{D}(\kappa)}+\Delta,\tau_{\mathcal{D}(\kappa)})/\varepsilon}\bigr)^2\Bigr]\Bigr\}\Bigr\}. \qquad (2.59)$$

Therefore, δ should be large compared to σ/t and we also need that Δ is at least of order $\mathcal{O}(\sqrt{\varepsilon|\log\sigma|})$. Then we see that after a time of order $\mathcal{O}(\sqrt{\varepsilon|\log\sigma|})$, the typical paths will have left $\mathcal{D}(\kappa)$ and, after another time of the same order, will reach a neighbourhood of $x_t^{\det,\sqrt{\varepsilon}}$, which scales with σ/t.

Finally, we can also estimate the probability of reaching the positive rather than the negative branch. Consider x_s, starting in x_0 at time $t_0 < 0$, and let $t > 0$. Without loss of generality, we may assume that $x_0 > 0$. The symmetry of f implies

$$\mathbb{P}^{t_0,x_0}\bigl\{x_t \geqslant 0\bigr\} = 1 - \frac{1}{2}\mathbb{P}^{t_0,x_0}\bigl\{\exists\, s \in [t_0,t) : x_s = 0\bigr\}, \qquad (2.60)$$

and therefore it is sufficient to estimate the probability for x_s to reach zero before time zero, for instance. We linearize the SDE (2.1) and use the fact that the solution x_s^0 of the linearized equation

$$\mathrm{d}x_s^0 = \frac{1}{\varepsilon}a(s)x_s^0\,\mathrm{d}s + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_s, \qquad x_{t_0}^0 = x_0 \qquad (2.61)$$

satisfies $x_s \leqslant x_s^0$ as long as x_s does not reach zero. For the Gaussian process x_s^0 we know

$$\mathbb{P}^{t_0,x_0}\bigl\{\exists\, s \in [t_0,t) : x_s^0 = 0\bigr\} = 2\Bigl(1 - \mathbb{P}^{t_0,x_0}\bigl\{x_t^0 \geqslant 0\bigr\}\Bigr) = 1 - \frac{1}{\sqrt{2\pi}}\int_{-u(t)}^{u(t)} \mathrm{e}^{-y^2/2}\,\mathrm{d}y, \quad (2.62)$$

where $u(t) = x_0\,\mathrm{e}^{\alpha(t,t_0)/\varepsilon}/\sqrt{v(t,t_0)}$ and $v(t,t_0)$ denotes the variance of x_t^0. For $t = 0$, $u(0)$ is of order $x_0\varepsilon^{1/4}\sigma^{-1}\,\mathrm{e}^{-const\;t_0^2/\varepsilon}$, see Lemma 4.2. Thus the probability in (2.62) is exponentially close to one for small ε, and we conclude that the probability for x_t to reach the positive branch rather than the negative one is exponentially close to $1/2$.

When the *global* behaviour of f is known, we can also investigate the long-time behaviour of the solutions x_t. For instance, in the special case $f(x,t) = tx - x^3$, under the assumption $\sigma^2 \leqslant const/|\log\varepsilon|$, it is unlikely that a path which is close to one of the stable equilibrium branches $\pm\sqrt{t}$ at some time of order 1, will switch to the other equilibrium branch again. This is a consequence of the fact that the distance between the equilibrium branches increases while the branches become more and more attractive. Along the lines of Section 3.1 it can be shown that the probability of ever reaching zero again decays like $\mathrm{e}^{-const/\sigma^2}$ in that case.

3 The motion near nonbifurcating equilibria

In this section we consider the nonlinear SDE (2.1) under Assumption 2.1 which guarantees the existence of a hyperbolic equilibrium branch. Section 3.1 is devoted to the stable case, while in Section 3.2, we consider the unstable case.

3.1 Stable case

We first consider the case of a stable equilibrium, that is, we assume that $a(t) \leqslant -a_0$ for all $t \in I$. We will start by analysing the linearization of (2.1) around a given deterministic solution. Proposition 3.4 shows that the solutions of the linearized equation are likely to remain in a strip of width $h\sqrt{\zeta(t)}$ around the deterministic solution. Here $\zeta(t)$ is related to the variance and will be analysed in Lemma 3.1. Proposition 3.7 allows to compare the trajectories of the linear and the nonlinear equation, and thus completes the proof of Theorem 2.4.

By Theorem 2.2, there exists a $c_0 > 0$ such that the deterministic solution x^{det} of (2.6) with initial condition $x_0^{\text{det}} = x_0$ satisfies

$$|x_t^{\text{det}} - x^\star(t)| \leqslant 2c_1\varepsilon + |x_0 - x^\star(0)|\, \mathrm{e}^{-a_0 t/2\varepsilon} \qquad \forall t \in I, \tag{3.1}$$

provided $|x_0 - x^\star(0)| \leqslant c_0$.

Let x_t denote the solution of the SDE (2.1), starting at time $t_0 = 0$ in some x_0. We are interested in the stochastic process $y_t = x_t - x_t^{\text{det}}$, which describes the deviation due to noise from the deterministic solution x_t^{det}. It obeys an SDE of the form

$$\mathrm{d}y_t = \frac{1}{\varepsilon}\big[\bar a(t) y_t + \bar b(y_t, t)\big]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad y_0 = 0, \tag{3.2}$$

where we have introduced the notations

$$\begin{aligned} \bar a(t) &= \bar a_\varepsilon(t) = \partial_x f(x_t^{\text{det}}, t) \\ \bar b(y,t) &= \bar b_\varepsilon(y,t) = f(x_t^{\text{det}} + y, t) - f(x_t^{\text{det}}, t) - \bar a(t) y. \end{aligned} \tag{3.3}$$

Taking ε and $|x_0 - x^\star(0)|$ sufficiently small, we may assume that there exists a constant $\bar d > 0$ such that $|x_t^{\text{det}} + y - x^\star(t)| \leqslant d$ whenever $|y| \leqslant \bar d$. It follows from Taylor's formula that for all $(y,t) \in [-\bar d, \bar d] \times I$,

$$|\bar b(y,t)| \leqslant M y^2 \tag{3.4}$$

$$|\bar a(t) - a(t)| \leqslant M\big(2c_1\varepsilon + |x_0 - x^\star(0)|\, \mathrm{e}^{-a_0 t/2\varepsilon}\big). \tag{3.5}$$

We may further assume that there are constants $\bar a_+ \geqslant \bar a_- > a_0/4$ such that

$$-\bar a_+ \leqslant \bar a(t) \leqslant -\bar a_- \qquad \forall t \in I. \tag{3.6}$$

Finally, the relation $\bar a'(t) = \partial_{xt} f(x_t^{\text{det}}, t) + \partial_{xx} f(x_t^{\text{det}}, t) \frac{1}{\varepsilon} f(x_t^{\text{det}}, t)$ implies the existence of a constant $c_2 > 0$ such that

$$|\bar a'(t)| \leqslant c_2 \Big(1 + |x_0 - x^\star(0)| \frac{\mathrm{e}^{-a_0 t/2\varepsilon}}{\varepsilon}\Big). \tag{3.7}$$

Our analysis will be based on a comparison between solutions of (3.2) and those of the linearized equation

$$\mathrm{d}y_t^0 = \frac{1}{\varepsilon} \bar a(t) y_t^0 \,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad y_0^0 = 0. \tag{3.8}$$

Its solution y_t^0 at time t is a Gaussian random variable with mean zero and variance

$$v(t) = \frac{\sigma^2}{\varepsilon} \int_0^t \mathrm{e}^{2\bar\alpha(t,s)/\varepsilon} \,\mathrm{d}s \qquad \text{where } \bar\alpha(t,s) = \int_s^t \bar a(u) \,\mathrm{d}u. \tag{3.9}$$

Note that (3.6) implies that $\bar\alpha(t,s) \leqslant -\bar a_-(t-s)$ whenever $t \geqslant s$, which implies in particular, that $v(t)$ is not larger than $\sigma^2/2\bar a_-$. We can, however, derive a more precise bound, which is useful when ε and $\mathrm{e}^{-a_0 t/2\varepsilon}$ are small. To do so, we introduce the function

$$\zeta(t) = \frac{1}{2|\bar a(0)|} \mathrm{e}^{2\bar\alpha(t)/\varepsilon} + \frac{1}{\varepsilon} \int_0^t \mathrm{e}^{2\bar\alpha(t,s)/\varepsilon} \,\mathrm{d}s, \qquad \text{where } \bar\alpha(t) = \bar\alpha(t,0). \tag{3.10}$$

Note that $v(t) \leqslant \sigma^2 \zeta(t)$, and that both functions differ by a term which becomes negligible as soon as $t > \mathcal{O}(\varepsilon|\log\varepsilon|)$. The behaviour of $\zeta(t)$ is characterized in the following lemma.

Lemma 3.1. *The function $\zeta(t)$ satisfies the following relations for all $t \in I$.*

$$\zeta(t) = \frac{1}{2|\bar a(t)|} + \mathcal{O}(\varepsilon) + \mathcal{O}\big(|x_0 - x^\star(0)| \,\mathrm{e}^{-a_0 t/2\varepsilon}\big) \tag{3.11}$$

$$\frac{1}{2\bar a_+} \leqslant \zeta(t) \leqslant \frac{1}{2\bar a_-} \tag{3.12}$$

$$\zeta'(t) \leqslant \frac{1}{\varepsilon} \tag{3.13}$$

Proof: By integration by parts, we obtain that

$$\zeta(t) = \frac{1}{-2\bar a(t)} - \frac{1}{2} \int_0^t \frac{\bar a'(s)}{\bar a(s)^2} \mathrm{e}^{2\bar\alpha(t,s)/\varepsilon} \,\mathrm{d}s. \tag{3.14}$$

Using (3.6) and (3.7) we get

$$\begin{aligned}
&\Big| \int_0^t \frac{\bar a'(s)}{\bar a(s)^2} \mathrm{e}^{2\bar\alpha(t,s)/\varepsilon} \,\mathrm{d}s \Big| \\
&\qquad \leqslant \frac{c_2}{\bar a_-^2} \int_0^t \mathrm{e}^{-2\bar a_-(t-s)/\varepsilon} \,\mathrm{d}s + \frac{c_2}{\bar a_-^2} \frac{|x_0 - x^\star(0)|}{\varepsilon} \int_0^t \mathrm{e}^{[-2\bar a_-(t-s) - a_0 s/2]/\varepsilon} \,\mathrm{d}s \\
&\qquad \leqslant \frac{c_2}{2\bar a_-^3} \varepsilon + \frac{c_2}{\bar a_-^2} \frac{|x_0 - x^\star(0)|}{2\bar a_- - a_0/2} \mathrm{e}^{-a_0 t/2\varepsilon},
\end{aligned} \tag{3.15}$$

which proves (3.11). We now observe that $\zeta(t)$ is a solution of the linear ODE

$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = \frac{1}{\varepsilon}\big(2\bar a(t)\zeta + 1\big), \qquad \zeta(0) = \frac{1}{2|\bar a(0)|}. \tag{3.16}$$

Since $\zeta(t) > 0$ and $\bar a(t) < 0$, we have $\zeta'(t) \leqslant 1/\varepsilon$. We also see that $\zeta'(t) \geqslant 0$ whenever $\zeta(t) \leqslant 1/2\bar a_+$ and $\zeta'(t) \leqslant 0$ whenever $\zeta(t) \geqslant 1/2\bar a_-$. Since $\zeta(0)$ belongs to the interval $[1/2\bar a_+, 1/2\bar a_-]$, $\zeta(t)$ must remain in this interval for all t. □

As we have already seen in (2.14), the probability of finding y_t^0 outside a strip of width much larger than $\sqrt{2v(t)}$ is very small. By Lemma 3.1, we now know that $\sqrt{2v(t)}$ behaves approximately like $\sigma|a(t)|^{-1/2}$. One of the key points of the present work is to show that the whole path $\{y_s\}_{0\leqslant s\leqslant t}$ remains in a strip of similar width with high probability. The strip will be defined with the help of $\zeta(t)$ instead of $v(t)$, because we need the width to be bounded away from zero, even for small t.

To investigate y^0_t we need to estimate the stochastic integral $\int_0^t \mathrm{e}^{\bar\alpha(t,u)/\varepsilon}\,\mathrm{d}W_u$. To do so, we would like to use the Bernstein-type inequality

$$\mathbb{P}\Bigl\{\sup_{0\leqslant s\leqslant t}\int_0^s \varphi(u)\,\mathrm{d}W_u \geqslant \delta\Bigr\} \leqslant \exp\Bigl\{-\frac{\delta^2}{2\int_0^t \varphi(u)^2\,\mathrm{d}u}\Bigr\}, \tag{3.17}$$

valid for Borel-measurable deterministic functions $\varphi(u)$. Unfortunately, this estimate cannot be applied directly, because in our case, the integrand depends explicitly on the upper integration limit. This is why we introduce a partition of the interval $[0,t]$.

Lemma 3.2. *Let $\rho : I \to \mathbb{R}_+$ be a measurable, strictly positive function. Fix $K \in \mathbb{N}$, and let $0 = u_0 \leqslant u_1 < \dots < u_K = t$ be a partition of the interval $[0,t]$. Then*

$$\mathbb{P}^{0,0}\Bigl\{\sup_{0\leqslant s\leqslant t}\frac{1}{\rho(s)}\Bigl|\frac{\sigma}{\sqrt{\varepsilon}}\int_0^s \mathrm{e}^{\bar\alpha(s,u)/\varepsilon}\,\mathrm{d}W_u\Bigr| \geqslant h\Bigr\} \leqslant 2\sum_{k=1}^K P_k, \tag{3.18}$$

where

$$P_k = \exp\Bigl\{-\frac12\frac{h^2}{\sigma^2}\Bigl(\inf_{u_{k-1}\leqslant s\leqslant u_k}\rho(s)^2\,\mathrm{e}^{2\bar\alpha(u_k,s)/\varepsilon}\Bigr)\Bigl(\frac{1}{\varepsilon}\int_0^{u_k}\mathrm{e}^{2\bar\alpha(u_k,s)/\varepsilon}\,\mathrm{d}s\Bigr)^{-1}\Bigr\}. \tag{3.19}$$

Proof: We have

$$\begin{aligned}
&\mathbb{P}^{0,0}\Bigl\{\sup_{0\leqslant s\leqslant t}\frac{1}{\rho(s)}\Bigl|\frac{\sigma}{\sqrt{\varepsilon}}\int_0^s \mathrm{e}^{\bar\alpha(s,u)/\varepsilon}\,\mathrm{d}W_u\Bigr| \geqslant h\Bigr\} \qquad (3.20)\\
&\quad= \mathbb{P}^{0,0}\Bigl\{\exists k\in\{1,\dots,K\} : \sup_{u_{k-1}\leqslant s\leqslant u_k}\frac{1}{\rho(s)}\Bigl|\int_0^s \mathrm{e}^{\bar\alpha(s,u)/\varepsilon}\,\mathrm{d}W_u\Bigr| \geqslant \frac{h\sqrt{\varepsilon}}{\sigma}\Bigr\}\\
&\quad\leqslant 2\sum_{k=1}^K \mathbb{P}^{0,0}\Bigl\{\sup_{u_{k-1}\leqslant s\leqslant u_k}\int_0^s \mathrm{e}^{-\bar\alpha(u)/\varepsilon}\,\mathrm{d}W_u \geqslant \frac{h\sqrt{\varepsilon}}{\sigma}\inf_{u_{k-1}\leqslant s\leqslant u_k}\rho(s)\,\mathrm{e}^{-\bar\alpha(s)/\varepsilon}\Bigr\}.
\end{aligned}$$

Applying the Bernstein inequality (3.17) to the last expression, we obtain (3.18). □

Remark 3.3. Note that in the proof of Lemma 3.2 we have not used the monotonicity of $s \mapsto \bar\alpha(t,s)$ so that the estimate (3.18) can also be applied in the case where $\bar a(s)$ changes sign.

We are now ready to derive an upper bound for the probability that y^0_s leaves a strip of appropriate width $h\rho(s)$ before time t. Taking $\rho(s) = \sqrt{\zeta(s)}$ will be a good choice since it leads to approximately constant P_k in (3.18).

Proposition 3.4. *There exists an $r = r(\bar a_+, \bar a_-)$ such that*

$$\mathbb{P}^{0,0}\Bigl\{\sup_{0\leqslant s\leqslant t}\frac{|y^0_s|}{\sqrt{\zeta(s)}} \geqslant h\Bigr\} \leqslant C(t,\varepsilon)\exp\Bigl\{-\frac12\frac{h^2}{\sigma^2}(1-r\varepsilon)\Bigr\}, \tag{3.21}$$

where

$$C(t,\varepsilon) = \frac{|\bar\alpha(t)|}{\varepsilon^2} + 2. \tag{3.22}$$

Proof: Let

$$K = \left\lceil \frac{|\bar{\alpha}(t)|}{2\varepsilon^2} \right\rceil. \tag{3.23}$$

For $k = 1, \dots, K-1$, we define the partition times u_k by the relation

$$|\bar{\alpha}(u_k)| = 2\varepsilon^2 k, \tag{3.24}$$

which is possible since $\bar{\alpha}(t)$ is continuous and decreasing. This definition implies in particular that $\bar{\alpha}(u_k, u_{k-1}) = -2\varepsilon^2$ and, therefore, $u_k - u_{k-1} \leqslant 2\varepsilon^2/\bar{a}_-$. Bounding the integral in (3.19) by $\zeta(u_k)$, we obtain

$$P_k \leqslant \exp\Bigl\{-\frac{1}{2}\frac{h^2}{\sigma^2} \inf_{u_{k-1}\leqslant s\leqslant u_k} \frac{\zeta(s)}{\zeta(u_k)}\, \mathrm{e}^{2\bar{\alpha}(u_k,s)/\varepsilon}\Bigr\}. \tag{3.25}$$

We have $\mathrm{e}^{2\bar{\alpha}(u_k,s)/\varepsilon} \geqslant \mathrm{e}^{-4\varepsilon}$ and

$$\zeta(s) - \zeta(u_k) = -\int_s^{u_k} \zeta'(u)\,\mathrm{d}u \geqslant -\frac{u_k - s}{\varepsilon}. \tag{3.26}$$

Since $\zeta(u_k) \geqslant 1/2\bar{a}_+$, this implies

$$P_k \leqslant \exp\Bigl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\Bigl(1 - 4\frac{\bar{a}_+}{\bar{a}_-}\varepsilon\Bigr)\,\mathrm{e}^{-4\varepsilon}\Bigr\}, \tag{3.27}$$

and the result follows from Lemma 3.2. □

Remark 3.5. If we only assume that $\bar{a}$ is Borel-measurable with $\bar{a}(t) \leqslant -\bar{a}_-$ for all $t \in I$, we still have

$$\mathbb{P}^{0,0}\Bigl\{\sup_{0\leqslant s\leqslant t} |y_s^0| \geqslant h/\sqrt{2\bar{a}_-}\Bigr\} \leqslant C(t,\varepsilon)\exp\Bigl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\,\mathrm{e}^{-4\varepsilon}\Bigr\}. \tag{3.28}$$

To prove this, we choose the same partition as before and bound the integral in (3.19) by $\varepsilon/2\bar{a}_-$.

We now return to the nonlinear equation (3.2), the solutions of which we want to compare to those of its linearization (3.8). To this end, we introduce the events

$$\Omega_t(h) = \Bigl\{\omega\colon \bigl|y_s(\omega)\bigr| < h\sqrt{\zeta(s)}\ \forall s \in [0,t]\Bigr\} \tag{3.29}$$

$$\Omega_t^0(h) = \Bigl\{\omega\colon \bigl|y_s^0(\omega)\bigr| < h\sqrt{\zeta(s)}\ \forall s \in [0,t]\Bigr\}. \tag{3.30}$$

Proposition 3.4 gives us an upper bound on the probability of the complement of $\Omega_t^0(h)$. The key point to control the nonlinear case is a relation between the sets Ω_t and Ω_t^0 (for slightly different values of h). This is done in Proposition 3.7 below.

Notation 3.6. *For two events Ω_1 and Ω_2, we write $\Omega_1 \overset{\text{a.s.}}{\subset} \Omega_2$ if $\mathbb{P}$-almost all $\omega \in \Omega_1$ belong to Ω_2.*

Proposition 3.7. *Let $\gamma = 2\sqrt{2\bar{a}_+}\,M/\bar{a}_-^2$ and assume that $h < \bar{d}\sqrt{\bar{a}_-/2} \wedge \gamma^{-1}$. Then*

$$\Omega_t(h) \overset{\text{a.s.}}{\subset} \Omega_t^0\Bigl(\bigl[1 + \frac{\gamma}{4}h\bigr]h\Bigr) \tag{3.31}$$

$$\Omega_t^0(h) \overset{\text{a.s.}}{\subset} \Omega_t\Bigl(\bigl[1 + \gamma h\bigr]h\Bigr). \tag{3.32}$$

PROOF:

1. The difference $z_s = y_s - y^0_s$ satisfies

$$\frac{\mathrm{d}z_s}{\mathrm{d}s} = \frac{1}{\varepsilon}\bigl[\bar{a}(s)z_s + \bar{b}(y^0_s + z_s, s)\bigr] \tag{3.33}$$

with $z_0 = 0$ $\mathbb{P}$-a.s. Now,

$$z_s = \frac{1}{\varepsilon}\int_0^s \mathrm{e}^{\bar{\alpha}(s,u)/\varepsilon}\,\bar{b}(y^0_u + z_u, u)\,\mathrm{d}u, \tag{3.34}$$

which implies

$$|z_s| \leqslant \frac{1}{\varepsilon}\int_0^s \mathrm{e}^{\bar{\alpha}(s,u)/\varepsilon}|\bar{b}(y_u, u)|\,\mathrm{d}u \tag{3.35}$$

for all $s \in [0,t]$.

2. Let us assume that $\omega \in \Omega_t(h)$. Then we have for all $s \in [0,t]$

$$|y_s(\omega)| \leqslant h\sqrt{\zeta(s)} \leqslant \frac{h}{\sqrt{2\bar{a}_-}} \leqslant \frac{\bar{d}}{2}, \tag{3.36}$$

and thus by (3.35),

$$|z_s(\omega)| \leqslant \frac{1}{\varepsilon}\int_0^s \mathrm{e}^{\bar{\alpha}(s,u)/\varepsilon}\,\frac{Mh^2}{2\bar{a}_-}\,\mathrm{d}u. \tag{3.37}$$

The integral on the right-hand side can be estimated by (3.12), yielding

$$\frac{1}{\varepsilon}\int_0^s \mathrm{e}^{\bar{\alpha}(s,u)/\varepsilon}\,\mathrm{d}u \leqslant 2\zeta_{2\varepsilon}(s) \leqslant \frac{1}{\bar{a}_-}. \tag{3.38}$$

Therefore,

$$|z_s(\omega)| \leqslant \frac{Mh^2}{2\bar{a}_-^2} \leqslant \frac{M\sqrt{\bar{a}_+}\,h}{\sqrt{2}\bar{a}_-^2}h\sqrt{\zeta(s)}, \tag{3.39}$$

which proves (3.31) because $|y^0_s(\omega)| \leqslant |y_s(\omega)| + |z_s(\omega)|$.

3. Let us now assume that $\omega \in \Omega^0_t(h)$. Then we have $|y^0_s(\omega)| \leqslant \bar{d}/2$ for all $s \in [0,t]$ as in (3.36). For $\delta = \gamma h$, we have $\delta < 1$ by assumption. We consider the first exit time

$$\tau = \inf\bigl\{s \in [0,t] \colon |z_s| \geqslant \delta h\sqrt{\zeta(s)}\bigr\} \in [0,t] \cup \{\infty\} \tag{3.40}$$

and the event

$$A = \Omega^0_t(h) \cap \bigl\{\omega \colon \tau(\omega) < \infty\bigr\}. \tag{3.41}$$

If $\omega \in A$, then for all $s \in [0,\tau(\omega)]$, we have $|y_s(\omega)| \leqslant (1+\delta)h\sqrt{\zeta(s)} \leqslant \bar{d}$, and thus by (3.35) and (3.38),

$$|z_s(\omega)| \leqslant \frac{1}{\varepsilon}\int_0^s \mathrm{e}^{\bar{\alpha}(s,u)/\varepsilon}\,\frac{M(1+\delta)^2h^2}{2\bar{a}_-}\,\mathrm{d}u \leqslant \frac{M(1+\delta)^2h^2}{2\bar{a}_-^2} < \delta h\sqrt{\zeta(s)}. \tag{3.42}$$

However, by the definition of τ, we have $|z_{\tau(\omega)}(\omega)| = \delta h\sqrt{\zeta(\tau(\omega))}$, which contradicts (3.42) for $s = \tau(\omega)$. Therefore $\mathbb{P}\{A\} = 0$, which implies that for almost all $\omega \in \Omega^0_t$, we have $|z_s(\omega)| < \delta h\sqrt{\zeta(s)}$ for all $s \in [0,t]$, and hence

$$|y_s(\omega)| < (1+\delta)h\sqrt{\zeta(s)} \quad \forall s \in [0,t] \tag{3.43}$$

for these ω, which proves (3.32). □

We close this subsection with a corollary which is Theorem 2.4, restated in terms of the process y_t. It is a direct consequence of Propositions 3.7 and 3.4.

Corollary 3.8. *There exist h_0 and ε_0, depending only on f, such that for $\varepsilon < \varepsilon_0$ and $h < h_0$,*

$$\mathbb{P}^{0,0}\Big\{\sup_{0\leqslant s\leqslant t} \frac{|y_s|}{\sqrt{\zeta(s)}} > h\Big\} \leqslant C(t,\varepsilon)\exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\big[1-\mathcal{O}(\varepsilon)-\mathcal{O}(h)\big]\Big\}. \tag{3.44}$$

3.2 Unstable case

We now consider a similar situation as in Section 3.1, but with an unstable equilibrium branch, that is, we assume that $a(t) \geqslant a_0 > 0$ for all $t \in I$. Our aim is to prove Theorem 2.6 which is equivalent to Proposition 3.10 below. The proof is again based on a comparison of solutions of the nonlinear equation (2.1) and its linearization around a given deterministic solution.

Theorem 2.2 shows the existence of a particular solution $\widehat{x}_t^{\text{det}}$ of the deterministic equation (2.6) such that $|\widehat{x}_t^{\text{det}} - x^\star(t)| \leqslant c_1\varepsilon$ for all $t \in I$. We are interested in the stochastic process $y_t = x_t - \widehat{x}_t^{\text{det}}$, which describes the deviation due to noise from this deterministic solution $\widehat{x}^{\text{det}}$. It obeys the SDE

$$\mathrm{d}y_t = \frac{1}{\varepsilon}\big[\bar{a}(t)y_t + \bar{b}(y_t,t)\big]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \tag{3.45}$$

where

$$\begin{aligned} \bar{a}(t) &= \bar{a}_\varepsilon(t) = \partial_x f(\widehat{x}_t^{\text{det}},t) \\ \bar{b}(y,t) &= \bar{b}_\varepsilon(y,t) = f(\widehat{x}_t^{\text{det}}+y,t) - f(\widehat{x}_t^{\text{det}},t) - \bar{a}(t)y \end{aligned} \tag{3.46}$$

are the analogs of $\bar{a}$ and $\bar{b}$ defined in (3.3). Taking ε sufficiently small, we may assume that there exist constants $\bar{a}_0, \bar{a}_1, \bar{d} > 0$, such that the following estimates hold for all $t \in I$ and all y such that $|y| \leqslant \bar{d}$:

$$\bar{a}(t) \geqslant \bar{a}_0, \qquad |\bar{a}'(t)| \leqslant \bar{a}_1, \qquad |\bar{b}(y,t)| \leqslant My^2. \tag{3.47}$$

The bound on $|\bar{a}'(t)|$ is a consequence of the analog of (3.7) together with the fact that $|\widehat{x}_0^{\text{det}} - x^\star(0)| = \mathcal{O}(\varepsilon)$.

We first consider the linear equation

$$\mathrm{d}y_t^0 = \frac{1}{\varepsilon}\bar{a}(t)y_t^0\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t. \tag{3.48}$$

Given the initial value y_0^0, the solution y_t^0 at time t is a Gaussian random variable with mean $y_0^0\,\mathrm{e}^{\bar{\alpha}(t)/\varepsilon}$ and variance

$$v(t) = \frac{\sigma^2}{\varepsilon}\int_0^t \mathrm{e}^{2\bar{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s, \tag{3.49}$$

where $\bar{\alpha}(t,s) = \int_s^t \bar{a}(u)\,\mathrm{d}u \geqslant \bar{a}_0(t-s)$ for $t \geqslant s$. The variance, which is growing exponentially fast, can be estimated with the help of the following lemma.

Lemma 3.9. *For $0 < \varepsilon < 2\bar{a}_0^2/\bar{a}_1$, one has*

$$\frac{1}{\varepsilon}\int_0^t \mathrm{e}^{2\bar{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s = \Big[\frac{\mathrm{e}^{2\bar{\alpha}(t)/\varepsilon}}{2\bar{a}(0)} - \frac{1}{2\bar{a}(t)}\Big]\big[1+\mathcal{O}(\varepsilon)\big]. \tag{3.50}$$

Proof: By integration by parts, we obtain that

$$\int_0^t \mathrm{e}^{2\bar{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s = \frac{\varepsilon}{2\bar{a}(0)}\,\mathrm{e}^{2\bar{\alpha}(t)/\varepsilon} - \frac{\varepsilon}{2\bar{a}(t)} - \frac{\varepsilon}{2}\int_0^t \frac{\bar{a}'(s)}{\bar{a}(s)^2}\,\mathrm{e}^{2\bar{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s, \tag{3.51}$$

which implies that

$$\Bigl[1-\frac{\varepsilon}{2}\frac{\bar{a}_1}{\bar{a}_0^2}\Bigr]\int_0^t \mathrm{e}^{2\bar{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s \leqslant \frac{\varepsilon}{2\bar{a}(0)}\,\mathrm{e}^{2\bar{\alpha}(t)/\varepsilon} - \frac{\varepsilon}{2\bar{a}(t)} \leqslant \Bigl[1+\frac{\varepsilon}{2}\frac{\bar{a}_1}{\bar{a}_0^2}\Bigr]\int_0^t \mathrm{e}^{2\bar{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s. \tag{3.52}$$

By our hypothesis on ε, the first term in brackets is positive. □

The following proposition, which restates Theorem 2.6 in terms of y_t, is the main result of this subsection.

Proposition 3.10. *There exist constants $\varepsilon_0, h_0 > 0$ such that for all $h \leqslant \sigma \wedge h_0$, all $\varepsilon \leqslant \varepsilon_0$ and for any given y_0 with $|y_0|\sqrt{2\bar{a}(0)} < h$, we have*

$$\mathbb{P}^{0,y_0}\Bigl\{\sup_{0\leqslant s\leqslant t}|y_s|\sqrt{2\bar{a}(s)} < h\Bigr\} \leqslant \sqrt{\mathrm{e}}\exp\Bigl\{-\kappa\frac{\sigma^2}{h^2}\frac{\bar{\alpha}(t)}{\varepsilon}\Bigr\}, \tag{3.53}$$

where $\kappa = \frac{\pi}{2\mathrm{e}}\bigl(1-\mathcal{O}(h)-\mathcal{O}(\varepsilon)\bigr)$.

Proof:

1. Let $K \in \mathbb{N}$ and let $0 = u_0 < u_1 < \dots < u_K = t$ be any partition of the interval $[0,t]$. We define the events

$$\begin{aligned} A_k &= \Bigl\{\omega\colon \sup_{u_k\leqslant s\leqslant u_{k+1}}|y_s|\sqrt{2\bar{a}(s)} < h\Bigr\} \\ B_k &= \Bigl\{\omega\colon |y_{u_k}|\sqrt{2\bar{a}(u_k)} < h\Bigr\} \supset A_{k-1}. \end{aligned} \tag{3.54}$$

Let q_k be a deterministic upper bound on $P_k = \mathbb{P}^{u_k,y_{u_k}}\{A_k\}$, valid on B_k. Then we have by the Markov property

$$\begin{aligned} &\mathbb{P}^{0,y_0}\Bigl\{\sup_{0\leqslant s\leqslant t}|y_s|\sqrt{2\bar{a}(s)} < h\Bigr\} \\ &= \mathbb{P}^{0,y_0}\Bigl\{\bigcap_{k=0}^{K-1}A_k\Bigr\} = \mathbb{E}^{0,y_0}\Bigl\{1_{\bigcap_{k=0}^{K-2}A_k}\mathbb{E}^{0,y_0}\bigl\{1_{A_K}\bigm|\{y_s\}_{0\leqslant s\leqslant u_{K-1}}\bigr\}\Bigr\} \\ &= \mathbb{E}^{0,y_0}\Bigl\{1_{\bigcap_{k=0}^{K-2}A_k}P_{K-1}\Bigr\} \leqslant q_{K-1}\mathbb{P}^{0,y_0}\Bigl\{\bigcap_{k=0}^{K-2}A_k\Bigr\} \leqslant \dots \leqslant \prod_{k=0}^{K-1}q_k. \end{aligned} \tag{3.55}$$

2. To define the partition, we set

$$K = \Bigl\lceil\frac{1}{\gamma}\frac{\bar{\alpha}(t)}{\varepsilon}\frac{\sigma^2}{h^2}\Bigr\rceil \tag{3.56}$$

for some $\gamma \in (0,1]$ to be chosen later, and

$$\bar{\alpha}(u_{k+1},u_k) = \gamma\varepsilon\frac{h^2}{\sigma^2}, \qquad k = 0,\dots,K-2. \tag{3.57}$$

Since $\bar{\alpha}(u_{k+1},u_k) \geqslant \bar{a}_0(u_{k+1}-u_k)$, we have $u_{k+1}-u_k \leqslant \frac{h^2}{\sigma^2}\frac{\gamma}{\bar{a}_0}\varepsilon$, and using Taylor's formula, we find for all $s\in[u_k,u_{k+1}]$ and all $k=0,\dots,K-1$

$$1-\frac{h^2}{\sigma^2}\frac{\bar{a}_1}{\bar{a}_0^2}\gamma\varepsilon \leqslant \frac{\bar{a}(s)}{\bar{a}(u_k)} \leqslant 1+\frac{h^2}{\sigma^2}\frac{\bar{a}_1}{\bar{a}_0^2}\gamma\varepsilon, \tag{3.58}$$

where $\bar{a}_1$ is the upper bound on $|\bar{a}'|$, see (3.47). In order to estimate P_k, we introduce linear approximations $(y_t^{(k)})_{t\in[u_k,u_{k+1}]}$ for $k\in\{0,\dots,K-2\}$, defined by

$$\mathrm{d}y_t^{(k)} = \frac{1}{\varepsilon}\bar{a}(t)y_t^{(k)} + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t^{(k)}, \qquad y_{u_k}^{(k)} = y_{u_k}, \tag{3.59}$$

where $W_t^{(k)} = W_t - W_{u_k}$ is a Brownian motion with $W_{u_k}^{(k)}=0$ which is independent of $\{W_s\colon 0\leqslant s\leqslant u_k\}$. If $\omega\in A_k$, we have for all $s\in[u_k,u_{k+1}]$

$$\begin{aligned}|y_s(\omega)-y_s^{(k)}(\omega)| &\leqslant \frac{1}{\varepsilon}\int_{u_k}^{s}\mathrm{e}^{\bar{\alpha}(s,u)/\varepsilon}|\bar{b}(y_u,u)|\,\mathrm{d}u\\ &\leqslant \frac{Mh^2}{2\bar{a}_0}\frac{\mathrm{e}^{\bar{\alpha}(u_{k+1},u_k)/\varepsilon}}{\bar{a}(u_k)}\bigl[1+\mathcal{O}(\varepsilon)\bigr] \leqslant r_0\frac{h^2}{\sqrt{2\bar{a}(s)}},\end{aligned} \tag{3.60}$$

where $r_0 = M\,\mathrm{e}(2\bar{a}_0^3)^{-1/2}+\mathcal{O}(\varepsilon)$. This shows that on A_k,

$$|y_s^{(k)}(\omega)| \leqslant \bigl[1+r_0h\bigr]\frac{h}{\sqrt{2\bar{a}(s)}} \qquad \forall s\in[u_k,u_{k+1}]. \tag{3.61}$$

3. We are now ready to estimate P_k. (3.61) shows that on B_k,

$$\begin{aligned}P_k &\leqslant \mathbb{P}^{u_k,y_{u_k}}\Bigl\{\sup_{u_k\leqslant s\leqslant u_{k+1}}|y_s^{(k)}|\sqrt{2\bar{a}(s)} < h(1+r_0h)\Bigr\}\\ &\leqslant \mathbb{P}^{u_k,y_{u_k}}\bigl\{|y_{u_{k+1}}^{(k)}|\sqrt{2\bar{a}(u_{k+1})} < h(1+r_0h)\bigr\}\\ &\leqslant \frac{1}{\sqrt{2\pi v_{u_{k+1}}^{(k)}}}\frac{2h(1+r_0h)}{\sqrt{2\bar{a}(u_{k+1})}},\end{aligned} \tag{3.62}$$

where $v_{u_{k+1}}^{(k)}$ denotes the conditional variance of $y_{u_{k+1}}^{(k)}$, given y_{u_k}. As in (3.50),

$$v_{u_{k+1}}^{(k)} = \frac{\sigma^2}{\varepsilon}\int_{u_k}^{u_{k+1}}\mathrm{e}^{2\bar{\alpha}(u_{k+1},s)/\varepsilon}\,\mathrm{d}s = \frac{\sigma^2}{2}\biggl[\frac{\mathrm{e}^{2\bar{\alpha}(u_{k+1},u_k)/\varepsilon}}{\bar{a}(u_k)} - \frac{1}{\bar{a}(u_{k+1})}\biggr]\bigl[1+\mathcal{O}(\varepsilon)\bigr]. \tag{3.63}$$

It follows that

$$\begin{aligned}\bar{a}(u_{k+1})v_{u_{k+1}}^{(k)} &\geqslant \frac{\sigma^2}{2}\Bigl[\mathrm{e}^{2\gamma h^2/\sigma^2}\frac{\bar{a}(u_{k+1})}{\bar{a}(u_k)}-1\Bigr]\bigl[1-\mathcal{O}(\varepsilon)\bigr]\\ &\geqslant \frac{\sigma^2}{2}\Bigl[\Bigl(1+2\gamma\frac{h^2}{\sigma^2}\Bigr)\Bigl(1-\frac{\bar{a}_1}{\bar{a}_0^2}\frac{h^2}{\sigma^2}\gamma\varepsilon\Bigr)-1\Bigr]\bigl[1-\mathcal{O}(\varepsilon)\bigr]\\ &\geqslant \gamma h^2\Bigl[1-\frac{\bar{a}_1}{2\bar{a}_0^2}(1+2\gamma)\varepsilon\Bigr]\bigl[1-\mathcal{O}(\varepsilon)\bigr]\\ &\geqslant \gamma h^2\bigl[1-\mathcal{O}(\varepsilon)\bigr].\end{aligned} \tag{3.64}$$

Inserting this into (3.62), we obtain for each $k=0,\dots,K-2$ on B_k the estimate

$$P_k \leqslant \frac{2h(1+r_0h)}{\sqrt{2\pi}}\frac{1}{\sqrt{2\gamma h^2}}\bigl[1+\mathcal{O}(\varepsilon)\bigr] = \frac{1}{\sqrt{\pi\gamma}}\bigl[1+\mathcal{O}(\varepsilon)+\mathcal{O}(h)\bigr] =: q. \tag{3.65}$$

Note that for any $\gamma \in (1/\pi, 1]$, there exist $h_0 > 0$ and $\varepsilon_0 > 0$ such that $q < 1$ for all $h \leqslant h_0$ and all $\varepsilon \leqslant \varepsilon_0$. Since $q_{K-1} = 1$ is an obvious bound, we obtain from (3.55)

$$\mathbb{P}^{0,y_0}\Big\{\sup_{0\leqslant s\leqslant t} |y_s|\sqrt{2\bar{a}(s)} < h\Big\} \leqslant q^{K-1} \leqslant \frac{1}{q}\exp\Big\{-\frac{\bar{\alpha}(t)}{\varepsilon}\frac{\sigma^2}{h^2}\frac{1}{2\gamma q^2}q^2\log\big(1/q^2\big)\Big\}. \quad (3.66)$$

Choosing γ so that $q^2 = 1/\mathrm{e}$ holds, yields almost the optimal exponent, and we obtain

$$\mathbb{P}^{0,y_0}\Big\{\sup_{0\leqslant s\leqslant t} |y_s|\sqrt{2\bar{a}(s)} < h\Big\} \leqslant \sqrt{\mathrm{e}}\exp\Big\{-\kappa\frac{\bar{\alpha}(t)}{\varepsilon}\frac{\sigma^2}{h^2}\Big\}. \quad (3.67)$$

□

4 Pitchfork bifurcation

4.1 Preliminaries

In this section, we consider the nonlinear SDE (2.1) in the case where f undergoes a supercritical pitchfork bifurcation, i.e., we require Assumption 2.7 to hold in a region $\mathcal{N} = \{(x,t) \in \mathbb{R}^2 \colon |x| \leqslant d,\ |t| \leqslant T\} \subset \mathcal{N}_0$. The noise intensity σ is assumed to satisfy $\sigma = \mathcal{O}(\sqrt{\varepsilon})$ throughout Section 4. Only in Subsection 4.3, this condition will be slightly strengthened to $\sigma|\log\sigma|^{3/2} = \mathcal{O}(\sqrt{\varepsilon})$.

Recall that we rescaled space and time in order to obtain (2.26). Using Taylor series and the symmetry assumptions, we may write for all $(x,t) \in \mathcal{N}$

$$\begin{aligned} f(x,t) &= a(t)x + b(x,t) = x\big[a(t) + g_0(x,t)\big] \\ \partial_x f(x,t) &= a(t) + g_1(x,t) \end{aligned} \quad (4.1)$$

where $a(t)$, $g_0(x,t)$, $g_1(x,t)$ are twice continuously differentiable functions satisfying

$$\begin{aligned} a(t) &= \partial_x f(0,t) = t + \mathcal{O}(t^2) & & \\ g_0(x,t) &= \big[-1 + \gamma_0(x,t)\big]x^2 & |g_0(x,t)| &\leqslant Mx^2 \\ g_1(x,t) &= \big[-3 + \gamma_1(x,t)\big]x^2 & |g_1(x,t)| &\leqslant 3Mx^2, \end{aligned} \quad (4.2)$$

with γ_0, γ_1 some continuous functions such that $\gamma_0(0,0) = \gamma_1(0,0) = 0$. The following standard result from bifurcation theory is easily obtained by applying the implicit function theorem, see [GH, p. 150] or [IJ, Section II.4] for instance. We state it without proof.

Proposition 4.1. *If T and d are sufficiently small, there exist twice continuously differentiable functions $x^\star, \bar{x} : (0,T] \to \mathbb{R}_+$ of the form*

$$\begin{aligned} x^\star(t) &= \sqrt{t}\big[1 + \mathcal{O}_T(1)\big] \\ \bar{x}(t) &= \sqrt{t/3}\big[1 + \mathcal{O}_T(1)\big] \end{aligned} \quad (4.3)$$

with the following properties:

- *the only solutions of $f(x,t) = 0$ in $\mathcal{N}$ are either of the form $(0,t)$, or of the form $(\pm x^\star(t), t)$ with $t > 0$;*
- *the only solutions of $\partial_x f(x,t) = 0$ in $\mathcal{N}$ are of the form $(\pm\bar{x}(t), t)$ with $t \geqslant 0$;*
- *the derivative of f at $\pm x^\star(t)$ is*

$$a^\star(t) = \partial_x f(x^\star(t),t) = -2t\big[1 + \mathcal{O}_T(1)\big]. \quad (4.4)$$

- *the derivatives of $x^\star(t)$ and $\bar{x}(t)$ satisfy*

$$\frac{\mathrm{d}x^\star}{\mathrm{d}t} = \frac{1}{2\sqrt{t}}[1+\mathcal{O}_T(1)], \qquad \frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = \frac{1}{2\sqrt{3t}}[1+\mathcal{O}_T(1)]. \tag{4.5}$$

As already pointed out in Section 2.3, there is no restriction in assuming T and d to be small. Thus we may assume that the terms $\mathcal{O}_T(1)$ are sufficiently small to do no harm. For instance, we may and will always assume that $a^\star(t) < 0$.

In the following subsections, we are going to analyse the dynamics in three different regions of the (t,x)-plane: near $x=0$ for $t \leqslant \sqrt{\varepsilon}$, near $x=0$ for $t \geqslant \sqrt{\varepsilon}$, and near $x = x^\star(t)$ for $t \geqslant \sqrt{\varepsilon}$.

First, in Subsection 4.2, we analyse the behaviour for $t \leqslant \sqrt{\varepsilon}$. Theorem 2.10 is proved in the same way as Theorem 2.4, the main difference lying in the behaviour of the variance which is investigated in Lemma 4.2.

Subsection 4.3 is devoted to the rather involved proof of Theorem 2.11. We start by giving some preparatory results. Proposition 4.7 estimates the probability of remaining in a smaller strip $\mathcal{S}$ in a similar way as Proposition 3.10. We then show in Lemma 4.8 that the paths are likely to leave $\mathcal{D}(\kappa)$ as well, unless the solution of a suitably chosen linear SDE returns to zero. The probability of such a return to zero is studied in Lemma 4.9. Finally, Theorem 2.11 is proved, the proof being based on an iterative scheme.

The last subsection analyses the motion after $\tau_{\mathcal{D}(\kappa)}$. Here, the main difficulty is to control the behaviour of the deterministic solutions, which are shown to approach $x^\star(t)$, cf. Proposition 4.11. We then prove that the paths of the random process are likely to stay in a neighbourhood of the deterministic solutions. The proof is similar to the corresponding proof in Section 3.1.

4.2 The behaviour for $t \leqslant \sqrt{\varepsilon}$

We begin by considering the linear SDE

$$\mathrm{d}x_t^0 = \frac{1}{\varepsilon}a(t)x_t^0\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t \tag{4.6}$$

with initial condition $x_{t_0}^0 = x_0$ at time $t_0 \in [-T,0)$. Let

$$v(t,t_0) = \frac{\sigma^2}{\varepsilon}\int_{t_0}^t \mathrm{e}^{2\alpha(t,s)/\varepsilon}\,\mathrm{d}s \tag{4.7}$$

denote the variance of x_t^0. As before, we now introduce a function $\zeta(t)$ which will allow us to define a strip that the process x_t is unlikely to leave before time $\sqrt{\varepsilon}$, see Corollary 4.5 below. Let

$$\zeta(t) = \frac{1}{2|a(t_0)|}\,\mathrm{e}^{2\alpha(t,t_0)/\varepsilon} + \frac{1}{\varepsilon}\int_{t_0}^t \mathrm{e}^{2\alpha(t,s)/\varepsilon}\,\mathrm{d}s. \tag{4.8}$$

The following lemma describes the behaviour of $\zeta(t)$.

Lemma 4.2. *There exist constants $c_\pm > 0$ such that*

$$\begin{aligned}
\frac{c_-}{|t|} &\leqslant \zeta(t) \leqslant \frac{c_+}{|t|} && \text{for } t_0 \leqslant t \leqslant -\sqrt{\varepsilon} \\
\frac{c_-}{\sqrt{\varepsilon}} &\leqslant \zeta(t) \leqslant \frac{c_+}{\sqrt{\varepsilon}} && \text{for } -\sqrt{\varepsilon} \leqslant t \leqslant \sqrt{\varepsilon} \\
\frac{c_-}{\sqrt{\varepsilon}}\,\mathrm{e}^{2\alpha(t)/\varepsilon} &\leqslant \zeta(t) \leqslant \frac{c_+}{\sqrt{\varepsilon}}\,\mathrm{e}^{2\alpha(t)/\varepsilon} && \text{for } \sqrt{\varepsilon} \leqslant t \leqslant T.
\end{aligned} \tag{4.9}$$

If, moreover, $a'(t) > 0$ on $[t_0,t]$, then $\zeta(t)$ is increasing on $[t_0,t]$.

Proof: First note that Equation (4.2) implies the existence of constants $a_+ \geqslant a_- > 0$ such that

$$\begin{aligned} a_+ t \leqslant a(t) \leqslant a_- t \qquad &\text{for } -T \leqslant t \leqslant 0 \\ a_- t \leqslant a(t) \leqslant a_+ t \qquad &\text{for } 0 \leqslant t \leqslant T. \end{aligned} \tag{4.10}$$

For $s \leqslant t \leqslant 0$, this implies $-a_+(s^2 - t^2) \leqslant 2\alpha(t,s) \leqslant -a_-(s^2-t^2)$. Integration by parts yields the relation

$$\frac{1}{\varepsilon}\int_{t_0}^t \mathrm{e}^{-a_\pm(s^2-t^2)/\varepsilon}\,\mathrm{d}s = \frac{1}{2a_\pm|t|} - \frac{\mathrm{e}^{-a_\pm(t_0^2-t^2)/\varepsilon}}{2a_\pm|t_0|} - \int_{t_0}^t \frac{\mathrm{e}^{-a_\pm(s^2-t^2)/\varepsilon}}{2a_\pm s^2}\,\mathrm{d}s. \tag{4.11}$$

Since the last two terms on the right-hand side are negative, the upper bound for $t \leqslant -\sqrt{\varepsilon}$ is immediate. For the corresponding lower bound, we use

$$\zeta(t) - \frac{\mathrm{e}^{2\alpha(t,t_0)/\varepsilon}}{2|a(t_0)|} \geqslant \frac{1}{\varepsilon}\int_{t_0\vee 2t}^t \mathrm{e}^{-a_+(s^2-t^2)/\varepsilon}\,\mathrm{d}s \geqslant \frac{1-\mathrm{e}^{-a_+((t_0\vee 2t)^2-t^2)/\varepsilon}}{2a_+|t_0\vee 2t|}, \tag{4.12}$$

where the last inequality is obtained by replacing $\mathrm{e}^{-a_+(s^2-t^2)/\varepsilon}$ by 1 on the right-hand side of (4.11). For $t \leqslant t_0/2$, we thus get $\zeta(t) \geqslant 1/(4a_+|t|)$, while for $t_0/2 < t \leqslant -\sqrt{\varepsilon}$, we find $\zeta(t) \geqslant (1-\mathrm{e}^{-3a_+})/(4a_+|t|)$.

In the case $|t| \leqslant \sqrt{\varepsilon}$, we use the relation

$$\zeta(t) = \zeta(-\sqrt{\varepsilon})\,\mathrm{e}^{2\alpha(t,-\sqrt{\varepsilon})/\varepsilon} + \frac{1}{\varepsilon}\int_{-\sqrt{\varepsilon}}^t \mathrm{e}^{2\alpha(t,s)/\varepsilon}\,\mathrm{d}s. \tag{4.13}$$

Since $|\alpha(t,s)| = \mathcal{O}(\varepsilon)$ for $|t|,|s| \leqslant \sqrt{\varepsilon}$, we conclude that $\zeta(t)$ remains of order $1/\sqrt{\varepsilon}$ for $|t| \leqslant \sqrt{\varepsilon}$. For $t \geqslant \sqrt{\varepsilon}$, we have

$$\mathrm{e}^{-2\alpha(t)/\varepsilon}\,\zeta(t) = \zeta(\sqrt{\varepsilon})\,\mathrm{e}^{-2\alpha(\sqrt{\varepsilon})/\varepsilon} + \frac{1}{\varepsilon}\int_{\sqrt{\varepsilon}}^t \mathrm{e}^{-2\alpha(s)/\varepsilon}\,\mathrm{d}s. \tag{4.14}$$

Now, $2\alpha(s) \geqslant -a_- s^2$ for $s \geqslant 0$ implies that the right-hand side remains of order $1/\sqrt{\varepsilon}$ for all t.

Finally, assume that $a'(t) > 0$ for all t, and recall that $\zeta(t)$ is the solution of the initial value problem

$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = \frac{2a(t)}{\varepsilon}\zeta + \frac{1}{\varepsilon}, \qquad \zeta(t_0) = \frac{1}{2|a(t_0)|}. \tag{4.15}$$

Since $\zeta(t) \geqslant 0$, $\zeta' > 0$ for all positive t. For negative t, ζ' is positive whenever the function $V(t) = \zeta(t) + 1/2a(t)$ is negative. We have $V(t_0) = 0$ and

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{2a(t)}{\varepsilon}V - \frac{a'(t)}{2a(t)^2}. \tag{4.16}$$

Since $V' < 0$ whenever $V = 0$, V can never become positive. This implies $\zeta' \geqslant 0$. □

The following proposition shows that the solution x_t^0 of the linearized equation (4.6) is likely to track the solution of the corresponding deterministic equation.

Proposition 4.3. *Assume that $-T \leqslant t_0 < t \leqslant \sqrt{\varepsilon}$. For sufficiently small ε,*

$$\mathbb{P}^{t_0,x_0}\Big\{\sup_{t_0\leqslant s\leqslant t}\frac{|x^0_s - x_0\,\mathrm{e}^{\alpha(s,t_0)/\varepsilon}|}{\sqrt{\zeta(s)}} > h\Big\} \leqslant C(t,\varepsilon)\exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\big[1-r(\varepsilon)\big]\Big\}, \qquad (4.17)$$

where

$$C(t,\varepsilon) = \frac{|\alpha(t,t_0)|}{\varepsilon^2} + \frac{a_+ + 4\sqrt{\varepsilon}+4}{\varepsilon} \qquad (4.18)$$

and where $r(\varepsilon) = \mathcal{O}(\varepsilon)$ for $t_0 \leqslant t \leqslant -\sqrt{\varepsilon}$, and $r(\varepsilon) = \mathcal{O}(\sqrt{\varepsilon})$ for $-\sqrt{\varepsilon} \leqslant t \leqslant \sqrt{\varepsilon}$.

Proof: We will only give the proof in the case $-\sqrt{\varepsilon} \leqslant t \leqslant \sqrt{\varepsilon}$ as this is the more interesting part. By Lemma 3.2, the probability in (4.17) is bounded by $2\sum_{k=1}^K P_k$, where

$$P_k = \exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\frac{1}{\zeta(u_k)}\inf_{u_{k-1}\leqslant u\leqslant u_k}\zeta(u)\,\mathrm{e}^{2\alpha(u_k,u)/\varepsilon}\Big\} \qquad (4.19)$$

for any partition $t_0 = u_0 < \dots < u_K = t$ of the interval $[t_0,t]$. The choice of the partition should reflect the different behaviour of x^0_s for $s \leqslant -\sqrt{\varepsilon}$ and for $-\sqrt{\varepsilon} \leqslant s \leqslant \sqrt{\varepsilon}$. We set

$$K_0 = \left\lceil \frac{-\alpha(-\sqrt{\varepsilon},t_0)}{2\varepsilon^2}\right\rceil, \qquad K = K_0 + \left\lceil\frac{t+\sqrt{\varepsilon}}{\varepsilon}\right\rceil \qquad (4.20)$$

and define the partition times by

$$\begin{aligned} -\alpha(u_k,t_0) &= 2\varepsilon^2 k && \text{for } 0 \leqslant k \leqslant K_0 - 1,\\ u_k &= -\sqrt{\varepsilon} + \varepsilon(k-K_0) && \text{for } K_0 \leqslant k \leqslant K-1. \end{aligned} \qquad (4.21)$$

Estimating P_k as in the proof of Proposition 3.4, we obtain

$$P_k \leqslant \exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\Big(1-\frac{2\varepsilon}{a_-c_-}\Big)\,\mathrm{e}^{-4\varepsilon}\Big\} \qquad \text{for } 0 \leqslant k \leqslant K_0-1, \qquad (4.22)$$

$$P_k \leqslant \exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\Big(1-\frac{\sqrt{\varepsilon}}{c_-}[1+2a_+c_+]\Big)\,\mathrm{e}^{-a_+\varepsilon}\Big\} \qquad \text{for } K_0 \leqslant k \leqslant K-1. \qquad (4.23)$$

Finally, let us note that

$$2K \leqslant \frac{|\alpha(-\sqrt{\varepsilon},t_0)|}{\varepsilon^2} + \frac{2}{\varepsilon}(t+\sqrt{\varepsilon}) + 4 \leqslant \frac{|\alpha(t,t_0)|}{\varepsilon^2} + \frac{a_+}{\varepsilon} + \frac{4}{\sqrt{\varepsilon}} + 4, \qquad (4.24)$$

which concludes the proof of the proposition. □

Let us now compare solutions of the two SDEs

$$\mathrm{d}x^0_t = \frac{1}{\varepsilon}a(t)x^0_t\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t \qquad x^0_{t_0} = x_0 \qquad (4.25)$$

$$\mathrm{d}x_t = \frac{1}{\varepsilon}f(x_t,t)\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t \qquad x_{t_0} = x_0, \qquad (4.26)$$

where $t_0 \in [-T,0)$. We define the events

$$\Omega^0_t(h) = \Big\{\omega\colon \big|x^0_s(\omega) - x_0\,\mathrm{e}^{\alpha(s,t_0)/\varepsilon}\big| \leqslant h\sqrt{\zeta(s)}\ \forall s\in[t_0,t]\Big\} \qquad (4.27)$$

$$\Omega_t(h) = \Big\{\omega\colon \big|x_s(\omega) - x_0\,\mathrm{e}^{\alpha(s,t_0)/\varepsilon}\big| \leqslant h\sqrt{\zeta(s)}\ \forall s\in[t_0,t]\Big\}. \qquad (4.28)$$

Proposition 4.3 gives us an upper bound on the probability of the complement of $\Omega^0_t(h)$. We now give relations between these events.

Proposition 4.4. *Let $t \in [t_0, \sqrt{\varepsilon}\,]$ and $|x_0| \leqslant h/\varepsilon^{1/4}$, where we assume $h^2 < \varepsilon/\gamma$ for $\gamma = M(1+2\sqrt{c_+})^3 c_+/\sqrt{c_-}$ and $h^2 \leqslant d^2\sqrt{\varepsilon}/(1+2\sqrt{c_+})^2$. Then*

$$\Omega_t(h) \overset{\text{a.s.}}{\subset} \Omega_t^0\Big(\Big[1+\gamma\frac{h^2}{\varepsilon}\Big]h\Big) \tag{4.29}$$

$$\Omega_t^0(h) \overset{\text{a.s.}}{\subset} \Omega_t\Big(\Big[1+\gamma\frac{h^2}{\varepsilon}\Big]h\Big). \tag{4.30}$$

The proof follows along the lines of the proof of the corresponding Proposition 3.7 in the case of nonbifurcating equilibria and we skip it here.

The two preceding propositions immediately imply the main result on the behaviour of the solution of the nonlinear equation (4.26) for $t \leqslant \sqrt{\varepsilon}$, i.e., Theorem 2.10, which we restate here for an arbitrary initial time $t_0 \in [-T, \sqrt{\varepsilon}\,]$.

Corollary 4.5. *Assume that $-T \leqslant t_0 < t \leqslant \sqrt{\varepsilon}$. Then there exists an $h_0 > 0$ such that for all $h \leqslant h_0\sqrt{\varepsilon}$ and all initial conditions x_0 with $|x_0| \leqslant h/\varepsilon^{1/4}$, the following estimate holds:*

$$\mathbb{P}^{t_0,x_0}\Big\{\sup_{t_0\leqslant s\leqslant t}\frac{|x_s - x_0\,\mathrm{e}^{\alpha(s,t_0)/\varepsilon}|}{\sqrt{\zeta(s)}} > h\Big\} \leqslant C(t,\varepsilon)\exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\big[1-r(\varepsilon)-\mathcal{O}(h^2/\varepsilon)\big]\Big\}, \tag{4.31}$$

where $C(t,\varepsilon)$ and $r(\varepsilon)$ are given in Proposition 4.3.

4.3 Escape from the origin

We now consider the SDE (2.1), written in the form

$$\mathrm{d}x_t = \frac{1}{\varepsilon}\big[a(t)x_t + b(x_t,t)\big]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \tag{4.32}$$

for $t \geqslant t_0 \geqslant \sqrt{\varepsilon}$, where we assume that $|x_{t_0}| \leqslant \tilde{x}(t_0)$. Our aim is to estimate for $\kappa \in (0, 2/3)$ the first exit time $\tau_{\mathcal{D}(\kappa)}$ of x_t from $\mathcal{D}(\kappa)$. Recall that $a(t) + \frac{1}{x}b(x,t) \geqslant \kappa a(t)$ holds in $\mathcal{D}(\kappa)$ by the definition of $\mathcal{D}(\kappa)$. Moreover, we have $a_- t \leqslant a(t) \leqslant a_+ t$, $0 \leqslant a'(t) \leqslant a_1$, and $|b(x,t)| \leqslant M|x|^3$ in $\mathcal{D}(\kappa)$.

We first state a result allowing to estimate the variance of the linearization of (4.32).

Lemma 4.6. *Let $a(t)$ be any continuously differentiable, strictly positive, increasing function, and set $\alpha(t,s) = \int_s^t a(u)\,\mathrm{d}u$. Then the integral*

$$v(t,s) = \frac{\sigma^2}{\varepsilon}\int_s^t \mathrm{e}^{2\alpha(t,u)/\varepsilon}\,\mathrm{d}u \tag{4.33}$$

satisfies the inequalities

$$\frac{\sigma^2}{2a(t)}\big[\mathrm{e}^{2\alpha(t,s)/\varepsilon} - 1\big] \leqslant v(t,s) \leqslant \frac{\sigma^2}{2a(s)}\,\mathrm{e}^{2\alpha(t,s)/\varepsilon}. \tag{4.34}$$

PROOF: Using integration by parts, we have

$$\mathrm{e}^{-2\alpha(t,s)/\varepsilon}\,v(t,s) = \sigma^2\Big[\frac{1}{2a(s)} - \frac{1}{2a(t)}\,\mathrm{e}^{-2\alpha(t,s)/\varepsilon} - \int_s^t \frac{a'(u)}{2a(u)^2}\,\mathrm{e}^{-2\alpha(u,s)/\varepsilon}\,\mathrm{d}u\Big]. \tag{4.35}$$

The upper bound follows immediately, and the lower bound is obtained by bounding the exponential in the last integral by 1. □

Our first step towards estimating $\tau_{\mathcal{D}(\kappa)}$ is to estimate the first exit time $\tau_{\mathcal{S}}$ from a smaller strip $\mathcal{S}$, defined as

$$\mathcal{S} = \Bigl\{(x,t)\colon \sqrt{\varepsilon} \leqslant t \leqslant T, |x| < \frac{h}{\sqrt{a(t)}}\Bigr\}, \tag{4.36}$$

where we will choose

$$h = \sqrt{\frac{2}{\kappa}}\,\sigma\sqrt{|\log\sigma|}. \tag{4.37}$$

Proposition 4.7. *Assume $h > \sigma$ and let $t_0 \geqslant \sqrt{\varepsilon}$ and $|x_0| \leqslant h/\sqrt{a(t_0)}$. Then, for any $\mu > 0$, we have*

$$\mathbb{P}^{t_0,x_0}\bigl\{\tau_{\mathcal{S}} \geqslant t\bigr\} \leqslant \Bigl(\frac{h}{\sigma}\Bigr)^{\mu} \exp\Bigl\{-\frac{\mu}{1+\mu}\frac{\alpha(t,t_0)}{\varepsilon}\Bigl[1-\mathcal{O}\Bigl(\frac{1}{\mu\log(h/\sigma)}\Bigr)\Bigr]\Bigr\} \tag{4.38}$$

under the condition

$$\Bigl(\frac{h}{\sigma}\Bigr)^{3+\mu}\Bigl(1+(1+\mu)\frac{\varepsilon}{t_0^2}\log\frac{h}{\sigma}\Bigr) \leqslant \mathcal{O}\Bigl(\frac{t_0^2}{\sigma^2}\Bigr). \tag{4.39}$$

PROOF: We introduce a partition $t_0 = u_0 < \dots < u_K = t$ of the interval $[t_0,t]$ via the relations

$$\alpha(u_k,u_{k-1}) = (1+\mu)\varepsilon\log\frac{h}{\sigma} \qquad \text{for } 1 \leqslant k < K = \Bigl\lceil \frac{\alpha(t,t_0)}{(1+\mu)\varepsilon\log(h/\sigma)}\Bigr\rceil, \tag{4.40}$$

and for each k, we define a linear approximation $(x_t^{(k)})_{t\in[u_k,u_{k+1}]}$ by

$$\mathrm{d}x_t^{(k)} = \frac{1}{\varepsilon}a(t)x_t^{(k)}\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t^{(k)} \qquad x_{u_k}^{(k)} = x_{u_k}, \tag{4.41}$$

where $W_t^{(k)} = W_t - W_{u_k}$. Assume that $|x_s|\sqrt{a(s)} \leqslant h$ for all $s \in [u_k,u_{k+1}]$. Then by Lemma 4.6

$$\begin{aligned} |x_s - x_s^{(k)}| &\leqslant \frac{1}{\varepsilon}\int_{u_k}^{s} |b(x_u,u)|\,\mathrm{e}^{\alpha(s,u)/\varepsilon}\,\mathrm{d}u \\ &\leqslant M\frac{h^3}{a(u_k)^{3/2}}\frac{1}{a(u_k)}\,\mathrm{e}^{\alpha(u_{k+1},u_k)/\varepsilon} \leqslant \frac{h}{\sqrt{a(s)}} \end{aligned} \tag{4.42}$$

for $s \in [u_k,u_{k+1}]$, provided our partition is chosen in such a way that for all k

$$h^2 \leqslant \frac{a_-^2}{M}\sqrt{\frac{a(u_k)}{a(u_{k+1})}}\,\mathrm{e}^{-\alpha(u_{k+1},u_k)/\varepsilon}\,t_0^2. \tag{4.43}$$

Since the partition satisfies

$$\sqrt{\frac{a(u_{k+1})}{a(u_k)}} \leqslant \Bigl(1+\frac{a_1}{a(u_k)}(u_{k+1}-u_k)\Bigr)^{1/2} \leqslant 1 + \frac{a_1}{2a_-^2}(1+\mu)\frac{\varepsilon}{t_0^2}\log\frac{h}{\sigma}, \tag{4.44}$$

we see that Condition (4.43) is satisfied whenever (4.39) holds.

Now, if $|x_{u_k}|\sqrt{a(u_k)} \leqslant h$, then we have

$$\begin{aligned}\mathbb{P}^{u_k,x_{u_k}}\Bigl\{\sup_{u_k\leqslant s\leqslant u_{k+1}}|x_s|\sqrt{a(s)}\leqslant h\Bigr\} &\leqslant \mathbb{P}^{u_k,x_{u_k}}\Bigl\{|x^{(k)}_{u_{k+1}}|\sqrt{a(u_{k+1})}\leqslant 2h\Bigr\} \\ &\leqslant \frac{4h}{\sqrt{2\pi v^{(k)}_{u_{k+1}}a(u_{k+1})}},\end{aligned} \tag{4.45}$$

where the variance

$$v^{(k)}_{u_{k+1}} = \frac{\sigma^2}{\varepsilon}\int_{u_k}^{u_{k+1}} \mathrm{e}^{2\alpha(u_{k+1},s)/\varepsilon}\,\mathrm{d}s \tag{4.46}$$

can be estimated by Lemma 4.6. We thus have by the Markov property

$$P = \mathbb{P}^{t_0,x_0}\Bigl\{\sup_{t_0\leqslant s\leqslant t}|x_s|\sqrt{a(s)}\leqslant h\Bigr\} \leqslant \prod_{k=0}^{K-1}\biggl(\frac{4}{\sqrt{2\pi}}\frac{h}{\sqrt{v^{(k)}_{u_{k+1}}a(u_{k+1})}}\wedge 1\biggr), \tag{4.47}$$

which immediately implies (4.38). □

We want to choose μ in such a way that $\mathbb{P}^{t_0,x_0}\{\tau_{\mathcal{S}} \geqslant t\} \leqslant (h/\sigma)^{\mu}\,\mathrm{e}^{-\kappa\alpha(t,t_0)/\varepsilon}$ holds with the same κ as in the definition of $\mathcal{D}(\kappa)$. We opt for $\mu = 2$, because this choice guarantees the above estimate for all $\kappa < 2/3$ without choosing a κ-dependent μ. For $h = (2/\kappa)^{1/2}\sigma\sqrt{|\log\sigma|}$ and small enough ε, Condition (4.39) becomes a consequence of the following slightly stronger condition

$$\sigma|\log\sigma|^{3/2} = \mathcal{O}(\sqrt{\varepsilon}), \tag{4.48}$$

which we will assume to be satisfied from now on for the rest of this subsection.

The second step is to control the probability that x_t returns to zero after it has left the strip $\mathcal{S}$. To do so, we will compare solutions of (4.32) with those of the linear equation

$$\mathrm{d}x^0_t = \frac{1}{\varepsilon}a_0(t)x^0_t\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \tag{4.49}$$

where $a_0(t) = \kappa a(t)$ satisfies $a_0(t) \leqslant f(x,t)/x$ in $\mathcal{D}(\kappa)$. The following lemma shows that this choice of $a_0(s)$ implies that $|x_s| \geqslant |x^0_s|$ holds as long as x_s does not return to zero (Fig. 3). This implies that if x^0_s does not return to zero before time t, then x_s is likely to leave $\mathcal{D}(\kappa)$ before time t without returning to zero.

Lemma 4.8. *Let $t_0 \geqslant \sqrt{\varepsilon}$ and assume that $0 < x_0 < \tilde{x}(t_0)$. We define*

$$\mathcal{D}^+(\kappa) = \bigl\{(x,s)\colon \sqrt{\varepsilon}\leqslant s\leqslant t \text{ and } 0 < x < \tilde{x}(s)\bigr\} \tag{4.50}$$

and denote by $\tau_{\mathcal{D}^+(\kappa)}$ the first exit time of x_s from $\mathcal{D}^+(\kappa)$. Let τ^0 be the time of first return to zero of x^0_s in $[t_0,t]$, where we set $\tau^0 = \infty$ if $x^0_s > 0$ for all $s \in [t_0,t]$. Then $x_s \geqslant x^0_s$ for all $s \leqslant \tau_{\mathcal{D}^+(\kappa)} \wedge t$ and

$$\begin{aligned}\mathbb{P}^{t_0,x_0}\Bigl\{0 < x_s < \tilde{x}(s)\ \forall s\in[t_0,t], \tau^0 = \infty\Bigr\} &\leqslant \mathbb{P}^{t_0,x_0}\Bigl\{0 < x^0_s < \tilde{x}(s)\ \forall s\in[t_0,t]\Bigr\} \\ &\leqslant \frac{\tilde{x}(t)\sqrt{a_0(t)}}{\sqrt{\pi}\sigma}\frac{\mathrm{e}^{-\kappa\alpha(t,t_0)/\varepsilon}}{\sqrt{1-\mathrm{e}^{-2\kappa\alpha(t,t_0)/\varepsilon}}}.\end{aligned} \tag{4.51}$$

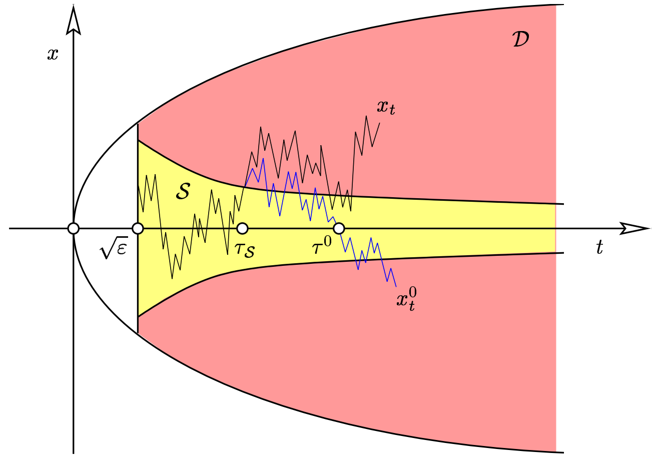

FIGURE 3. Assume the path x_t exits the region $\mathcal{S}$ at time $\tau_{\mathcal{S}}$, say by passing through the upper boundary of $\mathcal{S}$. We introduce a process x_t^0, starting on the same boundary at time $\tau_{\mathcal{S}}$, which obeys the linear SDE (4.49). Let τ^0 be the time of first return to zero of x_t^0. Then x_t lies above x_t^0 for $\tau_{\mathcal{S}} < t \leqslant \tau^0$. In case x_t also becomes negative, the two processes may cross each other. The probability of x_t^0 ever returning to zero is bounded by σ^2. If x_t^0 does not return to zero, x_t is likely to leave $\mathcal{D} = \mathcal{D}(\kappa)$.

PROOF:

1. Let $g(x,s) = f(x,s) - a_0(s)x$. By assumption, $g(x,s)$ is non-negative for $(x,s) \in \mathcal{D}^+(\kappa)$. The difference $z_s = x_s - x_s^0$ satisfies the equation

$$z_s = z_{t_0} + \frac{1}{\varepsilon}\int_{t_0}^{s} \bigl[g(x_u,u) + a_0(u)z_u\bigr]\,\mathrm{d}u \tag{4.52}$$

with $z_{t_0} = 0$. Since $g(x_s,s) \geqslant 0$ for $t_0 \leqslant s \leqslant \tau_{\mathcal{D}^+(\kappa)} \wedge t$,

$$z_s \geqslant z_{t_0} + \frac{1}{\varepsilon}\int_0^s a_0(u) z_u \,\mathrm{d}u, \tag{4.53}$$

follows for all such s and, therefore, Gronwall's lemma yields

$$z_s \geqslant z_{t_0}\,\mathrm{e}^{\kappa\alpha(s,t_0)/\varepsilon} = 0 \qquad \text{for all } s \in [t_0, \tau_{\mathcal{D}^+(\kappa)} \wedge t]. \tag{4.54}$$

This shows $x_s \geqslant x_s^0$ for those s. Now assume $\tau_{\mathcal{D}^+(\kappa)} = \infty$ and $\tau^0 = \infty$. Then, (4.54) implies that $0 < x_s^0 \leqslant x_s < \tilde{x}(s)$ for all $s \leqslant t$, which shows the first inequality in (4.51).

2. x_s^0 being distributed according to a normal law, we have

$$\begin{aligned}\mathbb{P}^{t_0,x_0}\bigl\{0 < x_s^0 < \tilde{x}(s)\ \forall s \in [t_0,t]\bigr\} &\leqslant \mathbb{P}^{t_0,x_0}\bigl\{0 < x_t^0 < \tilde{x}(t)\bigr\} \\ &\leqslant \frac{\tilde{x}(t)}{\sqrt{2\pi v_0(t,t_0)}},\end{aligned} \tag{4.55}$$

where the variance $v_0(t,t_0)$ of x_t^0 can be estimated by Lemma 4.6. This proves the second inequality in (4.51). □

The previous lemma is useful only if we can control the probability that the solution x_t^0 of the linearized equation returns to zero. The following result estimates this probability and its density.

Lemma 4.9. *Let $t_0 \geqslant \sqrt{\varepsilon}$ and assume that $x^0_{t_0} = \rho > \sigma/\sqrt{a_0(t_0)}$. Denote by τ^0 the time of the first return of x^0_t to zero. Then we have*

$$\mathbb{P}^{t_0,\rho}\{\tau^0 < t\} \leqslant \mathbb{P}^{t_0,\rho}\{\tau^0 < \infty\} \leqslant \mathrm{e}^{-a_0(t_0)\rho^2/\sigma^2} \tag{4.56}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}^{t_0,\rho}\{\tau^0 < t\} \leqslant \frac{2}{\sqrt{\pi}}\sqrt{a_0(t_0)}\,\frac{\rho}{\sigma}\,\mathrm{e}^{-a_0(t_0)\rho^2/\sigma^2}\,\frac{1}{\varepsilon}\sqrt{a_0(t)a_0(t_0)}\,\frac{\mathrm{e}^{-2\kappa\alpha(t,t_0)/\varepsilon}}{\sqrt{1-\mathrm{e}^{-2\kappa\alpha(t,t_0)/\varepsilon}}}. \tag{4.57}$$

PROOF:

1. Since by symmetry, $\mathbb{P}^{\tau^0,0}\{x^0_t \geqslant 0\} = \frac12$ on $\{\tau^0 < t\}$, we have by the strong Markov property
$$\mathbb{P}^{t_0,\rho}\{x^0_t \geqslant 0 | \tau^0 < t\} = \frac{1}{2}. \tag{4.58}$$
We now observe that
$$\begin{aligned}\mathbb{P}^{t_0,\rho}\{x^0_t \geqslant 0\} &= \mathbb{P}^{t_0,\rho}\{x^0_t \geqslant 0, \tau^0 \geqslant t\} + \mathbb{P}^{t_0,\rho}\{x^0_t \geqslant 0, \tau^0 < t\}\\ &= \mathbb{P}^{t_0,\rho}\{\tau^0 \geqslant t\} + \mathbb{P}^{t_0,\rho}\{x^0_t \geqslant 0|\tau^0 < t\}\mathbb{P}^{t_0,\rho}\{\tau^0 < t\}\\ &= 1 - \mathbb{P}^{t_0,\rho}\{\tau^0 < t\} + \tfrac12\mathbb{P}^{t_0,\rho}\{\tau^0 < t\}\\ &= 1 - \tfrac12\mathbb{P}^{t_0,\rho}\{\tau^0 < t\},\end{aligned} \tag{4.59}$$
which implies
$$\mathbb{P}^{t_0,\rho}\{\tau^0 < t\} = 2\bigl[1 - \mathbb{P}^{t_0,\rho}\{x^0_t \geqslant 0\}\bigr] = 2\mathbb{P}^{t_0,\rho}\{x^0_t < 0\}. \tag{4.60}$$
2. By Lemma 4.6, the variance $v_0(t,t_0)$ of x^0_t satisfies
$$\Xi = \frac{\rho^2\,\mathrm{e}^{2\kappa\alpha(t,t_0)/\varepsilon}}{2v_0(t,t_0)} \geqslant a_0(t_0)\frac{\rho^2}{\sigma^2}, \tag{4.61}$$
and we thus have
$$\begin{aligned}\mathbb{P}^{t_0,\rho}\{x^0_t < 0\} &= \frac{1}{\sqrt{2\pi v_0(t,t_0)}}\int_{-\infty}^0 \exp\Bigl\{-\frac{(x-\rho\,\mathrm{e}^{\kappa\alpha(t,t_0)/\varepsilon})^2}{2v_0(t,t_0)}\Bigr\}\,\mathrm{d}x\\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\frac{\rho\,\mathrm{e}^{\kappa\alpha(t,t_0)/\varepsilon}}{\sqrt{v_0(t,t_0)}}} \mathrm{e}^{-y^2/2}\,\mathrm{d}y \leqslant \frac12\,\mathrm{e}^{-\Xi},\end{aligned} \tag{4.62}$$
which proves (4.56), using (4.60) and (4.61).
3. In order to compute the derivative of $\mathbb{P}^{t_0,\rho}\{x^0_t < 0\}$, we first note that
$$\frac{\mathrm{d}}{\mathrm{d}t}v_0(t,t_0) = \frac{\sigma^2}{\varepsilon} + \frac{2a_0(t)}{\varepsilon}v_0(t,t_0). \tag{4.63}$$
Differentiating the second line of (4.62), we get
$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{P}^{t_0,\rho}\{x^0_t < 0\} &= \frac{1}{\sqrt{2\pi}}\exp\Bigl\{-\frac{\rho^2\,\mathrm{e}^{2\kappa\alpha(t,t_0)/\varepsilon}}{2v_0(t,t_0)}\Bigr\}\frac{\mathrm{d}}{\mathrm{d}t}\Bigl[-\frac{\rho\,\mathrm{e}^{\kappa\alpha(t,t_0)/\varepsilon}}{\sqrt{v_0(t,t_0)}}\Bigr]\\ &= \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\Xi}\,\frac{\rho}{2}\frac{\sigma^2}{\varepsilon}\frac{\mathrm{e}^{\kappa\alpha(t,t_0)/\varepsilon}}{v_0(t,t_0)^{3/2}}\\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\rho}\frac{\sigma^2}{\varepsilon}\frac{\mathrm{e}^{-\kappa\alpha(t,t_0)/\varepsilon}}{\sqrt{v_0(t,t_0)}}\,\Xi\,\mathrm{e}^{-\Xi}.\end{aligned} \tag{4.64}$$
Since $\Xi > a_0(t_0)\rho^2/\sigma^2 > 1$ and $\Xi\,\mathrm{e}^{-\Xi}$ is decreasing for $\Xi > 1$, we obtain the second bound (4.57) by using again (4.60) and Lemma 4.6. □

Assume for the moment that x_t^0 starts "on the border" of $\mathcal{S}$, i.e. in $\rho(t_0) = h/\sqrt{a(t_0)} = \sqrt{\kappa}h/\sqrt{a_0(t_0)}$. Then, by our choice $h = (2/\kappa)^{1/2}\sigma\sqrt{|\log\sigma|}$, Estimate (4.56) shows that the probability for x_t^0 to return to zero cannot exceed $\mathrm{e}^{-a_0(t_0)\rho^2/\sigma^2} = \sigma^2$.

We are now ready to prove the main estimate on the first exit time $\tau_{\mathcal{D}(\kappa)}$, which is the most important of our results. Since the proof is rather involved, we restate Theorem 2.11 here for convenience.

Proposition 4.10 (Theorem 2.11)**.** *Let $t_0 \geqslant \sqrt{\varepsilon}$ and $|x_0| \leqslant \tilde{x}(t_0)$. Then*

$$\mathbb{P}^{t_0,x_0}\bigl\{\tau_{\mathcal{D}(\kappa)} \geqslant t\bigr\} \leqslant C_0\,\tilde{x}(t)\sqrt{a(t)}\,\frac{|\log\sigma|}{\sigma}\biggl(1+\frac{\alpha(t,t_0)}{\varepsilon}\biggr)\frac{\mathrm{e}^{-\kappa\alpha(t,t_0)/\varepsilon}}{\sqrt{1-\mathrm{e}^{-2\kappa\alpha(t,t_0)/\varepsilon}}}, \tag{4.65}$$

where $C_0 > 0$ is a (numerical) constant.

The strategy of the proof can be summarized as follows. The paths are likely to leave $\mathcal{S}$ after a short time. Then there are two possibilities. Either the solution x_t^0 of the linear equation (4.49) does not return to zero, and Lemma 4.8 shows that x_t is likely to leave $\mathcal{D}(\kappa)$ as well. Or x_t^0 does return to zero. Using the (strong) Markov property and integrating over the distribution of the time of such a (first) return to zero, we obtain an integral equation for an upper bound on the probability of remaining in $\mathcal{D}(\kappa)$. Finally, this integral equation is solved by iterations.

Proof of Proposition 4.10.

1. We first introduce some notations. Let

$$\Phi_t(s,x) = \mathbb{P}^{s,x}\bigl\{\tau_{\mathcal{D}(\kappa)} \geqslant t\bigr\} = \mathbb{P}^{s,x}\Bigl\{\sup_{s\leqslant u\leqslant t}\frac{|x_u|}{\tilde{x}(u)} < 1\Bigr\}, \tag{4.66}$$

and define $\rho(t) = h/\sqrt{a(t)}$. We may assume that $\rho(t) \leqslant \tilde{x}(t)$ for all t (otherwise we replace $\tilde{x}$ by its maximum with ρ). For $t \geqslant s \geqslant \sqrt{\varepsilon}$ we define the quantities

$$q_t(s) = \sup_{|x|\leqslant\rho(s)}\Phi_t(s,x), \tag{4.67}$$

$$Q_t(s) = \sup_{\rho(s)\leqslant|x|\leqslant\tilde{x}(s)}\Phi_t(s,x). \tag{4.68}$$

2. Let us first consider the case $|x| \leqslant \rho(s)$. Recall that $\mathcal{S} = \{(x,t)\colon |x| < \rho(t)\}$. By Proposition 4.7 and the strong Markov property, we have the estimate

$$\begin{aligned}\Phi_t(s,x) &= \mathbb{P}^{s,x}\bigl\{\tau_{\mathcal{S}} \geqslant t\bigr\} + \mathbb{P}^{s,x}\Bigl\{\tau_{\mathcal{S}} < t, \sup_{\tau_{\mathcal{S}}\leqslant u\leqslant t}\frac{|x_u|}{\tilde{x}(u)} < 1\Bigr\}\\ &\leqslant \Bigl(\frac{h}{\sigma}\Bigr)^2\mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon} + \mathbb{E}^{s,x}\Bigl\{1_{\{\tau_{\mathcal{S}}<t\}}\mathbb{P}^{\tau_{\mathcal{S}},x_{\tau_{\mathcal{S}}}}\Bigl\{\sup_{\tau_{\mathcal{S}}\leqslant u\leqslant t}\frac{|x_u|}{\tilde{x}(u)} < 1\Bigr\}\Bigr\}\\ &\leqslant \Bigl(\frac{h}{\sigma}\Bigr)^2\mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon} + \mathbb{E}^{s,x}\bigl\{1_{[s,t)}(\tau_{\mathcal{S}})Q_t(\tau_{\mathcal{S}})\bigr\}.\end{aligned} \tag{4.69}$$

The second term can be estimated by integration by parts, see Lemma A.1. Let $\overline{Q}_t(u)$ be any upper bound on $Q_t(u)$ satisfying the hypotheses on g in that lemma. Since $Q_t(u) \leqslant Q_t(t) = 1$, we may assume that $\overline{Q}_t(t) = 1$. Application of (A.1) with $G(u) = 1 - (h/\sigma)^2\,\mathrm{e}^{-\kappa\alpha(u,s)/\varepsilon}$ shows that the second term in (4.69) is bounded by

$$\Bigl(\frac{h}{\sigma}\Bigr)^2\mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon} + \kappa\Bigl(\frac{h}{\sigma}\Bigr)^2\int_s^t\overline{Q}_t(u)\frac{a(u)}{\varepsilon}\,\mathrm{e}^{-\kappa\alpha(u,s)/\varepsilon}\,\mathrm{d}u. \tag{4.70}$$

We have thus obtained the inequality

$$q_t(s) \leqslant 2\Big(\frac{h}{\sigma}\Big)^2 \operatorname{e}^{-\kappa\alpha(t,s)/\varepsilon} + \kappa\Big(\frac{h}{\sigma}\Big)^2 \int_s^t \overline{Q}_t(u)\frac{a(u)}{\varepsilon}\operatorname{e}^{-\kappa\alpha(u,s)/\varepsilon}\,\mathrm{d}u. \tag{4.71}$$

3. Consider now the case $|x| \in [\rho(s), \tilde{x}(s)]$. Since $x \mapsto f(x,t)$ is an odd function, $\Phi_t(s,x) = \Phi_t(s,-x)$ follows. Hence we may assume that $x > 0$. We consider the linear SDE (4.49) with initial condition $x^0_s = x$, and denote by τ^0 the time of the first return of x^0_t to zero. Then we have

$$\Phi_t(s,x) = \mathbb{P}^{s,x}\Big\{\tau^0 \geqslant t, \sup_{s\leqslant u\leqslant t}\frac{|x_u|}{\tilde{x}(u)} < 1\Big\} + \mathbb{P}^{s,x}\Big\{\tau^0 < t, \sup_{s\leqslant u\leqslant t}\frac{|x_u|}{\tilde{x}(u)} < 1\Big\}, \tag{4.72}$$

and Lemma 4.8 yields

$$\mathbb{P}^{s,x}\Big\{\tau^0 \geqslant t, \sup_{s\leqslant u\leqslant t}\frac{|x_u|}{\tilde{x}(u)} < 1\Big\} \leqslant \frac{\tilde{x}(t)\sqrt{\kappa a(t)}}{\sqrt{\pi}\sigma}\frac{\operatorname{e}^{-\kappa\alpha(t,s)/\varepsilon}}{\sqrt{1-\operatorname{e}^{-2\kappa\alpha(t,s)/\varepsilon}}}. \tag{4.73}$$

The second term in (4.72) can be estimated using the density of the random variable τ^0, for which Lemma 4.9 gives the bound

$$\psi_{\tau^0}(u) = \frac{\mathrm{d}}{\mathrm{d}u}\mathbb{P}^{s,x}\{\tau^0 < u\} \leqslant \frac{2\kappa^{3/2}}{\sqrt{\pi}}\frac{h}{\sigma}\operatorname{e}^{-\kappa h^2/\sigma^2}\frac{a(u)}{\varepsilon}\frac{\operatorname{e}^{-2\kappa\alpha(u,s)/\varepsilon}}{\sqrt{1-\operatorname{e}^{-2\kappa\alpha(u,s)/\varepsilon}}}. \tag{4.74}$$

We obtain

$$\begin{aligned}
\mathbb{P}^{s,x}\Big\{\tau^0 < t, \sup_{s\leqslant u\leqslant t}\frac{|x_u|}{\tilde{x}(u)} < 1\Big\} &\leqslant \mathbb{E}^{s,x}\Big\{1_{\{\tau^0<t\}}\mathbb{P}^{\tau^0,x_{\tau^0}}\Big\{\sup_{\tau^0\leqslant u\leqslant t}\frac{|x_u|}{\tilde{x}(u)} < 1\Big\}\Big\}\\
&= \int_s^t \psi_{\tau^0}(u)\Phi_t(u,x_u)\,\mathrm{d}u\\
&\leqslant \int_s^t \psi_{\tau^0}(u)\big[q_t(u) + Q_t(u)\big]\,\mathrm{d}u.
\end{aligned} \tag{4.75}$$

4. Before inserting the estimate (4.71) for $q_t(u)$, we shall introduce some notations and provide bounds for certain integrals needed in the sequel. Let

$$g(t,s) = \frac{\operatorname{e}^{-\kappa\alpha(t,s)/\varepsilon}}{\sqrt{1-\operatorname{e}^{-2\kappa\alpha(t,s)/\varepsilon}}} \tag{4.76}$$

and $\phi = \operatorname{e}^{-\kappa\alpha(t,s)/\varepsilon}$. Then

$$\int_s^t \frac{a(u)}{\varepsilon}\operatorname{e}^{-\kappa\alpha(u,s)/\varepsilon}g(u,s)\,\mathrm{d}u \leqslant \int_s^t \frac{a(u)}{\varepsilon}g(u,s)\,\mathrm{d}u \leqslant \frac{\pi}{2\kappa} \leqslant \frac{2}{\kappa} \tag{4.77}$$

$$\int_s^t \frac{a(u)}{\varepsilon}\operatorname{e}^{-\kappa\alpha(u,s)/\varepsilon}g(t,u)g(u,s)\,\mathrm{d}u = \frac{\phi}{2\kappa}\int_0^1 \frac{\mathrm{d}x}{\sqrt{x(1-x)}} = \frac{\pi}{2\kappa}\phi < \frac{2}{\kappa}\phi \tag{4.78}$$

$$\begin{aligned}
\int_s^t \frac{a(u)}{\varepsilon}\operatorname{e}^{-\kappa\alpha(u,s)/\varepsilon}g(t,u)\,\mathrm{d}u &\leqslant \frac{\phi}{\kappa}\int_0^{\sqrt{1-\phi^2}}\frac{1}{1-x^2}\,\mathrm{d}x = \frac{\phi}{\kappa}\frac{1}{2}\log\frac{1+\sqrt{1-\phi^2}}{1-\sqrt{1-\phi^2}}\\
&\leqslant \frac{\phi}{\kappa}\log\frac{2}{\phi} \leqslant \Big[\frac{1}{\kappa} + \frac{\alpha(t,s)}{\varepsilon}\Big]\operatorname{e}^{-\kappa\alpha(t,s)/\varepsilon},
\end{aligned} \tag{4.79}$$

where we used the changes of variables $\operatorname{e}^{-2\kappa\alpha(u,s)/\varepsilon} = x(1-\phi^2) + \phi^2$ in (4.78) and $x^2 = 1 - \operatorname{e}^{-2\kappa\alpha(t,u)/\varepsilon}$ in (4.79).

5. Now we are ready to return to our estimate on $\int_s^t \psi_{\tau^0}(u) q_t(u) \, du$, compare (4.75). Inserting the bound (4.71) on $q_t(u)$ yields two summands, the first one being

$$
\begin{aligned}
2\Big(\frac{h}{\sigma}\Big)^2 & \int_s^t \psi_{\tau^0}(u) \, e^{-\kappa\alpha(t,u)/\varepsilon} \, du \\
& \leqslant \frac{4\kappa^{3/2}}{\sqrt{\pi}} \Big(\frac{h}{\sigma}\Big)^3 e^{-\kappa h^2/\sigma^2} \int_s^t \frac{a(u)}{\varepsilon} \frac{e^{-2\kappa\alpha(u,s)/\varepsilon}}{\sqrt{1 - e^{-2\kappa\alpha(u,s)/\varepsilon}}} \, e^{-\kappa\alpha(t,u)/\varepsilon} \, du \\
& \leqslant 2\sqrt{\pi\kappa} \Big(\frac{h}{\sigma}\Big)^3 e^{-\kappa h^2/\sigma^2} \, e^{-\kappa\alpha(t,s)/\varepsilon},
\end{aligned}
\tag{4.80}
$$

where we used (4.77) to bound the integral. The second summand is

$$
\begin{aligned}
\kappa\Big(\frac{h}{\sigma}\Big)^2 & \int_s^t \psi_{\tau^0}(u) \int_u^t \overline{Q}_t(v) \frac{a(v)}{\varepsilon} \, e^{-\kappa\alpha(v,u)/\varepsilon} \, dv \, du \\
& \leqslant \kappa\sqrt{\pi\kappa} \Big(\frac{h}{\sigma}\Big)^3 e^{-\kappa h^2/\sigma^2} \int_s^t \overline{Q}_t(v) \frac{a(v)}{\varepsilon} \, e^{-\kappa\alpha(v,s)/\varepsilon} \, dv,
\end{aligned}
\tag{4.81}
$$

where we used (4.77) again.
We can now collect terms. Introducing the abbreviations

$$
C = \max\Big\{ \frac{\tilde{x}(t)\sqrt{\kappa a(t)}}{\sqrt{\pi}\sigma}, 1 \Big\} \qquad \text{and} \qquad c = \sqrt{\pi\kappa} \Big(\frac{h}{\sigma}\Big)^3 e^{-\kappa h^2/\sigma^2},
\tag{4.82}
$$

the previous inequalities imply that

$$
Q_t(s) \leqslant C g(t,s) + c \, e^{-\kappa\alpha(t,s)/\varepsilon} + c \int_s^t \overline{Q}_t(u) \frac{a(u)}{\varepsilon} \, e^{-\kappa\alpha(u,s)/\varepsilon} \big[1 + g(u,s)\big] \, du.
\tag{4.83}
$$

6. We will now iterate the bounds on $Q_t(s)$. This will show the existence of two series $\{a_n\}_{n\geqslant 1}$ and $\{b_n\}_{n\geqslant 1}$ such that

$$
Q_t(s) \leqslant C g(t,s) + a_n \, e^{-\kappa\alpha(t,s)/\varepsilon} + b_n \qquad \forall n.
\tag{4.84}
$$

To do so, we need to assume that

$$
c\Big(\frac{\alpha(T,t_0)}{\varepsilon} + \frac{2}{\kappa}\Big) = \sqrt{\pi\kappa}\Big(\frac{\alpha(T,t_0)}{\varepsilon} + \frac{2}{\kappa}\Big)\Big(\frac{h}{\sigma}\Big)^3 e^{-\kappa h^2/\sigma^2} \leqslant \frac{1}{2}.
\tag{4.85}
$$

By our choice (4.37) of h, this condition reduces to

$$
\frac{1}{\kappa} \sigma |\log\sigma|^{3/4} = \mathcal{O}(\sqrt{\varepsilon}),
\tag{4.86}
$$

which is satisfied for small enough ε by our assumption (4.48) on σ.
Using the trivial bound $\overline{Q}_t(u) = 1$ in (4.83), we find that (4.84) holds with $a_1 = c$ and $b_1 = 3c/\kappa$. Inserting (4.84) into (4.83) again, we get

$$
\begin{aligned}
Q_t(s) & \leqslant C g(t,s) + c \, e^{-\kappa\alpha(t,s)/\varepsilon} \\
& \quad + c \int_s^t \Big[C g(t,u) + a_n \, e^{-\kappa\alpha(t,u)/\varepsilon} + b_n \Big] \frac{a(u)}{\varepsilon} \, e^{-\kappa\alpha(u,s)/\varepsilon} \big[1 + g(u,s)\big] \, du \\
& \leqslant C g(t,s) + c\Big[1 + C\Big(\frac{\alpha(t,s)}{\varepsilon} + \frac{3}{\kappa}\Big) + a_n\Big(\frac{\alpha(t,s)}{\varepsilon} + \frac{2}{\kappa}\Big)\Big] e^{-\kappa\alpha(t,s)/\varepsilon} + \frac{3c}{\kappa} b_n.
\end{aligned}
\tag{4.87}
$$

By induction, we find

$$
\begin{aligned}
a_{n+1} &= c\Big[1 + C\Big(\frac{\alpha(t,s)}{\varepsilon} + \frac{3}{\kappa}\Big)\Big]\sum_{j=0}^{n-1}\Big[c\Big(\frac{\alpha(t,s)}{\varepsilon} + \frac{2}{\kappa}\Big)\Big]^j + c\Big[c\Big(\frac{\alpha(t,s)}{\varepsilon} + \frac{2}{\kappa}\Big)\Big]^n \\
&\leqslant \Big[1 + C\Big(\frac{\alpha(t,s)}{\varepsilon} + \frac{3}{\kappa}\Big)\Big]\frac{c}{1 - c\big(\frac{\alpha(t,s)}{\varepsilon} + \frac{2}{\kappa}\big)} \qquad (4.88) \\
b_{n+1} &= \Big(\frac{3c}{\kappa}\Big)^{n+1} \qquad (4.89)
\end{aligned}
$$

as a possible choice, where we have used the fact that $c(\alpha(t,s)/\varepsilon + 2/\kappa) \leqslant \frac{1}{2}$ by the hypothesis (4.85). Taking the limit $n \to \infty$, and using $c \leqslant \frac{\kappa}{4} \leqslant \frac{1}{4}$, we obtain

$$
Q_t(s) \leqslant Cg(t,s) + \frac{1}{2}\big(1 + 3C\big)\,\mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon} \leqslant 3Cg(t,s). \qquad (4.90)
$$

In order to obtain also a bound on $q_t(s)$, we insert the above bound on $Q_t(s)$ into (4.71), which yields

$$
\begin{aligned}
q_t(s) &\leqslant 2\Big(\frac{h}{\sigma}\Big)^2 \mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon} + 3\kappa C\Big(\frac{h}{\sigma}\Big)^2 \int_s^t \frac{a(u)}{\varepsilon}\,\mathrm{e}^{-\kappa\alpha(u,s)/\varepsilon} g(t,u)\,\mathrm{d}u \\
&\leqslant \Big[2 + 3\kappa C\Big(\frac{1}{\kappa} + \frac{\alpha(t,s)}{\varepsilon}\Big)\Big]\Big(\frac{h}{\sigma}\Big)^2 \mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon} \qquad (4.91)
\end{aligned}
$$

by (4.79). This proves the proposition, and therefore Theorem 2.11, by taking the sum of the above estimates on $q_t(s)$ and $Q_t(s)$. □

4.4 Approach to $x^\star(t)$

We finally turn to the behaviour after the time $\tau = \tau_{\mathcal{D}(\kappa)} > \sqrt{\varepsilon}$, when x_t leaves the set $\mathcal{D}(\kappa)$. By symmetry, we can restrict the analysis to the case $x_\tau = \tilde{x}(\tau)$. Our aim is to prove that with high probability, x_t soon reaches a neighbourhood of $x^\star(t)$. From now on, we always assume $\kappa \in (1/2, 2/3)$.

We start by analysing the solution $x_t^{\det,\tau}$ of the deterministic equation

$$
\varepsilon\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t) \qquad (4.92)
$$

with initial condition $x_\tau^{\det,\tau} = \tilde{x}(\tau)$. Recall from (2.45) that $\tilde{a}(t) = \partial_x f(\tilde{x}(t),t) \leqslant -\eta a(t)$, for any constant η satisfying $\eta \leqslant 2 - 3\kappa - \mathcal{O}_T(1)$.

Proposition 4.11. *For sufficiently small ε and T,*

$$
\begin{aligned}
&\tilde{x}(t) \leqslant x_t^{\det,\tau} \leqslant x^\star(t) \qquad (4.93) \\
&0 \leqslant x^\star(t) - x_t^{\det,\tau} \leqslant C\Big[\frac{\varepsilon}{t^{3/2}} + \big(x^\star(\tau) - \tilde{x}(\tau)\big)\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}\Big] \qquad (4.94) \\
&0 \leqslant x_t^{\det,\sqrt{\varepsilon}} - x_t^{\det,\tau} \leqslant \big(x_\tau^{\det,\sqrt{\varepsilon}} - \tilde{x}(\tau)\big)\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon} \qquad (4.95)
\end{aligned}
$$

for all $t \in [\tau, T]$ and all $\tau \in [\sqrt{\varepsilon}, T]$, where $C > 0$ is a constant depending only on f.

Proof:

1. Whenever $x_t^{\det,\tau} = x^\star(t)$, we have
$$\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}\big(x^\star(t) - x_t^{\det,\tau}\big) = \varepsilon \frac{\mathrm{d}x^\star(t)}{\mathrm{d}t} - f(x^\star(t),t) = \varepsilon \frac{\mathrm{d}x^\star(t)}{\mathrm{d}t} \geqslant 0, \tag{4.96}$$
which shows that $x_t^{\det,\tau}$ can never become larger than $x^\star(t)$. Similarly, whenever $x_t^{\det,\tau} = \tilde{x}(t)$, we get
$$\begin{aligned}\varepsilon \frac{\mathrm{d}}{\mathrm{d}t}\big(x_t^{\det,\tau} - \tilde{x}(t)\big) &= f(\tilde{x}(t),t) - \varepsilon \frac{\mathrm{d}\tilde{x}(t)}{\mathrm{d}t} \\ &= \kappa \tilde{x}(t) t\big[1 + \mathcal{O}_T(1)\big] - \varepsilon \frac{\tilde{x}(t)}{2t}\big[1 + \mathcal{O}_T(1)\big] > 0\end{aligned} \tag{4.97}$$
provided $\kappa > \frac{1}{2}[1 + \mathcal{O}_T(1)]$, which shows that $x_t^{\det,\tau}$ can never become smaller than $\tilde{x}(t)$. This completes the proof of (4.93).
2. We now introduce the difference $y_t^{\det,\tau} = x^\star(t) - x_t^{\det,\tau}$. Using Taylor's formula, one immediately obtains that $y_t^{\det,\tau}$ satisfies the ODE
$$\varepsilon \frac{\mathrm{d}y}{\mathrm{d}t} = a^\star(t) y + b^\star(y,t) + \varepsilon x^{\star\prime}(t) \tag{4.98}$$
where
$$\begin{aligned} a^\star(t) &\leqslant -a_0^\star t \\ 0 \leqslant b^\star(y,t) &\leqslant M^\star \sqrt{t}\, y^2 \\ x^{\star\prime}(t) &\leqslant \frac{K^\star}{\sqrt{t}}, \end{aligned} \tag{4.99}$$
with $a_0^\star = 2[1 + \mathcal{O}_T(1)]$, $M^\star = 3[1 + \mathcal{O}_T(1)]$ and $K^\star = \frac{1}{2}[1 + \mathcal{O}_T(1)]$. We first consider the particular solution $\widehat{y}_t^{\det}$ of (4.98) starting at time $4\sqrt{\varepsilon}$ in $\widehat{y}_{4\sqrt{\varepsilon}}^{\det} = 0$. By (4.96), we know that $\widehat{y}_t^{\det} \geqslant 0$ for all $t \geqslant 4\sqrt{\varepsilon}$. We will use the fact that
$$\begin{aligned}\int_\tau^t \frac{1}{\sqrt{s}}\, \mathrm{e}^{-a_0^\star (t^2 - s^2)/4\varepsilon}\, \mathrm{d}s &\leqslant \int_\tau^t \frac{1}{\sqrt{s}}\, \mathrm{e}^{-a_0^\star t(t-s)/4\varepsilon}\, \mathrm{d}s \\ &\leqslant \frac{4\varepsilon}{a_0^\star t^{3/2}} \int_0^\xi \frac{\mathrm{e}^{-u}}{\sqrt{1 - u/\xi}}\, \mathrm{d}u < c_0 \frac{\varepsilon}{t^{3/2}},\end{aligned} \tag{4.100}$$
where $c_0 = 8/a_0^\star$. We have used the transformation $s = t - 4\varepsilon u/(a_0^\star t)$, introduced $\xi = a_0^\star t^2/4\varepsilon$ and bounded the last integral by 2. We now introduce the first exit time $\hat{\tau} = \inf\{t \geqslant 4\sqrt{\varepsilon} \colon \widehat{y}_t^{\det} \geqslant c_0 \varepsilon t^{-3/2}\}$. For $4\sqrt{\varepsilon} \leqslant t \leqslant \hat{\tau}$, we have
$$a^\star(t) y + b^\star(y,t) \leqslant \Big(-a_0^\star t + M^\star \sqrt{t}\, c_0 \frac{\varepsilon}{t^{3/2}}\Big) y \leqslant -a_0^\star \Big(1 - \frac{c_0 M^\star}{16 a_0^\star}\Big) t y. \tag{4.101}$$
Since $M^\star/(a_0^\star)^2 = \frac{3}{4}[1 + \mathcal{O}_T(1)]$, the term in brackets can be assumed to be larger than $\frac{1}{2}$. Hence (4.98) shows that
$$\varepsilon \frac{\mathrm{d}\widehat{y}^{\det}}{\mathrm{d}t} \leqslant -\frac{a_0^\star}{2} t \widehat{y}^{\det} + \varepsilon \frac{K^\star}{\sqrt{t}}, \tag{4.102}$$

which implies

$$\widehat{y}_t^{\det} \leqslant K^\star \int_\tau^t \frac{e^{-a_0^\star(t^2-s^2)/4\varepsilon}}{\sqrt{s}} \, ds < K^\star c_0 \frac{\varepsilon}{t^{3/2}}. \tag{4.103}$$

Since $K^\star = \frac{1}{2}[1 + \mathcal{O}_T(1)]$, we obtain $\widehat{y}_t^{\det} < c_0 \varepsilon t^{-3/2}$, and thus $\hat{\tau} = \infty$. This shows

$$0 \leqslant \widehat{y}_t^{\det} \leqslant K^\star c_0 \frac{\varepsilon}{t^{3/2}} \qquad \text{for } 4\sqrt{\varepsilon} \leqslant t \leqslant T. \tag{4.104}$$

3. Let $\tau \geqslant \sqrt{\varepsilon}$ and $0 \leqslant y_1 < y_2 \leqslant x^\star(\tau) - \tilde{x}(\tau)$ be given. Let $y_t^{(1)}$ and $y_t^{(2)}$ be solutions of (4.98) with initial conditions $y_\tau^{(1)} = y_1$ and $y_\tau^{(2)} = y_2$, respectively. Then there exists a $\theta \in [0,1]$ such that the difference $z_t = y_t^{(2)} - y_t^{(1)}$ satisfies

$$\varepsilon \frac{dz}{dt} = -\partial_x f(x^\star(t) - y_t^{(1)} - \theta z, t) \leqslant -\eta a(t) z, \tag{4.105}$$

where we have used (4.93) and the definition of η in (2.45). It follows that

$$0 \leqslant y_t^{(2)} - y_t^{(1)} \leqslant (y_2 - y_1) \, e^{-\eta\alpha(t,\tau)/\varepsilon}, \tag{4.106}$$

which proves (4.95) in particular. If $\tau \geqslant 4\sqrt{\varepsilon}$, we can use the relation $x^\star(t) - x_t^{\det,\tau} = \widehat{y}_t^{\det} + (y_t^{\det,\tau} - \widehat{y}_t^{\det})$ to show that

$$x^\star(t) - x_t^{\det,\tau} \leqslant K^\star c_0 \frac{\varepsilon}{t^{3/2}} + \big(x^\star(\tau) - \tilde{x}(\tau)\big) \, e^{-\eta\alpha(t,\tau)/\varepsilon}, \tag{4.107}$$

which proves (4.94) for $\tau \geqslant 4\sqrt{\varepsilon}$. Finally, if $\sqrt{\varepsilon} \leqslant \tau \leqslant 4\sqrt{\varepsilon}$, we can use the fact that $x^\star(t) - x_t^{\det,\tau} \leqslant x^\star(t) - x_t^{\det,4\sqrt{\varepsilon}}$ to prove that (4.94) holds for some constant $C > 0$. □

Let us now consider the process $y_t = y_t^\tau = x_t - x_t^{\det,\tau}$, starting at time τ in $y_\tau = 0$, which describes the deviation due to noise from the deterministic solution $x_t^{\det,\tau}$. It satisfies the SDE

$$dy_t = \frac{1}{\varepsilon}\big[a^\tau(t) y_t + b^\tau(y_t, t)\big] \, dt + \frac{\sigma}{\sqrt{\varepsilon}} \, dW_t, \tag{4.108}$$

where we have introduced

$$\begin{aligned} a^\tau(t) &= \partial_x f(x_t^{\det,\tau}, t) \\ b^\tau(y,t) &= f(x_t^{\det,\tau} + y, t) - f(x_t^{\det,\tau}) - a^\tau(t) y. \end{aligned} \tag{4.109}$$

The following bounds are direct consequences of Taylor's formula and Proposition 4.11:

$$a^\star(t) \leqslant a^\tau(t) \leqslant \tilde{a}(t) \tag{4.110}$$

$$a^\tau(t) = a^\star(t) + \mathcal{O}\Big(\frac{\varepsilon}{t}\Big) + \mathcal{O}(t \, e^{-\eta\alpha(t,\tau)/\varepsilon}) \tag{4.111}$$

$$(a^\tau)'(t) = \mathcal{O}\Big(1 + \frac{t^2}{\varepsilon} \, e^{-\eta\alpha(t,\tau)/\varepsilon}\Big) \tag{4.112}$$

$$|b^\tau(y,t)| \leqslant 3My^2\big(x^\star(t) + |y|\big), \qquad \text{valid for } x^\star(t) + |y| \leqslant d. \tag{4.113}$$

For comparison, we will also consider the linear SDE

$$dy_t^0 = \frac{1}{\varepsilon} a^\tau(t) y_t^0 \, dt + \frac{\sigma}{\sqrt{\varepsilon}} \, dW_t. \tag{4.114}$$

Let $\alpha^\tau(t,s) = \int_s^t a^\tau(u)\,\mathrm{d}u$ and denote by

$$v^\tau(t) = \frac{\sigma^2}{\varepsilon}\int_\tau^t \mathrm{e}^{2\alpha^\tau(t,s)/\varepsilon}\,\mathrm{d}s \tag{4.115}$$

the variance of y_t^0. Again we introduce and investigate a function

$$\zeta^\tau(t) = \frac{1}{2|\tilde{a}(\tau)|}\,\mathrm{e}^{2\alpha^\tau(t,\tau)/\varepsilon} + \frac{1}{\varepsilon}\int_\tau^t \mathrm{e}^{2\alpha^\tau(t,s)/\varepsilon}\,\mathrm{d}s. \tag{4.116}$$

Lemma 4.12. *The function $\zeta^\tau(t)$ satisfies the following relations for $\tau \leqslant t \leqslant T$:*

$$\zeta^\tau(t) = \frac{1}{2|\tilde{a}(t)|} + \mathcal{O}\Big(\frac{\varepsilon}{t^3}\Big) + \mathcal{O}\Big(\frac{1}{t}\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}\Big) \tag{4.117}$$

$$\frac{1}{2|a^\star(t)|} \leqslant \zeta^\tau(t) \leqslant \frac{1}{2|\tilde{a}(\tau)|} \tag{4.118}$$

$$(\zeta^\tau)'(t) \leqslant \frac{1}{\varepsilon}. \tag{4.119}$$

Proof:

1. By integration by parts, we find

$$\zeta^\tau(t) = \frac{1}{2|\tilde{a}(t)|} - \frac{1}{2}\int_\tau^t \frac{(a^\tau)'(s)}{a^\tau(s)^2}\,\mathrm{e}^{2\alpha^\tau(t,s)/\varepsilon}\,\mathrm{d}s. \tag{4.120}$$

The relation $|a^\tau(s)| \geqslant |\tilde{a}(s)| \geqslant \eta|a(s)|$ together with (4.112) yields

$$\bigg|\int_\tau^t \frac{(a^\tau)'(s)}{a^\tau(s)^2}\,\mathrm{e}^{2\alpha^\tau(t,s)/\varepsilon}\,\mathrm{d}s\bigg| \leqslant \mathit{const}\int_\tau^t \Big(\frac{1}{s^2} + \frac{1}{\varepsilon}\,\mathrm{e}^{-\eta\alpha(s,\tau)/\varepsilon}\Big)\,\mathrm{e}^{-2\eta\alpha(t,s)/\varepsilon}\,\mathrm{d}s. \tag{4.121}$$

The second term in brackets gives a contribution of order $\frac{1}{t}\,\mathrm{e}^{-\eta\alpha(t,\tau)/\varepsilon}$. In order to estimate the contribution of the first term, we perform the change of variables $u = \eta(t^2 - s^2)/2\varepsilon$, thereby obtaining

$$\int_\tau^t \frac{1}{s^2}\,\mathrm{e}^{-\eta(t^2-s^2)/2\varepsilon}\,\mathrm{d}s = \frac{\varepsilon}{\eta t^3}\int_0^{\xi-\xi_0} \frac{\mathrm{e}^{-u}}{(1-u/\xi)^{3/2}}\,\mathrm{d}u \leqslant \frac{\varepsilon}{\eta t^3}\Big[2^{3/2} + 2\frac{\xi^{3/2}\,\mathrm{e}^{-\xi/2}}{\sqrt{\xi_0}}\Big], \tag{4.122}$$

where $\xi = \eta t^2/2\varepsilon$ and $\xi_0 = \eta\tau^2/2\varepsilon$. The last inequality is obtained by splitting the integral at $\xi/2$. Using the fact that $t^3\,\mathrm{e}^{-\eta t^2/4\varepsilon} \leqslant (6\varepsilon/\eta)^{3/2}\,\mathrm{e}^{-3/2}$ for all $t \geqslant 0$, we reach the conclusion that this integral is bounded by a constant times ε/t^3, which completes the proof of (4.117).

2. We now use the fact that $\zeta^\tau(t)$ solves the ODE

$$\frac{\mathrm{d}\zeta^\tau}{\mathrm{d}t} = \frac{1}{\varepsilon}\big(2a^\tau(t)\zeta^\tau + 1\big), \qquad \zeta^\tau(\tau) = \frac{1}{2|\tilde{a}(\tau)|}. \tag{4.123}$$

Then, (4.119) is an immediate consequence of this relation, and (4.118) is obtained from the fact that

$$\frac{\mathrm{d}\zeta^\tau(t)}{\mathrm{d}t} = \frac{1}{\varepsilon}\Big(-\frac{|a^\tau(t)|}{|\tilde{a}(\tau)|} + 1\Big) \leqslant 0, \tag{4.124}$$

whenever $\zeta^\tau(t) = 1/2|\tilde{a}(\tau)|$, and

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\zeta^\tau(t) - \frac{1}{2|a^\star(t)|}\Big) = \frac{1}{\varepsilon}\Big(-\frac{|a^\tau(t)|}{|a^\star(t)|} + 1\Big) - \frac{a^{\star\prime}(t)}{2a^\star(t)^2} \geqslant 0, \tag{4.125}$$

whenever $\zeta^\tau(t) = 1/2|a^\star(\tau)|$. Here we used (4.110) and the monotonicity of $\tilde{a}(t)$ for small t. □

We note that Lemma 4.12 and the bounds (4.110) on a^τ imply the existence of constants $c_+ \geqslant c_- > 0$, depending only on f and T, such that

$$\frac{c_-}{t} \leqslant \zeta^\tau(t) \leqslant \frac{c_+}{t} \qquad \forall t \in [\tau, T]. \tag{4.126}$$

We can now easily prove that y_t^0 remains in a strip of width $h\sqrt{\zeta^\tau}$ with high probability, in much the same way as in Proposition 3.4.

Proposition 4.13. *For sufficiently small T and ε, and all $t \in [\tau, T]$,*

$$\mathbb{P}^{\tau,0}\Bigl\{\sup_{\tau\leqslant s\leqslant t} \frac{|y_s^0|}{\sqrt{\zeta^\tau(s)}} \geqslant h\Bigr\} \leqslant C^\tau(t,\varepsilon)\exp\Bigl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\bigl[1-r(\varepsilon)\bigr]\Bigr\}, \tag{4.127}$$

where $r(\varepsilon) = \mathcal{O}(\varepsilon)$ and

$$C^\tau(t,\varepsilon) = \frac{|\alpha^\tau(t,\tau)|}{\varepsilon^2} + 2. \tag{4.128}$$

Proof: Let $K = \lceil |\alpha^\tau(t,\tau)|/2\varepsilon^2 \rceil$ and define a partition $\tau = u_0 < \dots < u_K = t$ of $[\tau, t]$ by

$$|\alpha^\tau(u_k,\tau)| = 2\varepsilon^2 k, \qquad k = 1, \dots, K-1. \tag{4.129}$$

Since $a^\tau(s) \leqslant \tilde{a}(s) \leqslant -\eta s/2$, we obtain $u_k - u_{k-1} \leqslant 4\varepsilon^2/(\eta u_{k-1})$ for all k. Now we can proceed as in the proof of Proposition 3.4. □

We can now compare the solutions of the linear and the nonlinear equation. To do so, we define the events

$$\Omega_t(h) = \bigl\{\omega\colon |y_s^\tau| < h\sqrt{\zeta^\tau(s)}\ \forall s \in [\tau, t]\bigr\} \tag{4.130}$$

$$\Omega_t^0(h) = \bigl\{\omega\colon |y_s^0| < h\sqrt{\zeta^\tau(s)}\ \forall s \in [\tau, t]\bigr\}. \tag{4.131}$$

The following proposition shows that y_t^τ and y_t^0 differ only slightly.

Proposition 4.14. *Let $\gamma = 1 \vee 48M(2+\sqrt{c_+})c_+^2/\sqrt{c_-}$ and assume $h < \tau/\gamma$ as well as $h \leqslant [d - x^\star(t)]\sqrt{\tau}/(2\sqrt{c_+})$. Then*

$$\Omega_t(h) \overset{\text{a.s.}}{\subset} \Omega_t^0\Bigl(\Bigl[1+\gamma\frac{h}{\tau}\Bigr]h\Bigr) \tag{4.132}$$

$$\Omega_t^0(h) \overset{\text{a.s.}}{\subset} \Omega_t\Bigl(\Bigl[1+\gamma\frac{h}{\tau}\Bigr]h\Bigr). \tag{4.133}$$

The proof is similar to the one of the corresponding result in the case of nonbifurcating equilibria, cf. Proposition 3.7.

Now, the following corollary is a direct consequence of the two preceding propositions.

Corollary 4.15. *There exists h_0 such that if $h < h_0\tau$, then*

$$\mathbb{P}^{\tau,\tilde{x}(\tau)}\Biggl\{\sup_{\tau\leqslant s\leqslant t} \frac{|x_s - x_s^{\det,\tau}|}{\sqrt{\zeta^\tau(s)}} > h\Biggr\} \leqslant C^\tau(t,\varepsilon)\exp\Bigl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\Bigl[1-\mathcal{O}(\varepsilon)-\mathcal{O}\Bigl(\frac{h}{\tau}\Bigr)\Bigr]\Bigr\}, \tag{4.134}$$

where $C^\tau(t,\varepsilon)$ is given by (4.128).

Appendix

The following lemma provides an estimate on expectation values, as used in Subsection 4.3. It is based on integration by parts.

Lemma A.1. *Let $\tau \geqslant s_0$ be a random variable satisfying $F_\tau(s) = \mathbb{P}\{\tau < s\} \geqslant G(s)$ for some continuously differentiable function G. Then*

$$\mathbb{E}\bigl\{1_{[s_0,t)}(\tau)g(\tau)\bigr\} \leqslant g(t)\bigl[F_\tau(t) - G(t)\bigr] + \int_{s_0}^{t} g(s)G'(s)\,\mathrm{d}s \tag{A.1}$$

holds for all $t > s_0$ and all functions $0 \leqslant g \leqslant 1$ satisfying the two conditions

- *there exists an $s_1 \in (s_0, \infty]$ such that g is continuously differentiable and increasing on (s_0, s_1);*
- *$g(s) = 1$ for all $s \geqslant s_1$.*

Proof: First note that for all $t \leqslant s_1$,

$$\begin{aligned}\int_{s_0}^{t} g'(s)\mathbb{P}\{\tau \geqslant s\}\,\mathrm{d}s &= \mathbb{E}\Bigl\{\int_{s_0}^{t\wedge\tau} g'(s)\,\mathrm{d}s\Bigr\} \\ &= \mathbb{E}\{g(t\wedge\tau)\} - g(s_0) \\ &= \mathbb{E}\{1_{[s_0,t)}(\tau)g(\tau)\} + g(t)\mathbb{P}\{\tau \geqslant t\} - g(s_0)\end{aligned} \tag{A.2}$$

which implies, by integration by parts,

$$\begin{aligned}\mathbb{E}\{1_{[s_0,t)}(\tau)g(\tau)\} &= \int_{s_0}^{t} g'(s)\bigl[1 - F_\tau(s)\bigr]\,\mathrm{d}s - g(t)\bigl[1 - F_\tau(t)\bigr] + g(s_0) \\ &\leqslant \int_{s_0}^{t} g(s)G'(s)\,\mathrm{d}s + g(t)\bigl[F_\tau(t) - G(t)\bigr],\end{aligned} \tag{A.3}$$

where we have used $F_\tau(s) \geqslant G(s)$ and $G(s_0) \leqslant F(s_0) = 0$. This proves the assertion in the case $t \leqslant s_1$. In the case $t > s_1$, we have

$$\begin{aligned}\mathbb{E}\{1_{[s_0,t)}(\tau)g(\tau)\} &= \mathbb{E}\{1_{[s_0,s_1)}(\tau)g(\tau)\} + \mathbb{P}\{\tau \in [s_1,t)\} \\ &\leqslant \int_{s_0}^{s_1} g(s)G'(s)\,\mathrm{d}s + g(s_1)\bigl[F_\tau(s_1) - G(s_1)\bigr] + \bigl[F_\tau(t) - F_\tau(s_1)\bigr] \\ &= \int_{s_0}^{t} g(s)G'(s)\,\mathrm{d}s - \bigl[G(t) - G(s_1)\bigr] + \bigl[F_\tau(t) - G(s_1)\bigr],\end{aligned} \tag{A.4}$$

where we have used that $g(s) = 1$ holds for all $s \in [s_1, t]$. This proves the assertion for $t > s_1$. □

References

[Ar] L. Arnold, *Random Dynamical Systems* (Springer-Verlag, Berlin, 1998).

[Ben] E. Benoît (Ed.), *Dynamic Bifurcations, Proceedings, Luminy 1990* (Springer-Verlag, Lecture Notes in Mathematics 1493, Berlin, 1991).

[Ber] N. Berglund, *Adiabatic Dynamical Systems and Hysteresis*, Thesis EPFL no 1800 (1998). Available at `http://dpwww.epfl.ch/instituts/ipt/berglund/these.html`

[BK] N. Berglund, H. Kunz, *Chaotic hysteresis in an adiabatically oscillating double well*, Phys. Rev. Letters **78**:1692–1694 (1997). N. Berglund, H. Kunz, *Memory effects and scaling laws in slowly driven systems*, J. Phys. A **32**:15–39 (1999).

[CF94] H. Crauel, F. Flandoli, *Attractors for random dynamical systems*, Probab. Theory Related Fields **100**:365–393 (1994).

[CF98] H. Crauel, F. Flandoli, *Additive noise destroys a pitchfork bifurcation*, J. Dynam. Differential Equations **10**:259–274 (1998).

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

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

[Ga] G. Gaeta, *Dynamical bifurcation with noise*, Int. J. Theoret. Phys. **34**:595–603 (1995).

[Gr] I. S. Gradšteĭn, *Applications 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**:263–286 (1953).

[GH] J. Guckenheimer, P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Springer-Verlag, New York, 1983).

[IJ] G. Iooss, D. D. Joseph, *Elementary Stability and Bifurcation Theory* (Springer-Verlag, New York, 1980).

[JL] K. M. Jansons, G. D. Lythe, *Stochastic calculus: Application to dynamic bifurcations and threshold crossings*, J. Stat. Phys. **90**:227–251 (1998).

[Ku] R. Kuske, *Probability densities for noisy delay bifurcations*, J. Stat. Phys. **96**:797–816 (1999).

[ME] P. Mandel, T. Erneux, *Laser Lorenz equations with a time-dependent parameter*, Phys. Rev. Letters **53**:1818–1820 (1984).

[Ne] A. I. Neishtadt, *Persistence of stability loss for dynamical bifurcations I, II*, Diff. Equ. **23**:1385–1391 (1987). Diff. Equ. **24**:171–176 (1988).

[Schm] B. Schmalfuß, *Invariant attracting sets of nonlinear stochastic differential equations*, Math. Res. **54**:217–228 (1989).

[Sh] M. A. Shishkova, *Examination of one system of differential equations with a small parameter in highest derivatives*, Dokl. Akad. Nauk SSSR **209**:576–579 (1973). [English transl.: Soviet Math. Dokl. **14**:384–387 (1973)].

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

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

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

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

Nils Berglund
Georgia Institute of Technology
Atlanta, GA 30332-0430, USA
and
Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstraße 39, 10117 Berlin, Germany
current address:
Department of Mathematics, ETH Zürich
ETH Zentrum, 8092 Zürich, Switzerland
E-mail address: berglund@math.ethz.ch

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