Markov Processes Relat. Fields **16**, 549–598 (2010)

Markov Processes and Related Fields

The Eyring – Kramers Law for Potentials with Nonquadratic Saddles

N. Berglund [1] **and B. Gentz** [2]

[1] Université d'Orléans, Laboratoire Mapmo, CNRS, UMR 6628, Fédération Denis Poisson, FR 2964, Bâtiment de Mathématiques, B.P. 6759, 45067 Orléans Cedex 2, France. E-mail: nils.berglund@univ-orleans.fr
[2] Faculty of Mathematics, University of Bielefeld, P.O. Box 10 01 31, 33501 Bielefeld, Germany. E-mail: gentz@math.uni-bielefeld.de

Received April 7, 2009, revised October 29, 2009

Abstract. The Eyring – Kramers law describes the mean transition time of an overdamped Brownian particle between local minima in a potential landscape. In the weak-noise limit, the transition time is to leading order exponential in the potential difference to overcome. This exponential is corrected by a prefactor which depends on the principal curvatures of the potential at the starting minimum and at the highest saddle crossed by an optimal transition path. The Eyring – Kramers law, however, does not hold whenever one or more of these principal curvatures vanishes, since it would predict a vanishing or infinite transition time. We derive the correct prefactor up to multiplicative errors that tend to one in the zero-noise limit. As an illustration, we discuss the case of a symmetric pitchfork bifurcation, in which the prefactor can be expressed in terms of modified Bessel functions, as well as bifurcations with two vanishing eigenvalues. The corresponding transition times are studied in a full neighbourhood of the bifurcation point. These results extend work by Bovier, Eckhoff, Gayrard and Klein [2], who rigorously analysed the case of quadratic saddles, using methods from potential theory.

Keywords: stochastic differential equations, exit problem, transition times, most probable transition path, large deviations, Wentzell – Freidlin theory, metastability, potential theory, capacities, subexponential asymptotics, pitchfork bifurcation

AMS Subject Classification: Primary 60J45, 31C15, Secondary 60J60, 37H20

Contents

1. Introduction

Consider the stochastic differential equation

$$\mathrm{d}x_t = -\nabla V(x_t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t, \tag{1.1}$$

where $V : \mathbb{R}^d \to \mathbb{R}$ is a confining potential. The Eyring−Kramers law [6, 15] describes the expected transition time τ between potential minima in the small-noise limit $\varepsilon \to 0$. In the one-dimensional case ($d = 1$), it has the following form. Assume x and y are quadratic local minima of V, separated by a unique quadratic local maximum z. Then the expected transition time from x to y satisfies

$$\mathbb{E}^x\{\tau\} \simeq \frac{2\pi}{\sqrt{V''(x)|V''(z)|}}\,\mathrm{e}^{[V(z)-V(x)]/\varepsilon}\,. \tag{1.2}$$

In the multidimensional case ($d \geqslant 2$), assume the local minima are separated by a unique saddle z, which is such that the Hessian $\nabla^2 V(z)$ admits a single negative eigenvalue $\lambda_1(z)$, while all other eigenvalues are strictly positive. Then the analogue of (1.2) reads

$$\mathbb{E}^x\{\tau\} \simeq \frac{2\pi}{|\lambda_1(z)|}\sqrt{\frac{\det(\nabla^2 V(z))}{\det(\nabla^2 V(x))}}\,\mathrm{e}^{[V(z)-V(x)]/\varepsilon}\,. \tag{1.3}$$

This expression has been generalised to situations where there are several alternative saddles allowing to go from x to y, and to potentials with more than two minima.

A long time has elapsed between the first presentation of the formula (1.3) by Eyring [6] and Kramers [15] and its rigorous mathematical proof (including a precise definition of what the symbol "$\simeq$" in (1.3) actually means). The exponential asymptotics were proved to be correct by Wentzell and Freidlin in the early Seventies, using the theory of large deviations [7, 17, 18]. While being very flexible, and allowing to study more general than gradient systems like (1.1), large deviations do not allow to obtain the prefactor of the transition time. An alternative approach is based on the fact that mean transition times obey certain elliptic partial differential equations, whose solutions can be approximated by WKB-theory (for a recent survey of these methods, see [14]). This approach provides formal asymptotic series expansions in ε, whose justification is, however, a difficult problem of analysis. A framework for such a rigorous justification is provided by microlocal analysis, which was primarily developed by Helffer and Sjöstrand to solve quantum mechanical tunnelling problems in the semiclassical limit [10–13]. Unfortunately, it turns out that when translated into terms of semiclassical analysis, the problem of proving the Eyring – Kramers formula becomes a particularly intricate one, known as "tunnelling through non-resonant wells". The first mathematically rigorous proof of (1.3) in arbitrary dimension (and its generalisations to more than two wells) was obtained by Bovier, Eckhoff, Gayrard and Klein [2], using a different approach based on potential theory and a variational principle. In [2], the Eyring – Kramers law is shown to hold with $a \simeq b$ meaning $a = b(1 + \mathcal{O}(\varepsilon^{1/2}|\log\varepsilon|))$. Finally, a full asymptotic expansion of the prefactor in powers of ε was proved to hold in [8, 9], using again analytical methods.

In this work, we are concerned with the case where the determinant of one of the Hessian matrices vanishes. In such a case, the expression (1.3) either diverges or goes to zero, which is obviously absurd. It seems reasonable (as has been pointed out, e.g., in [16]) that one has to take into account higher-order terms of the Taylor expansion of the potential at the stationary points when estimating the transition time. Of course, cases with degenerate Hessian are in a sense not generic, so why should we care about this situation at all? The answer is that as soon as the potential depends on a parameter, degenerate stationary points are bound to occur, most notably at bifurcation points, i.e., where the number of saddles varies as the parameter changes. See, for instance, [3, 4] for an analysis of a naturally arising parameter-dependent system displaying a series of symmetry-breaking bifurcations. For this particular system, an analysis of the subexponential asymptotics of metastable transition times in the synchronisation regime has been given recently in [1], with a careful control of the dimension-dependence of the error terms.

In order to study sharp asymptotics of expected transition times, we rely on the potential-theoretic approach developed in [2, 5]. In particular, the expected transition time can be expressed in terms of so-called Newtonian capacities between sets, which can in turn be estimated by a variational principle involving Dirichlet forms. The main new aspect of the present work is that we estimate capacities in cases involving nonquadratic saddles.

In the non-degenerate case, saddles are easy to define: they are stationary points at which the Hessian has exactly one strictly negative eigenvalue, all other eigenvalues being strictly positive. When the determinant of the Hessian vanishes, the situation is not so simple, since the nature of the stationary point depends on higher-order terms in the Taylor expansion. We thus start, in Section 2, by defining and classifying saddles in degenerate cases. In Section 3, we estimate capacities for the most generic cases, which then allows us to derive expressions for the expected transition times. In Section 4, we extend these results to a number of bifurcation scenarios arising in typical applications, that is, we consider parameter-dependent potentials for parameter values in a full neighbourhood of a critical parameter value yielding non-quadratic saddles. Section 5 contains the proofs of the main results.

2. Classification of nonquadratic saddles

We consider a continuous, confining potential $V : \mathbb{R}^d \to \mathbb{R}$, bounded below by some $a_0 \in \mathbb{R}$ and having exponentially tight level sets, that is,

$$\int_{\{x\in\mathbb{R}^d \colon V(x)\geqslant a\}} \mathrm{e}^{-V(x)/\varepsilon}\,\mathrm{d}x \leqslant C(a)\,\mathrm{e}^{-a/\varepsilon} \quad \forall a \geq a_0, \tag{2.1}$$

with $C(a)$ bounded above and uniform in $\varepsilon \leqslant 1$. We start by giving a topological definition of saddles, before classifying saddles for sufficiently differentiable potentials V.

2.1. Topological definition of saddles

We start by introducing the notion of a gate between two sets A and B. Roughly speaking, a *gate* is a set that cannot be avoided by those paths going from A to B which stay as low as possible in the potential landscape. *Saddles* will then be defined as particular points in gates.

It is useful to introduce some terminology and notations:

- For $x, y \in \mathbb{R}^d$, we denote by $\gamma : x \to y$ a *path* from x to y, that is, a continuous function $\gamma\colon [0,1] \to \mathbb{R}^d$ such that $\gamma(0) = x$ and $\gamma(1) = y$.
- The *communication height* between x and y is the highest potential value no path leading from x to y can avoid reaching, even when staying as low as possible, i.e.,
$$\overline{V}(x,y) = \inf_{\gamma\colon x\to y} \sup_{t\in[0,1]} V(\gamma(t)). \tag{2.2}$$
Note that $\overline{V}(x,y) \geqslant V(x) \vee V(y)$, with equality holding, for instance, in cases where x and y are "on the same side of a mountain slope".
- The communication height between two sets $A, B \subset \mathbb{R}^d$ is given by
$$\overline{V}(A,B) = \inf_{x\in A, y\in B} \overline{V}(x,y). \tag{2.3}$$
We denote by $\mathcal{G}(A,B) = \{z \in \mathbb{R}^d \colon V(z) = \overline{V}(A,B)\}$ the level set of $\overline{V}(A,B)$.
- The *set of minimal paths* from A to B is
$$\mathcal{P}(A,B) = \left\{\gamma\colon x \to y \,\middle|\, x \in A, y \in B, \sup_{t\in[0,1]} V(\gamma(t)) = \overline{V}(A,B)\right\}. \tag{2.4}$$

The following definition is taken from [2].

Definition 2.1. A *gate* $G(A,B)$ is a minimal subset of $\mathcal{G}(A,B)$ such that all minimal paths $\gamma \in \mathcal{P}(A,B)$ must intersect $G(A,B)$.

Let us consider some examples in dimension $d = 2$ (Figure 1):

- In uninteresting cases, e.g. for A and B on the same side of a slope, the gate $G(A,B)$ is a subset of $A \cup B$. We will not be concerned with such cases.
- If on the way from A to B, one has to cross one "mountain pass" z which is higher than all other passes, then $G(A,B) = \{z\}$ (Figure 1a).
- If there are several passes at communication height $\overline{V}(A,B)$ between A and B, between which one can choose, then the gate $G(A,B)$ is the union of these passes (Figure 1b).

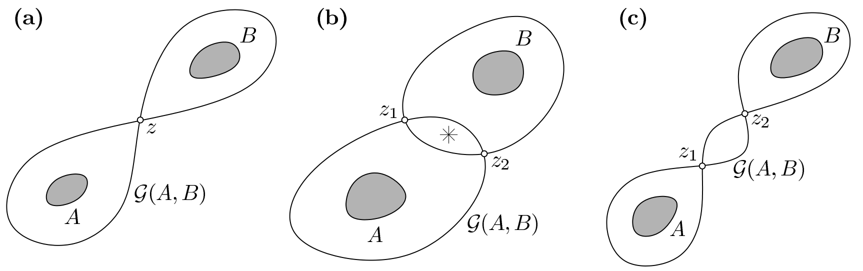

Figure 1. Examples of potentials and gates. **(a)** $G(A,B) = \{z\}$. **(b)** $G(A,B) = \{z_1, z_2\}$. **(c)** $G(A,B) = \{z_1\}$ or $\{z_2\}$. Here curves show level lines, shaded areas indicate potential wells and the star marks a potential maximum.

- If when going from A to B, one has to cross several passes in a row, all at communication height $\overline{V}(A,B)$, then the gate $G(A,B)$ is not uniquely defined: any of the passes will form a gate (Figure 1c).
- If A and B are separated by a ridge of constant altitude $\overline{V}(A,B)$, then the whole ridge is the gate $G(A,B)$.
- If the potential contains a flat part separating A from B, at height $\overline{V}(A,B)$, then any curve in this part separating the two sets is a gate.

We now proceed to defining saddles as particular cases of isolated points in gates. However, the definition should be independent of the choice of sets A and B. In order to do this, we start by introducing notions of valleys (cf. Figure 2):

- The *closed valley* of a point $x \in \mathbb{R}^d$ is the set

$$\mathcal{CV}(x) = \bigl\{y \in \mathbb{R}^d \colon \overline{V}(y,x) = V(x)\bigr\}. \tag{2.5}$$

  It is straightforward to check that $\mathcal{CV}(x)$ is closed and path-connected.

- The *open valley* of a point $x \in \mathbb{R}^d$ is the set

$$\mathcal{OV}(x) = \bigl\{y \in \mathcal{CV}(x) \colon V(y) < V(x)\bigr\} . \tag{2.6}$$

  It is again easy to check that $\mathcal{OV}(x)$ is open. Note however that if the potential contains horizontal parts, then $\mathcal{CV}(x)$ need not be the closure of $\mathcal{OV}(x)$ (Figure 2c). Also note that $\mathcal{OV}(x)$ need not be path-connected (Figure 2b). We will use this fact to define a saddle.

Let $\mathcal{B}_\varepsilon(x) = \{y \in \mathbb{R}^d \colon \|y-x\|_2 < \varepsilon\}$ denote the open ball of radius ε, centred in x.

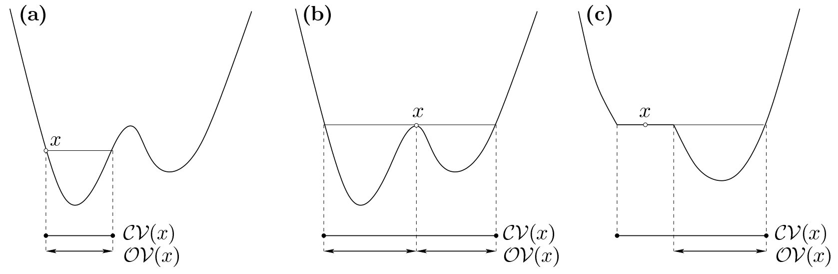

Figure 2. Examples of potentials and valleys. In case **(b)**, x is a saddle.

Definition 2.2. A *saddle* is a point $z \in \mathbb{R}^d$ such that there exists $\varepsilon > 0$ for which

1. $\mathcal{OV}(z) \cap \mathcal{B}_\varepsilon(z)$ is non-empty and not path-connected.
2. $(\mathcal{OV}(z) \cup \{z\}) \cap \mathcal{B}_\varepsilon(z)$ is path-connected.

The link between saddles and gates is made clear by the following two results.

Proposition 2.1. *Let z be a saddle. Assume $\mathcal{OV}(z)$ is not path-connected*[1]*, and let A and B belong to different path-connected components of $\mathcal{OV}(z) \cap \mathcal{B}_\varepsilon(z)$. Then $z \in G(A, B)$.*

Proof. Choose points $x \in A$ and $y \in B$ and a path $\gamma \colon A \to B$. Since A and B belong to different path-connected components of $\mathcal{OV}(z)$, the path γ must leave $\mathcal{OV}(z)$, which implies $\sup_{t \in [0,1]} V(\gamma(t)) \geqslant V(z)$. Since $(\mathcal{OV}(z) \cup \{z\}) \cap \mathcal{B}_\varepsilon(z)$ is path-connected, we can find a path $\gamma \colon A \to B$ staying all the time in this set, and thus for this path, $\sup_{t \in [0,1]} V(\gamma(t)) = V(z)$. As a consequence, the communication height $\overline{V}(x, y)$ equals $V(z)$, i.e., γ belongs to the set $\mathcal{P}(A, B)$ of minimal paths. Since $\mathcal{OV}(z)$ is not path-connected, we have found a path $\gamma \in \mathcal{P}(A, B)$ which must contain z, and thus $z \in G(A, B)$. □

Proposition 2.2. *Let A and B be two disjoint sets, and let $z \in G(A, B)$. Assume that z is isolated in the sense that there exists $\varepsilon > 0$ such that $\mathcal{B}^*_\varepsilon(z) := \mathcal{B}_\varepsilon(z) \setminus \{z\}$ is disjoint from the union of all gates $G(A, B)$ between A and B. Then z is a saddle.*

Proof. Consider the set

$$D = \bigcup_{\gamma \in \mathcal{P}(A,B)} \bigcup_{t \in [0,1]} \gamma(t) \tag{2.7}$$

[1]We need to make this assumption globally, in order to rule out situations where z is not the lowest saddle between two domains.

of all points contained in minimal paths from A to B. We claim that $D = \mathcal{CV}(z)$.

On one hand, if $x \in D$ then there exists a minimal path from A to B containing x. We follow this path backwards from x to A. Then there is a (possibly different) minimal path leading from the first path's endpoint in A through z to B. By gluing together these paths, we obtain a minimal path connecting x and z. This path never exceeds the potential value $V(z)$, which proves $\overline{V}(x,z) = V(z)$. Thus $x \in \mathcal{CV}(z)$, and $D \subset \mathcal{CV}(z)$ follows.

On the other hand, pick $y \in \mathcal{CV}(z)$. Then there is a path $\gamma_1 \colon y \to z$ along which the potential does not exceed $V(z)$. Inserting this path (twice, going back and forth) in a minimal path $\gamma \in \mathcal{P}(A,B)$ containing z, we get another minimal path from A to B, containing y. This proves $y \in D$, and thus the inverse inclusion $\mathcal{CV}(z) \subset D$.

Now pick $x \in A$ and $y \in B$. There must exist a minimal path $\gamma \colon x \to y$, containing z, with the property that V is strictly smaller than $V(z)$ on $\gamma([0,1]) \cap \mathcal{B}^*_\varepsilon(z)$, since otherwise we would contradict the assumption that z be isolated. We can thus pick x' on γ between x and z and y' on γ between z and y such that $V(x') < V(z)$ and $V(y') < V(z)$. Hence we have $x', y' \in \mathcal{OV}(z)$ and any minimal path from x' to y' staying in $\mathcal{B}_\varepsilon(z)$ has to cross $z \notin \mathcal{OV}(z)$. This shows that $\mathcal{OV}(z) \cap \mathcal{B}_\varepsilon(z)$ is non-empty and not path-connected. Finally, take any $x, y \in \mathcal{OV}(z) \cap \mathcal{B}_\varepsilon(z) \subset D$. Then we can connect them by a path $\gamma \ni z$, and making ε small enough we may assume that V is strictly smaller than $V(z)$ on $\gamma([0,1]) \cap \mathcal{B}^*_\varepsilon(z)$, i.e., $\gamma([0,1]) \setminus \{z\} \subset \mathcal{OV}(z)$. This proves that $(\mathcal{OV}(z) \cup \{z\}) \cap \mathcal{B}_\varepsilon(z)$ is path-connected. □

2.2. Classification of saddles for differentiable potentials

Let us show that for sufficiently smooth potentials, our definition of saddles is consistent with the usual definition of nondegenerate saddles. Then we will start classifying degenerate saddles.

Proposition 2.3. *Let V be of class $\mathcal{C}^1$, and let z be a saddle. Then z is a stationary point of V, i.e., $\nabla V(z) = 0$.*

Proof. Suppose, to the contrary, that $\nabla V(z) \neq 0$. We may assume $z = 0$ and $V(z) = 0$. Choose local coordinates in which $\nabla V(0) = (a, 0, \dots, 0)$ with $a > 0$. By the implicit-function theorem, there exists a differentiable function $h : \mathbb{R}^{d-1} \to \mathbb{R}^d$ such that all solutions of the equation $V(x) = 0$ in a small ball $\mathcal{B}_\varepsilon(0)$ are of the form $x_1 = h(x_2, \dots, x_d)$. Furthermore, $V(\varepsilon, 0, \dots, 0) = a\varepsilon + o(\varepsilon)$ is positive for $\varepsilon > 0$ and negative for $\varepsilon < 0$. By continuity, $V(x)$ is positive for $x_1 > h(x_2, \dots, x_d)$ and negative for $x_1 < h(x_2, \dots, x_d)$, showing that $\mathcal{OV}(z) \cap \mathcal{B}_\varepsilon(0)$ is path-connected. Hence z is not a saddle. □

Proposition 2.4. *Assume V is of class $\mathcal{C}^2$, and let z be a saddle. Then*

1. *The Hessian $\nabla^2 V(z)$ has at least one eigenvalue smaller or equal than 0.*

2. *The Hessian $\nabla^2 V(z)$ has at most one eigenvalue strictly smaller than* 0.

Proof. Denote the eigenvalues of $\nabla^2 V(z)$ by $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_d$. We may again assume that $z = 0$ and $V(0) = 0$, and choose a basis in which the Hessian is diagonal. Then

$$V(x) = \frac{1}{2} \sum_{i=1}^{d} \lambda_i x_i^2 + o(\|x\|_2^2). \tag{2.8}$$

1. Assume, to the contrary that $\lambda_1 > 0$. Then $V > 0$ near $z = 0$, so that $\mathcal{OV}(z) = \emptyset$, and $z = 0$ is not a saddle.

2. Suppose, to the contrary, that $\lambda_1 \leqslant \lambda_2 < 0$, and fix a small $\delta > 0$. Since

$$V(x) = -\frac{1}{2}|\lambda_1| x_1^2 - \frac{1}{2}|\lambda_2| x_2^2 + \frac{1}{2} \sum_{i=3}^{d} \lambda_i x_i^2 + o(\|x\|_2^2), \tag{2.9}$$

we can find an $\varepsilon = \varepsilon(\delta) \in (0, \delta)$ such that for any fixed $(x_3, \ldots, x_d)$ of length less than ε, the set $\{(x_1, x_2) \colon x_1^2 + x_2^2 < \delta^2, V(x) < 0\}$ is path-connected (topologically, it is an annulus). This implies that

$$\{(x_1, \ldots, x_d) : x_1^2 + x_2^2 < \delta^2, V(x) < 0, \|(x_3, \ldots, x_d)\|_2 < \varepsilon\}$$

is also path-connected. Hence $\mathcal{OV}(z) \cap \mathcal{B}_\varepsilon(z)$ is path-connected.

□

Proposition 2.5. *Assume V is of class $\mathcal{C}^2$, and let z be a nondegenerate stationary point, i.e. such that* $\det(\nabla^2 V(z)) \neq 0$. *Then z is a saddle if and only if $\nabla^2 V(z)$ has exactly one strictly negative eigenvalue.*

Proof. Denote the eigenvalues of $\nabla^2 V(z)$ by $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_d$. If z is a saddle, then the previous result implies that $\lambda_1 < 0 < \lambda_2$. Conversely, if $\lambda_1 < 0 < \lambda_2$, we have

$$V(x) = -\frac{1}{2}|\lambda_1| x_1^2 + \frac{1}{2} \sum_{i=2}^{d} \lambda_i x_i^2 + o(\|x\|_2^2). \tag{2.10}$$

Thus for fixed small $(x_2, \ldots, x_d)$, the set $\{x_1 \colon |x_1| < \varepsilon, V(x) < 0\}$ is not path-connected, as it does not contain 0. Thus $\mathcal{OV}(z) \cap \mathcal{B}_\varepsilon(z)$ is not path-connected (it is topologically the interior of a double cone). However, for $x_1 = 0$, adding the origin makes the set path-connected, so that $(\mathcal{OV}(z) \cap \mathcal{B}_\varepsilon(z)) \cup \{z\}$ is path-connected. □

We can now classify all candidates for saddles in the following way. Let $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_d$ be the eigenvalues of the Hessian $\nabla^2 V(z)$ of a stationary point z, arranged in increasing order. Then the following cases may occur:

1. $\lambda_1 < 0$:
   (a) $\lambda_2 > 0$: z is a nondegenerate saddle.
   (b) $\lambda_2 = 0$:
      i. $\lambda_3 > 0$: z is a singularity of codimension 1.
      ii. $\lambda_3 = 0$:
         A. $\lambda_4 > 0$: z is a singularity of codimension 2.
         B. $\lambda_4 = 0$: z is a singularity of codimension larger than 2.
2. $\lambda_1 = 0$:
   (a) $\lambda_2 > 0$: z is a singularity of codimension 1.
   (b) $\lambda_2 = 0$:
      i. $\lambda_3 > 0$: z is a singularity of codimension 2.
      ii. $\lambda_3 = 0$: z is a singularity of codimension larger than 2.

One can of course push further the classification, including all singularities up to codimension d.

2.3. Singularities of codimension 1

We assume in this subsection that the potential V is of class $\mathcal{C}^4$ and that z is a stationary point of V with the Hessian $\nabla^2 V(z)$ having one vanishing eigenvalue. We may assume $z = 0$ and $V(z) = 0$. According to Proposition 2.4, there are two cases to be considered:

1. $\lambda_1 < 0$, $\lambda_2 = 0$ and $0 < \lambda_3 \leqslant \dots \leqslant \lambda_d$.
2. $\lambda_1 = 0$ and $0 < \lambda_2 \leqslant \dots \leqslant \lambda_d$.

It will be convenient for the purpose of this subsection to relabel the first two eigenvalues in such a way that $\lambda_1 = 0$, while $\lambda_2 \neq 0$ can be positive or negative and to choose a basis in which $\nabla^2 V(z) = \operatorname{diag}(0, \lambda_2, \dots, \lambda_d)$. For potentials V of class $\mathcal{C}^r$ and $i_1, \dots, i_r \in \{1, \dots, d\}$, we introduce the notation

$$V_{i_1 \dots i_r} = \frac{1}{n_1! \dots n_d!} \frac{\partial}{\partial x_{i_1}} \dots \frac{\partial}{\partial x_{i_r}} V(z), \quad n_j = \#\{k \colon i_k = j\}, \tag{2.11}$$

with the convention that the i_k's are always in increasing order.

Under the above assumptions, the potential admits a Taylor expansion of the form

$$V(x) = \frac{1}{2} \sum_{i=2}^{d} \lambda_i x_i^2 + \sum_{1 \leqslant i \leqslant j \leqslant k \leqslant d} V_{ijk} x_i x_j x_k + \sum_{1 \leqslant i \leqslant j \leqslant k \leqslant l \leqslant d} V_{ijkl} x_i x_j x_k x_l + \mathcal{O}(\|x\|_2^4), \tag{2.12}$$

and the theory of normal forms allows to simplify this expression. Indeed, there exists a change of variables $x = y+g(y)$, with g a polynomial function, such that the potential expressed in the new variables has as few as possible terms of low order in its Taylor expansion. In general, only a few so-called resonant terms cannot be eliminated, and are thus essential for describing the local dynamics.

Proposition 2.6. *There exists a polynomial change of variables $x = y + g(y)$, where g is a polynomial with terms of degree 2 and 3, such that*

$$V(y + g(y)) = \frac{1}{2}\sum_{i=2}^{d} \lambda_i y_i^2 + C_3 y_1^3 + C_4 y_1^4 + \mathcal{O}(\|y\|_2^4), \tag{2.13}$$

where

$$C_3 = V_{111}, \quad C_4 = V_{1111} - \frac{1}{2}\sum_{j=2}^{d} \frac{V_{11j}^2}{\lambda_j}. \tag{2.14}$$

The proof uses standard normal form theory and will be given in Appendix A. Note that if V is of class $\mathcal{C}^5$, then (2.13) holds with $\mathcal{O}(\|y\|_2^4)$ replaced by $\mathcal{O}(\|y\|_2^5)$. Let us now apply the result to derive an easy to verify necessary condition for a point z to be a saddle.

Corollary 2.1.

1. *Assume $\lambda_2 < 0$. Then the point z is*
   - *a saddle if $C_3 = 0$ and $C_4 > 0$;*
   - *not a saddle if $C_3 \neq 0$ or $C_4 < 0$.*
2. *Assume $\lambda_2 > 0$. Then the point z is*
   - *a saddle if $C_3 = 0$ and $C_4 < 0$;*
   - *not a saddle if $C_3 \neq 0$ or $C_4 > 0$.*

Proof. Consider first the case $C_3 \neq 0$. For simplicity, let us restrict to $d = 2$. In a neighbourhood of $z = 0$, any solution to the equation $V(y) = 0$ must satisfy

$$y_2^2 = -\frac{2C_3}{\lambda_2} y_1^3 - \frac{2C_4}{\lambda_2} y_1^4 + \mathcal{O}(\|y\|_2^4). \tag{2.15}$$

Thus, solutions exist for y_1 with $y_1 C_3/\lambda_2 < 0$. Plugging the ansatz

$$y_2 = \pm\sqrt{-\frac{2C_3}{\lambda_2} y_1^3}\,\bigl[1 + r_2(y_1)\bigr] \tag{2.16}$$

into the relation $V(y) = 0$, dividing by y_1^3 and applying the implicit-function theorem to the pair (r_2, y_1) in the resulting equation shows that there is a unique curve through the origin on which the potential vanishes. Since for $y_2 = 0$, the potential has the same sign as y_1, we conclude that 0 is not a saddle. Now just note that the proof is similar in dimension $d > 2$.

Consider next the case $C_3 = 0$ and $\lambda_2 C_4 > 0$. If $\lambda_2 > 0$, one sees that the origin is a local minimum. If $\lambda_2 < 0$, then for fixed $(y_3, \dots, y_d)$ and sufficiently small ε, the set $\{(y_1, y_2)\colon y_1^2 + y_2^2 < \varepsilon^2, V(y) < 0\}$ is path-connected (topologically, it is an annulus). Hence $\mathcal{OV}(0) \cup \mathcal{B}_\varepsilon(0)$ is path-connected.

Finally, if $C_3 = 0$ and $\lambda_2 C_4 < 0$, then either the set $\{y_1\colon V(y) < 0\}$ for fixed $(y_2, \dots, y_d)$ or the set $\{y_2\colon V(y) < 0\}$ for fixed $(y_1, y_3, \dots, y_d)$ is not path-connected, so that the valley of 0 is locally split into two disconnected components, joined at the origin. □

Remark 2.1. The normal-form transformation $x \mapsto x + g(x)$ can also be applied when $\lambda_1 \neq 0$. The result is exactly the same normal form as in (2.13), except that there is an additional term $\frac{1}{2}\lambda_1 y_1^2$. This observation is useful as it allows to study the system with a unique transformation of variables in a full neighbourhood of the bifurcation point.

Remark 2.2. One easily checks that if V is of class $\mathcal{C}^r$, one can construct higher-order normal forms by eliminating all terms which are not of the form y_1^k with $k \leqslant r$. In other words, there exists a polynomial $g(y)$ with terms of degree between 2 and $r - 1$ such that

$$V(y + g(y)) = \frac{1}{2}\sum_{i=2}^{d} \lambda_i y_i^2 + C_3 y_1^3 + C_4 y_1^4 + \dots + C_r y_1^r + \mathcal{O}(\|y\|_2^r). \qquad (2.17)$$

In general, however, there is no simple expression of the coefficients of the normal form in terms of the original Taylor coefficients of V. Note that for $z = 0$ to be a saddle, the first index q such that $C_q \neq 0$ has to be even with $\lambda_2 C_q < 0$.

2.4. Singularities of codimension 2

In this subsection, we shall assume that the potential V is of class $\mathcal{C}^r$ for some $r \geqslant 3$ and that z is a stationary point of V with the Hessian $\nabla^2 V(z)$ having two vanishing eigenvalues. We may assume $z = 0$ and $V(z) = 0$. According to Proposition 2.4, there are two cases to be considered:

1. $\lambda_1 < 0$, $\lambda_2 = \lambda_3 = 0$ and $0 < \lambda_4 \leqslant \dots \leqslant \lambda_d$.
2. $\lambda_1 = \lambda_2 = 0$ and $0 < \lambda_3 \leqslant \dots \leqslant \lambda_d$.

For the purpose of this subsection, it will be convenient to relabel the first three eigenvalues in such a way that $\lambda_1 = \lambda_2 = 0$, while $\lambda_3 \neq 0$ can be positive or negative. We choose a basis in which $\nabla^2 V(z) = \mathrm{diag}(0, 0, \lambda_3, \dots, \lambda_d)$. The potential thus admits a Taylor expansion of the form

$$\begin{aligned} V(x) = \frac{1}{2}\sum_{i=3}^{d} \lambda_i x_i^2 &+ \sum_{1 \leqslant i \leqslant j \leqslant k \leqslant d} V_{ijk} x_i x_j x_k \\ &+ \sum_{1 \leqslant i \leqslant j \leqslant k \leqslant l \leqslant d} V_{ijkl} x_i x_j x_k x_l + \dots . \end{aligned} \tag{2.18}$$

The theory of normal forms immediately yields the following result.

Proposition 2.7. *There exists a polynomial change of variables $x = y + g(y)$, where g is a polynomial with terms of degree between 2 and r, such that*

$$V(y + g(y)) = \frac{1}{2}\sum_{i=3}^{d} \lambda_i y_i^2 + \sum_{k=3}^{r} V_k(y_1, y_2) + o(\|y\|_2^r), \tag{2.19}$$

where each V_k is a homogeneous polynomial of order k, i.e.,

$$\begin{aligned} V_3(y_1, y_2) &= V_{111} y_1^3 + V_{112} y_1^2 y_2 + V_{122} y_1 y_2^2 + V_{222} y_2^3, \\ V_4(y_1, y_2) &= V_{1111} y_1^4 + V_{1112} y_1^3 y_2 + V_{1122} y_1^2 y_2^2 + V_{1222} y_1 y_2^3 + V_{2222} y_2^4, \end{aligned} \tag{2.20}$$

and similarly for the higher-order terms.

The proof uses standard normal form theory and follows along the lines of the proof of Proposition 2.6, given in Appendix A. Therefore, we refrain from giving its details here. Let us remark that if V is of class $\mathcal{C}^{r+1}$, then (2.19) holds with $o(\|y\|_2^r)$ replaced by $\mathcal{O}(\|y\|_2^{r+1})$.

We can again apply the result to derive an easy to verify necessary condition for a point z to be a saddle. Let p be the smallest k such that V_k is not identically zero. Generically, we will have $p = 3$, but other values of p are quite possible, for instance due to symmetries (see Section 4.3 for examples). We call *discriminant* the polynomial[2]

$$\Delta(t) = \begin{cases} V_p(t, 1) = V_{1\dots11} t^p + V_{1\dots12} t^{p-1} + \dots + V_{2\dots22}, & \text{if } V_{1\dots11} \neq 0, \\ V_p(1, t) = V_{2\dots22} t^p + V_{2\dots21} t^{p-1} + \dots + V_{1\dots11}, & \text{if } V_{1\dots11} = 0. \end{cases} \tag{2.21}$$

The following corollary as well as Table 1 provide necessary conditions for z to be a saddle, expressed in terms of $\Delta(t)$ and the sign of λ_3.

[2] The rôles of y_1 and y_2 being interchangeable, both definitions in (2.21) are equivalent via the transformation $t \mapsto 1/t$ and multiplication by t^p. We choose to make this distinction in order to avoid having to introduce "roots at infinity". Equivalently, one could work on the projective line $\mathbb{R}P^1$.

	$\lambda_3 < 0$	$\lambda_3 > 0$
All roots of $\Delta(t)$ real and simple	Not a saddle	**Saddle**
No real root of $\Delta(t)$ and $\Delta(t) > 0$	**Saddle**	Not a saddle (local minimum)
No real root of $\Delta(t)$ and $\Delta(t) < 0$	Not a saddle	Not a saddle

Table 1. Classification of stationary points with double-zero eigenvalue with the help of the discriminant $\Delta(t)$ and the sign of λ_3.

Corollary 2.2.

1. *Assume $d > 2$ and $\lambda_3 < 0$. Then the point z is*
   - *not a saddle if the discriminant has one or several real roots, all of them simple;*
   - *a saddle if the discriminant has no real root and is positive;*
   - *not a saddle if the discriminant has no real root and is negative.*
2. *Assume $d = 2$ or $\lambda_3 > 0$. Then the point z is*
   - *a saddle if the discriminant has one or several real roots, all of them simple;*
   - *not a saddle* (*a local minimum*) *if the discriminant has no real root and is positive;*
   - *not a saddle if the discriminant has no real root and is negative.*

Proof. In order to determine the open valley $\mathcal{OV}(z)$ of $z = 0$, we first look for solutions of $V(y_1, y_2, 0, \dots, 0) = 0$ near the origin. Such solutions are necessarily of the form $y_1 = y_2(t + R(y_2))$ for some t, where $R(y_2)$ goes to zero continuously as $y_2 \to 0$. Plugging in, dividing by y_2^p and setting $y_2 = 0$, one sees that t must be a root of the discriminant. Applying the implicit function theorem to the pair (y_2, R), one finds that there is a unique R if this root is simple. Thus whenever the discriminant has real roots, all of them simple, $V(y_1, y_2, 0, \dots, 0)$ changes sign in a neighbourhood of the origin, the regions of constant sign being shaped like sectors. If $\lambda_3 < 0$, we have to distinguish between three cases:

1. The discriminant has simple real roots. In the plane $y_3 = \dots = y_d = 0$, there are several disconnected regions in which V is negative. However these regions merge when y_3 becomes nonzero. Hence all regions can be connected by a path leaving the plane $y_3 = \dots = y_d = 0$, so that the origin is not a saddle.

2. The discriminant has no real roots and is positive. Then for each fixed $(y_1, y_2, y_4, \dots)$, the set of y_3 such that V is negative is not path-connected, and consequently the set $\{y \colon V(y) < 0\}$ cannot be path-connected either. By adding the origin, the latter set becomes path-connected, showing that 0 is indeed a saddle.

3. The discriminant has no real roots and is negative. Then either $d = 3$ and the origin is a local maximum, or $d > 3$ and the set of $(y_4, \dots, y_d)$ for which V is negative is path-connected for each fixed (y_1, y_2, y_3). Thus any two points in the open valley close to the origin can be connected (connect both endpoints to points in the set $y_4 = \dots = y_d = 0$ by a path with constant (y_1, y_2, y_3), and then connect the two paths within the set $y_4 = \dots = y_d = 0$). Thus the origin cannot be a saddle.

The proofs are analogous in the case $\lambda_3 > 0$. $\square$

We note that if $\lambda_3 < 0$ and the degree p of the discriminant is odd, then the origin is usually not a saddle. For the sake of brevity, we do not discuss here cases in which the discriminant has nonsimple roots, because then the behaviour depends on higher-order terms in the Taylor expansion and there is a large number of cases to distinguish.

2.5. Singularities of higher codimension

Generalisation to nonquadratic saddles of higher codimension is now quite obvious. If the potential V is of class $\mathcal{C}^r$ and the first q eigenvalues $\lambda_1, \dots, \lambda_q$ of the Hessian $\nabla^2 V(z)$ are equal to zero, the normal form can be written as

$$V(y + g(y)) = \frac{1}{2} \sum_{i=q+1}^{d} \lambda_i y_i^2 + \sum_{k=3}^{r} V_k(y_1, \dots, y_q) + \mathcal{O}(\|y\|_2^r), \tag{2.22}$$

where again each V_k is a homogeneous polynomial of degree k. Let p denote the smallest k such that V_k is not identically zero. Then the role of the discriminant is now played by its analogue $\Delta(t_1, \dots, t_{q-1}) = V_p(t_1, t_2, \dots, t_{q-1}, 1)$, and the classification is analogous to the one in Table 1 (with λ_{q+1} instead of λ_3). Equivalently, one can study the sign of V_p on a sphere of constant radius.

3. First-passage times for nonquadratic saddles

3.1. Some potential theory

Let $(x_t)_t$ be the solution of the stochastic differential equation (1.1). Given a measurable set $A \subset \mathbb{R}^d$, we denote by $\tau_A = \inf\{t > 0 \colon x_t \in A\}$ the first-hitting time of A. For sets[3] $A, B \subset \mathbb{R}^d$, the quantity $h_{A,B}(x) = \mathbb{P}^x\{\tau_A < \tau_B\}$ is known

[3] All subsets of $\mathbb{R}^d$ we consider from now on will be assumed to be *regular*, that is, their complement is a region with continuously differentiable boundary.

to satisfy the boundary value problem

$$\begin{cases} Lh_{A,B}(x) = 0, & \text{for } x \in (A \cup B)^c, \\ h_{A,B}(x) = 1, & \text{for } x \in A, \\ h_{A,B}(x) = 0, & \text{for } x \in B, \end{cases} \tag{3.1}$$

where $L = \varepsilon\Delta - \langle \nabla V(\cdot), \nabla \rangle$ is the infinitesimal generator of the diffusion $(x_t)_t$. By analogy with the electrical potential created between two conductors at potentials 1 and 0, respectively, $h_{A,B}$ is called the *equilibrium potential* of A and B. More generally, one can define an equilibrium potential $h^\lambda_{A,B}$, defined by a boundary value problem similar to (3.1), but with $Lh^\lambda_{A,B} = \lambda h^\lambda_{A,B}$. However, we will not need this generalisation here.

The *capacity* of the sets A and B is again defined in analogy with electrostatics as the total charge accumulated on one conductor of a capacitor, for unit potential difference. The most useful expression for our purpose is the integral, or *Dirichlet form*,

$$\operatorname{cap}_A(B) = \varepsilon \int\limits_{(A\cup B)^c} \mathrm{e}^{-V(x)/\varepsilon} \, \|\nabla h_{A,B}(x)\|_2^2 \, \mathrm{d}x =: \Phi_{(A\cup B)^c}(h_{A,B}). \tag{3.2}$$

We will use the fact that the equilibrium potential $h_{A,B}$ minimises the Dirichlet form $\Phi_{(A\cup B)^c}$, i.e.,

$$\operatorname{cap}_A(B) = \inf_{h \in \mathcal{H}_{A,B}} \Phi_{(A\cup B)^c}(h). \tag{3.3}$$

Here $\mathcal{H}_{A,B}$ is the space of twice weakly differentiable functions, whose derivatives up to order 2 are in L^2, and which satisfy the boundary conditions in (3.1).

Proposition 6.1 in [2] shows (under some assumptions which can be relaxed to suit our situation) that if x is a (quadratic) local minimum of the potential, then the expected first-hitting time of a set B is given by

$$\mathbb{E}^x\{\tau_B\} = \frac{\int_{B^c} \mathrm{e}^{-V(y)/\varepsilon} \, h_{\mathcal{B}_\varepsilon(x),B}(y) \, \mathrm{d}y}{\operatorname{cap}_{\mathcal{B}_\varepsilon(x)}(B)}. \tag{3.4}$$

The numerator can be estimated by the Laplace method, using some rough *a priori* estimates on the equilibrium potential $h_{\mathcal{B}_\varepsilon(x),B}$ (which is close to 1 in a neighbourhood of x, and negligibly small elsewhere). In the generic situation where x is indeed a quadratic local minimum, and the saddle z forms the gate from x to B, it is known that

$$\int\limits_{B^c} \mathrm{e}^{-V(y)/\varepsilon} \, h_{B_\varepsilon(x),B}(y) \, \mathrm{d}y = \frac{(2\pi\varepsilon)^{d/2}}{\sqrt{\det(\nabla^2 V(x))}} \, \mathrm{e}^{-V(x)/\varepsilon} \left[1 + \mathcal{O}(\varepsilon^{1/2}|\log\varepsilon|)\right], \tag{3.5}$$

cf. [2, Equation (6.13)]. Thus, the crucial quantity to be computed is the capacity in the denominator. In the simplest case of a quadratic saddle z whose Hessian has eigenvalues $\lambda_1 < 0 < \lambda_2 \leqslant \ldots \leqslant \lambda_d$, one finds

$$\operatorname{cap}_{\mathcal{B}_\varepsilon(x)}(B) = \frac{1}{2\pi}\sqrt{\frac{(2\pi\varepsilon)^d|\lambda_1|}{\lambda_2\ldots\lambda_d}}\,\mathrm{e}^{-V(z)/\varepsilon}\left[1+\mathcal{O}(\varepsilon^{1/2}|\log\varepsilon|)\right], \tag{3.6}$$

cf. [2, Theorem 5.1], which implies the standard Eyring – Kramers formula (1.3).

In the sequel, we shall thus estimate the capacity in cases where the gate between $A = \mathcal{B}_\varepsilon(x)$ and B is a non-quadratic saddle. Roughly speaking, the central result, which is proved in Section 5, states that if the normal form of the saddle is of the type

$$V(y) = -u_1(y_1) + u_2(y_2,\ldots,y_q) + \frac{1}{2}\sum_{j=q+1}^{d}\lambda_j y_j^2 + \mathcal{O}(\|y\|_2^{r+1}), \tag{3.7}$$

where the functions u_1 and u_2 satisfy appropriate growth conditions, then

$$\operatorname{cap}_A(B) = \varepsilon\,\frac{\int_{\mathcal{B}_{\delta_2}(0)}\mathrm{e}^{-u_2(y_2,\ldots,y_q)/\varepsilon}\,\mathrm{d}y_2\ldots\mathrm{d}y_q}{\int_{-\delta_1}^{\delta_1}\mathrm{e}^{-u_1(y_1)/\varepsilon}\,\mathrm{d}y_1}\prod_{j=q+1}^{d}\sqrt{\frac{2\pi\varepsilon}{\lambda_j}}\left[1+R(\varepsilon)\right], \tag{3.8}$$

for certain $\delta_1 = \delta_1(\varepsilon) > 0$, $\delta_2 = \delta_2(\varepsilon) > 0$. The remainder $R(\varepsilon)$ goes to zero as $\varepsilon \to 0$, with a speed depending on u_1 and u_2. Once this result is established, the computation of capacities is reduced to the computation of the integrals in (3.8).

Remark 3.1. If x is a nonquadratic local minimum, the integral in (3.5) can also be estimated easily by standard Laplace asymptotics. Hence the extension of the Eyring – Kramers formula to flat local minima is straightforward, and the real difficulty lies in the effect of flat saddles.

3.2. Transition times for codimension 1 singular saddles

We assume in this section that the potential is of class $\mathcal{C}^5$ at least, as this allows a better control of the error terms. Consider first the case of a saddle z such that the Hessian matrix $\nabla^2 V(z)$ has eigenvalues $\lambda_1 = 0 < \lambda_2 \leqslant \lambda_3 \leqslant \ldots \leqslant \lambda_d$. In this case, the unstable direction at the saddle is non-quadratic while all stable directions are quadratic. According to Corollary 2.1, in the most generic case the potential admits a normal form

$$V(y) = -C_4 y_1^4 + \frac{1}{2}\sum_{j=2}^{d}\lambda_j y_j^2 + \mathcal{O}(\|y\|_2^5) \tag{3.9}$$

with $C_4 > 0$, i.e., the unstable direction is quartic. (Note that the saddle z is at the origin 0 of this coordinate system.)

We are interested in transition times between sets A and B for which the gate $G(A,B)$ consists only of the saddle z. In other words, we assume that any minimal path $\gamma \in \mathcal{P}(A,B)$ admits z as unique point of highest altitude. This does not exclude the existence of other stationary points in $\mathcal{OV}(z)$, i.e., the potential seen along the path γ may have several local minima and maxima.

Theorem 3.1. *Assume z is a saddle whose normal form satisfies* (3.9). *Let A and B belong to different path-connected components of $\mathcal{OV}(z)$ and assume that $G(A,B) = \{z\}$. Then*

$$\operatorname{cap}_A(B) = \frac{2C_4^{1/4}}{\Gamma(1/4)} \sqrt{\frac{(2\pi)^{d-1}}{\lambda_2 \dots \lambda_d}} \, \varepsilon^{d/2+1/4} \, \mathrm{e}^{-V(z)/\varepsilon} \bigl[1 + \mathcal{O}(\varepsilon^{1/4} |\log \varepsilon|^{5/4})\bigr], \quad (3.10)$$

where Γ denotes the Euler Gamma function.

The proof is given in Section 5.3. In the case of a quadratic local minimum x, combining this result with Estimate (3.5) immediately yields the following result on first-hitting times.

Corollary 3.1. *Assume z is a saddle whose normal form satisfies* (3.9). *Let O be one of the path-connected components of $\mathcal{OV}(z)$, and suppose that the minimum of V in O is attained at a unique point x and is quadratic. Let B belong to a different path-connected component of $\mathcal{OV}(z)$, with $G(\{x\},B) = \{z\}$. Then the expected first-hitting time of B satisfies*

$$\mathbb{E}^x\bigl\{\tau_B\bigr\} = \frac{\Gamma(1/4)}{2C_4^{1/4}} \sqrt{\frac{2\pi\lambda_2 \dots \lambda_d}{\det(\nabla^2 V(x))}} \varepsilon^{-1/4} \, \mathrm{e}^{[V(z)-V(x)]/\varepsilon} \bigl[1 + \mathcal{O}(\varepsilon^{1/4} |\log \varepsilon|^{5/4})\bigr]. \quad (3.11)$$

Note in particular that unlike in the case of a quadratic saddle, the subexponential asymptotics depends on ε to leading order, namely proportionally to $\varepsilon^{-1/4}$.

Remark 3.2.

1. If the gate $G(A,B)$ contains several isolated saddles, the capacity is obtained simply by adding the contributions of each individual saddle. In other words, just as in electrostatics, for capacitors in parallel the equivalent capacity is obtained by adding the capacities of individual capacitors.

2. We can extend this result to even flatter unstable directions. Assume that the potential V is of class $2p+1$ for some $p \geqslant 2$, and that the normal form

at the origin reads

$$V(y) = -C_{2p} y_1^{2p} + \frac{1}{2} \sum_{j=2}^{d} \lambda_j y_j^2 + \mathcal{O}(\|y\|_2^{2p+1}), \tag{3.12}$$

with $C_{2p} > 0$. Then a completely analogous proof shows that (3.10) is to be replaced by

$$\operatorname{cap}_A(B) = \frac{pC_{2p}^{1/2p}}{\Gamma(1/2p)} \sqrt{\frac{(2\pi)^{d-1}}{\lambda_2 \dots \lambda_d}}\, \varepsilon^{d/2+(p-1)/2p} \bigl[1+\mathcal{O}(\varepsilon^{1/2p} |\log\varepsilon|^{(2p+1)/2p})\bigr]. \tag{3.13}$$

Consequently, if the assumptions of Corollary 3.1 on the minimum x and the set B are satisfied, then

$$\begin{aligned} \mathbb{E}^x\{\tau_B\} = {} & \frac{\Gamma(1/2p)}{pC_{2p}^{1/2p}} \sqrt{\frac{2\pi\lambda_2 \dots \lambda_d}{\det(\nabla^2 V(x))}} \varepsilon^{-(p-1)/2p} \\ & \times \mathrm{e}^{[V(z)-V(x)]/\varepsilon} \bigl[1 + \mathcal{O}(\varepsilon^{1/2p} |\log\varepsilon|^{(2p+1)/2p})\bigr]. \end{aligned} \tag{3.14}$$

Note that the subexponential prefactor of the expected first-hitting time now behaves like $\varepsilon^{-(p-1)/2p}$. As p varies from 1 to ∞, i.e., as the unstable direction becomes flatter and flatter, the prefactor's dependence on ε changes from order 1 to order $1/\sqrt{\varepsilon}$.

Consider next the case of a saddle z such that the Hessian matrix $\nabla^2 V(z)$ has eigenvalues $\lambda_1 < \lambda_2 = 0 < \lambda_3 \leqslant \dots \leqslant \lambda_d$. In this case, all directions, whether stable or unstable, are quadratic but for one of the stable directions. According to Corollary 2.1, in the most generic case the potential admits a normal form

$$V(y) = -\frac{1}{2}|\lambda_1| y_1^2 + C_4 y_2^4 + \frac{1}{2} \sum_{j=3}^{d} \lambda_j y_j^2 + \mathcal{O}(\|y\|_2^5) \tag{3.15}$$

with $C_4 > 0$, i.e., the non-quadratic stable direction is quartic.

Theorem 3.2. *Assume z is a saddle whose normal form satisfies* (3.15). *Let A and B belong to different path-connected components of $\mathcal{OV}(z)$ and assume that $G(A,B) = \{z\}$. Then*

$$\operatorname{cap}_A(B) = \frac{\Gamma(1/4)}{2C_4^{1/4}} \sqrt{\frac{(2\pi)^{d-3}|\lambda_1|}{\lambda_3 \dots \lambda_d}}\, \varepsilon^{d/2-1/4}\, \mathrm{e}^{-V(z)/\varepsilon} \bigl[1 + \mathcal{O}(\varepsilon^{1/4} |\log\varepsilon|^{5/4})\bigr]. \tag{3.16}$$

The proof is given in Section 5.3. In the case of a quadratic local minimum x, combining this result with Estimate (3.5) immediately yields the following result on first-hitting times.

Corollary 3.2. *Assume z is a saddle whose normal form satisfies* (3.15)*. Let O be one of the path-connected components of $\mathcal{OV}(z)$, and suppose that the minimum of V in O is attained at a unique point x and is quadratic. Let B belong to a different path-connected component of $\mathcal{OV}(z)$, with $G(\{x\}, B) = \{z\}$. Then the expected first-hitting time of B satisfies*

$$\mathbb{E}^x\{\tau_B\} = \frac{2C_4^{1/4}}{\Gamma(1/4)}\sqrt{\frac{(2\pi)^3\lambda_3\dots\lambda_d}{|\lambda_1|\det(\nabla^2 V(x))}}\varepsilon^{1/4}\,\mathrm{e}^{[V(z)-V(x)]/\varepsilon}\,\bigl[1+\mathcal{O}(\varepsilon^{1/4}|\log\varepsilon|^{5/4})\bigr]. \tag{3.17}$$

Note again the ε-dependence of the prefactor, which is now proportional to $\varepsilon^{1/4}$ to leading order. A similar result is easily obtained in the case of the leading term in the normal form having order y_2^{2p} for some $p \geqslant 2$. In particular, the prefactor of the transition time then has leading order $\varepsilon^{(p-1)/2p}$:

$$\begin{aligned}\mathbb{E}^x\{\tau_B\} = {}&\frac{pC_{2p}^{1/2p}}{\Gamma(1/2p)}\sqrt{\frac{(2\pi)^3\lambda_3\dots\lambda_d}{|\lambda_1|\det(\nabla^2 V(x))}}\varepsilon^{(p-1)/2p}\\ &\times\mathrm{e}^{[V(z)-V(x)]/\varepsilon}\,\bigl[1+\mathcal{O}(\varepsilon^{1/2p}|\log\varepsilon|^{(2p+1)/2p})\bigr].\end{aligned} \tag{3.18}$$

As p varies from 1 to ∞, i.e., as the non-quadratic stable direction becomes flatter and flatter, the prefactor's dependence on ε changes from order 1 to order $\sqrt{\varepsilon}$.

3.3. Transition times for higher-codimension singular saddles

We assume in this section that the potential is at least of class $\mathcal{C}^5$. Consider the case of a saddle z such that the Hessian matrix $\nabla^2 V(z)$ has eigenvalues

$$\lambda_1 < 0 = \lambda_2 = \lambda_3 < \lambda_4 \leqslant \dots \leqslant \lambda_d. \tag{3.19}$$

In this case, the unstable direction is quadratic, while two of the stable directions are non-quadratic. Proposition 2.7 shows that near the saddle, the potential admits a normal form

$$V(y) = -\frac{1}{2}|\lambda_1|y_1^2 + V_3(y_2,y_3) + V_4(y_2,y_3) + \frac{1}{2}\sum_{j=4}^{d}\lambda_j y_j^2 + \mathcal{O}(\|y\|_2^5), \tag{3.20}$$

with V_3 and V_4 homogeneous polynomials of degree 3 and 4, respectively. If V_3 does not vanish identically, Corollary 2.2 shows that z is typically not a saddle.

We assume thus that V_3 is identically zero, and, again in view of Corollary 2.2, that the discriminant

$$\Delta(t) = V_{2222}t^4 + V_{2223}t^3 + V_{2233}t^2 + V_{2333}t + V_{3333} \tag{3.21}$$

has no real roots, and is positive with $V_{2222} > 0$. It is convenient to introduce polar coordinates, writing

$$V_4(r\cos\varphi, r\sin\varphi) = r^4 k(\varphi), \tag{3.22}$$

where we may assume that $k(\varphi)$ is bounded above and below by strictly positive constants $K_+ \geqslant K_-$. Then we have the following result, which is proved in Section 5.3.

Theorem 3.3. *Assume z is a saddle whose normal form satisfies* (3.20)*, with $V_3 \equiv 0$ and $V_4 > 0$. Suppose, the discriminant $\Delta(t)$ has no real roots and satisfies $V_{2222} > 0$. Let A and B belong to different path-connected components of $\mathcal{OV}(z)$ and assume that $G(A,B) = \{z\}$. Then*

$$\operatorname{cap}_A(B) = \frac{\sqrt{\pi}}{4}\int_0^{2\pi} \frac{\mathrm{d}\varphi}{k(\varphi)^{1/2}} \sqrt{\frac{(2\pi)^{d-4}|\lambda_1|}{\lambda_4\dots\lambda_d}}\,\varepsilon^{d/2-1/2} \times \mathrm{e}^{-V(z)/\varepsilon}\bigl[1 + \mathcal{O}(\varepsilon^{1/4}|\log\varepsilon|^{5/4})\bigr]. \tag{3.23}$$

Corollary 3.3. *Assume z is a saddle whose normal form satisfies* (3.20)*. Let O be one of the path-connected components of $\mathcal{OV}(z)$, and assume that the minimum of V in O is reached at a unique point x, which is quadratic. Let B belong to a different path-connected component of $\mathcal{OV}(z)$, with $G(\{x\},B) = \{z\}$. Then the expected first-hitting time of B satisfies*

$$\mathbb{E}^x\{\tau_B\} = \frac{4}{\sqrt{\pi}\int_0^{2\pi}\frac{\mathrm{d}\varphi}{k(\varphi)^{1/2}}} \sqrt{\frac{(2\pi)^4\lambda_4\dots\lambda_d}{|\lambda_1|\det(\nabla^2 V(x))}}\,\varepsilon^{1/2} \times \mathrm{e}^{[V(z)-V(x)]/\varepsilon}\bigl[1 + \mathcal{O}(\varepsilon^{1/4}|\log\varepsilon|^{5/4})\bigr]. \tag{3.24}$$

The prefactor is now proportional to $\varepsilon^{1/2}$ instead of being proportional to $\varepsilon^{1/4}$, which is explained by the presence of two vanishing eigenvalues.

Remark 3.3. This result admits two straightforward generalisations to less generic situations:

1. If the potential is of class $\mathcal{C}^{2p+1}$ and the first nonvanishing coefficient of the normal form has even degree $2p$, $p \geqslant 2$, then a completely analogous

proof shows that

$$\operatorname{cap}_A(B) = \frac{1}{2p}\Gamma\biggl(\frac{1}{p}\biggr)\int\limits_0^{2\pi}\frac{\mathrm{d}\varphi}{k(\varphi)^{1/p}}\sqrt{\frac{(2\pi)^{d-4}|\lambda_1|}{\lambda_4\dots\lambda_d}}\,\varepsilon^{d/2-(p-1)/p}\,\mathrm{e}^{-V(z)/\varepsilon} \times \bigl[1+\mathcal{O}(\varepsilon^{1/2p}|\log\varepsilon|^{(2p+1)/2p})\bigr]. \tag{3.25}$$

Consequently, if the assumptions of Corollary 3.3 on the minimum x and the set B are satisfied, then

$$\mathbb{E}^x\{\tau_B\} = \frac{2p}{\Gamma\bigl(\frac{1}{p}\bigr)}\frac{1}{\int_0^{2\pi}\frac{\mathrm{d}\varphi}{k(\varphi)^{1/p}}}\sqrt{\frac{(2\pi)^4\lambda_4\dots\lambda_d}{|\lambda_1|\det(\nabla^2V(x))}}\,\varepsilon^{(p-1)/p}\,\mathrm{e}^{[V(z)-V(x)]/\varepsilon} \times \bigl[1+\mathcal{O}(\varepsilon^{1/2p}|\log\varepsilon|^{(2p+1)/2p})\bigr]. \tag{3.26}$$

2. If all eigenvalues from λ_2 to λ_q are equal to zero, for some $q \geqslant 4$ and the first nonvanishing coefficient of the normal form has even degree $2p$, $p \geqslant 2$, then the prefactor of the capacity has order $\varepsilon^{d/2-(q-1)(p-1)/2p}$, and involves a $(q-2)$-dimensional integral over the angular part of the leading term in the normal form.

The other important codimension-two singularity occurs for a saddle z such that the Hessian matrix $\nabla^2 V(z)$ has eigenvalues

$$0 = \lambda_1 = \lambda_2 < \lambda_3 \leqslant \dots \leqslant \lambda_d. \tag{3.27}$$

In this case, Corollary 2.2 states that z is a saddle when the discriminant of the normal form has one or more real roots, all of them simple. As a consequence, there can be more than two valleys meeting at the saddle. This actually induces a serious difficulty for the estimation of the capacity. The reason is that for this estimation, one needs an *a priori* bound on the equilibrium potential $h_{A,B}$ in the valleys, some distance away from the saddle. In cases with only two valleys, $h_{A,B}$ is very close to 1 in the valley containing A, and very close to 0 in the valley containing B. When there are additional valleys, however, one would first have to obtain an *a priori* estimate on the value of $h_{A,B}$ in these valleys, which is not at all straightforward, except perhaps in situations involving symmetries. We will not discuss this case here.

Arguably, the study of singular saddles satisfying (3.27) is less important, because they are less stable against perturbations of the potential. Namely, the flatness of the potential around the saddle implies that like for the saddle–node bifurcation, there exist perturbations of the potential that do no longer admit a saddle close by. As a consequence, there is no potential barrier creating metastability for these perturbations.

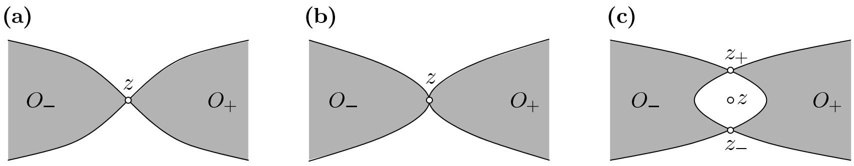

Figure 3. Saddles and open valleys of the normal-form potential (4.2), in a two-dimensional case, **(a)** for $\lambda_2 > 0$, **(b)** for $\lambda_2 = 0$ and **(c)** for $\lambda_2 < 0$. The system undergoes a transversal pitchfork bifurcation at $\lambda_2 = 0$.

4. Bifurcations

While the results in the previous section describe the situation for nonquadratic saddles, that is, at a bifurcation point, they do not incorporate the transition from quadratic to nonquadratic saddles. In order to complete the picture, we now give a description of the metastable behaviour in a full neighbourhood of a bifurcation point of a parameter-dependent potential. We will always assume that V is of class $\mathcal{C}^5$.

We shall discuss a few typical examples of bifurcations, which we will illustrate on the model potential

$$V_\gamma(x) = \sum_{i=1}^{N} U(x_i) + \frac{\gamma}{4}\sum_{i=1}^{N}(x_i - x_{i+1})^2, \tag{4.1}$$

introduced in [3]. Here x_i denotes the position of a particle attached to site i of the lattice $\mathbb{Z}/N\mathbb{Z}$, $U(x_i) = (1/4)x_i^4 - (1/2)x_i^2$ is a local double-well potential acting on that particle, and the second sum describes a harmonic ferromagnetic interaction between neighbouring particles (with the identification $x_{N+1} = x_1$). Indeed, for $N = 2$ and $\gamma = 1/2$, the origin is a nonquadratic saddle of codimension 1 of the potential (4.1), while for all $N \geqslant 3$, the origin is a nonquadratic saddle of codimension 2 when $\gamma = (2\sin^2(\pi/N))^{-1}$.

4.1. Transversal symmetric pitchfork bifurcation

Let us assume that the potential V depends continuously on a parameter γ, and that for $\gamma = \gamma^\star$, $z = 0$ is a nonquadratic saddle of V, with normal form (3.15). A symmetric pitchfork bifurcation occurs when for γ near $\gamma^\star$, the normal form has the expression

$$V(y) = \frac{1}{2}\lambda_1(\gamma)y_1^2 + \frac{1}{2}\lambda_2(\gamma)y_2^2 + C_4(\gamma)y_2^4 + \frac{1}{2}\sum_{j=3}^{d}\lambda_j(\gamma)y_j^2 + \mathcal{O}(\|y\|_2^5), \tag{4.2}$$

where $\lambda_2(\gamma^\star) = 0$, while $\lambda_1(\gamma^\star) < 0$, $C_4(\gamma^\star) > 0$, and similarly for the other quantities. We assume here that V is even in y_2, which is the most common

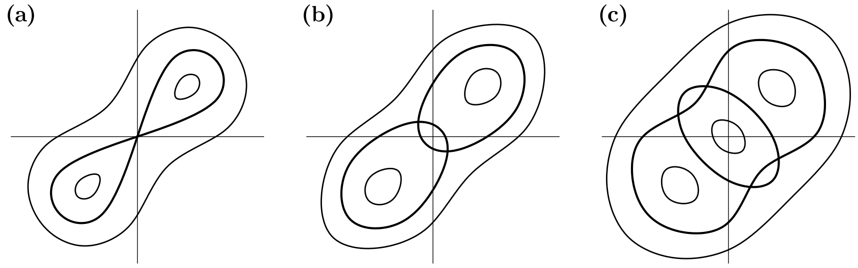

Figure 4. Level lines of the potential (4.5), **(a)** for $\gamma > 1/2$, **(b)** for $1/2 > \gamma > 1/3$ and **(c)** for $1/3 > \gamma > 0$. A transversal pitchfork bifurcation occurs at $\gamma = 1/2$, and a longitudinal pitchfork bifurcation occurs at $\gamma = 1/3$.

situation in which pitchfork bifurcations are observed. For simplicity, we shall usually refrain from indicating the γ-dependence of the eigenvalues in the sequel. All quantities except λ_2 are assumed to be bounded away from zero as γ varies.

When $\lambda_2 > 0$, $z = 0$ is a quadratic saddle. When $\lambda_2 < 0$, $z = 0$ is no longer a saddle (the origin then having a two-dimensional unstable manifold), but there exist two saddles $z_\pm$ with coordinates $z_\pm = \bigl(0, \pm\sqrt{|\lambda_2|/4C_4} + \mathcal{O}(\lambda_2), 0, \dots, 0\bigr) + \mathcal{O}(\lambda_2^2)$ (Figure 3). Let us denote the eigenvalues of $\nabla^2 V(z_\pm)$ by $\mu_1, \dots, \mu_d$. In fact, for $\lambda_2 < 0$ we have

$$\mu_2 = -2\lambda_2 + \mathcal{O}(|\lambda_2|^{3/2}), \quad \mu_j = \lambda_j + \mathcal{O}(|\lambda_2|^{3/2}) \quad \text{for } j \in \{1, 3, \dots, d\}. \tag{4.3}$$

Finally, the value of the potential on the saddles $z_\pm$ satisfies

$$V(z_+) = V(z_-) = V(z) - \frac{\lambda_2^2}{16C_4} + \mathcal{O}(|\lambda_2|^{5/2}). \tag{4.4}$$

Example 4.1. For $N = 2$ particles, the potential (4.1) reads

$$V(x_1, x_2) = U(x_1) + U(x_2) + \frac{\gamma}{2}(x_1 - x_2)^2. \tag{4.5}$$

Performing a rotation by $\pi/4$ yields the equivalent potential

$$\widehat{V}(y_1, y_2) = -\frac{1}{2}y_1^2 - \frac{1-2\gamma}{2}y_2^2 + \frac{1}{8}(y_1^4 + 6y_1^2y_2^2 + y_2^4), \tag{4.6}$$

which immediately shows that the origin $(0,0)$ is a stationary point with $\lambda_1(\gamma) = -1$ and $\lambda_2(\gamma) = -(1-2\gamma)$. For $\gamma > \gamma^\star = 1/2$, the origin is thus a quadratic saddle, at "altitude"0. It serves as a gate between the local minima located at $y = (\pm 1/\sqrt{2}, 0)$. As γ decreases below $\gamma^\star$, two new saddles appear at $y = (0, \pm\sqrt{2(1-2\gamma)}$ (cf. Figure 4 which shows the potential's level lines in the

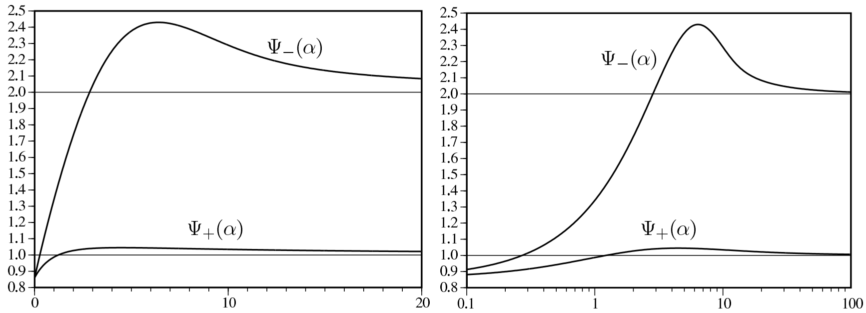

Figure 5. The functions $\Psi_\pm(\alpha)$, shown on a linear and on a logarithmic scale.

original variables (x_1, x_2)). They have a positive eigenvalue $\mu_2(\gamma) = 2(2\gamma - 1)$, and the "altitude" $-(1/2)(1 - 2\gamma)^2$. There is thus a pitchfork bifurcation at $\gamma = 1/2$. Note that another pitchfork bifurcation, affecting the new saddles, occurs at $\gamma = 1/3$.

Our main result is the following sharp estimate of the capacity.

Theorem 4.1. *Assume the normal form at $z=0$ satisfies (4.2). Let A and B belong to different path-connected components of $\mathcal{OV}(z)$ (respectively of $\mathcal{OV}(z_\pm)$ if $\lambda_2 < 0$). Assume further that $G(A,B) = \{z\}$ (resp. $G(A,B) = \{z_-, z_+\}$ if $\lambda_2 < 0$). Then for $\lambda_2 > 0$,*

$$\operatorname{cap}_A(B) = \sqrt{\frac{(2\pi)^{d-2}|\lambda_1|}{[\lambda_2 + (2\varepsilon C_4)^{1/2}]\lambda_3 \dots \lambda_d}}\, \Psi_+\Big(\frac{\lambda_2}{(2\varepsilon C_4)^{1/2}}\Big)\varepsilon^{d/2} \times \mathrm{e}^{-V(z)/\varepsilon}\big[1 + R_+(\varepsilon, \lambda_2)\big], \tag{4.7}$$

while for $\lambda_2 < 0$,

$$\operatorname{cap}_A(B) = \sqrt{\frac{(2\pi)^{d-2}|\mu_1|}{[\mu_2 + (2\varepsilon C_4)^{1/2}]\mu_3 \dots \mu_d}}\, \Psi_-\Big(\frac{\mu_2}{(2\varepsilon C_4)^{1/2}}\Big)\varepsilon^{d/2} \times \mathrm{e}^{-V(z_\pm)/\varepsilon}\big[1 + R_-(\varepsilon, \mu_2)\big]. \tag{4.8}$$

The functions Ψ_+ and Ψ_- are bounded above and below uniformly on $\mathbb{R}_+$. They admit the explicit expressions

$$\Psi_+(\alpha) = \sqrt{\frac{\alpha(1+\alpha)}{8\pi}}\,\mathrm{e}^{\alpha^2/16}\,K_{1/4}\Big(\frac{\alpha^2}{16}\Big),$$
$$\Psi_-(\alpha) = \sqrt{\frac{\pi\alpha(1+\alpha)}{32}}\,\mathrm{e}^{-\alpha^2/64}\Big[I_{-1/4}\Big(\frac{\alpha^2}{64}\Big) + I_{1/4}\Big(\frac{\alpha^2}{64}\Big)\Big], \tag{4.9}$$

where $K_{1/4}$ and $I_{\pm 1/4}$ denote the modified Bessel functions of the second and first kind, respectively. In particular,

$$\lim_{\alpha\to+\infty} \Psi_+(\alpha) = 1, \qquad \lim_{\alpha\to+\infty} \Psi_-(\alpha) = 2, \tag{4.10}$$

and

$$\lim_{\alpha\to 0} \Psi_+(\alpha) = \lim_{\alpha\to 0} \Psi_-(\alpha) = \frac{\Gamma(1/4)}{2^{5/4}\sqrt{\pi}} \simeq 0.8600. \tag{4.11}$$

Finally, the error terms satisfy

$$|R_\pm(\varepsilon,\lambda)| \leqslant C\Big[\frac{\varepsilon|\log\varepsilon|^3}{\max\{|\lambda|,(\varepsilon|\log\varepsilon|)^{1/2}\}}\Big]^{1/2}. \tag{4.12}$$

The functions $\Psi_\pm(\alpha)$ are universal, in the sense that they will be the same for all symmetric pitchfork bifurcations, regardless of the details of the system. They are shown in Figure 5. Note in particular that they are not monotonous, but both admit a maximum.

Corollary 4.1. *Assume the normal form at $z = 0$ satisfies* (4.2).

- *If $\lambda_2 > 0$, assume that the minimum of V in one of the path-connected components of $\mathcal{OV}(z)$ is reached at a unique point x, which is quadratic. Let B belong to a different path-connected component of $\mathcal{OV}(z)$, with $G(\{x\},B) = \{z\}$. Then the expected first-hitting time of B satisfies*

$$\mathbb{E}^x\big\{\tau_B\big\} = 2\pi\sqrt{\frac{[\lambda_2+(2\varepsilon C_4)^{1/2}]\lambda_3\dots\lambda_d}{|\lambda_1|\det(\nabla^2 V(x))}} \times \frac{e^{[V(z)-V(x)]/\varepsilon}}{\Psi_+\big(\lambda_2/(2\varepsilon C_4)^{1/2}\big)}\big[1+R_+(\varepsilon,\lambda_2)\big]. \tag{4.13}$$

- *If $\lambda_2 < 0$, assume that the minimum of V in one of the path-connected components of $\mathcal{OV}(z_+) = \mathcal{OV}(z_-)$ is reached at a unique point x, which is quadratic. Let B belong to a different path-connected component of $\mathcal{OV}(z_\pm)$, with $G(\{x\},B) = \{z_+,z_-\}$. Then the expected first-hitting time of B satisfies*

$$\mathbb{E}^x\big\{\tau_B\big\} = 2\pi\sqrt{\frac{[\mu_2+(2\varepsilon C_4)^{1/2}]\mu_3\cdots\mu_d}{|\mu_1|\det(\nabla^2 V(x))}} \times \frac{e^{[V(z_\pm)-V(x)]/\varepsilon}}{\Psi_-\big(\mu_2/(2\varepsilon C_4)^{1/2}\big)}\big[1+R_-(\varepsilon,\mu_2)\big]. \tag{4.14}$$

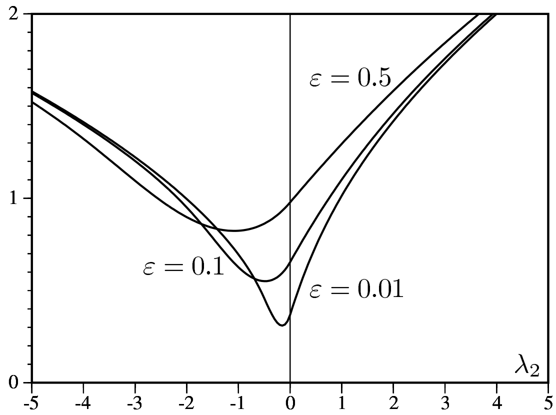

Figure 6. The prefactor of the expected transition time near a pitchfork bifurcation, as a function of the bifurcation parameter λ_2, shown for three different values of ε. (To be precise, we show the function $\lambda_2 \mapsto \sqrt{\lambda_2 + \varepsilon^{1/2}}/\Psi_+(\lambda_2/\varepsilon^{1/2})$ for $\lambda_2 > 0$ and the function $\lambda_2 \mapsto \sqrt{-2\lambda_2 + \varepsilon^{1/2}}/\Psi_-(-2\lambda_2/\varepsilon^{1/2})$ for $\lambda_2 < 0$.)

When λ_2 is bounded away from zero, the expression (4.13) reduces to the usual Eyring – Kramers formula (1.3). When $\lambda_2 \to 0$, it converges to the limiting expression (3.17). The function Ψ_+ controls the crossover between the two regimes, which takes place when λ_2 is of order $\varepsilon^{1/2}$. In fact, when $\lambda_2 \ll \varepsilon^{1/2}$, there is a saturation effect, in the sense that the system behaves as if the curvature of the potential were bounded below by $(2\varepsilon C_4)^{1/2}$. Similar remarks apply to the expression (4.14), the only difference being a factor 1/2 in the prefactor when μ_2 is bounded away from 0 (cf. (4.10)), which is due to the fact that the gate between x and B then contains two saddles.

The λ_2-dependence of the prefactor is shown in Figure 6. It results from the combined effect of the term under the square root and the factors $\Psi_\pm$. Note in particular that the minimal value of the prefactor is located at a negative value of λ_2, which can be shown to be of order $\varepsilon^{1/2}$.

4.2. Longitudinal symmetric pitchfork bifurcation

Consider now the case where for $\gamma = \gamma^\star$, $z = 0$ is a nonquadratic saddle of V, with normal form (3.9). Then a slightly different variant of symmetric pitchfork bifurcation occurs, when for γ near $\gamma^\star$, the normal form has the expression

$$V(y) = \frac{1}{2}\lambda_1(\gamma)y_1^2 - C_4(\gamma)y_1^4 + \frac{1}{2}\sum_{j=2}^{d}\lambda_j(\gamma)y_j^2 + \mathcal{O}(\|y\|_2^5), \tag{4.15}$$

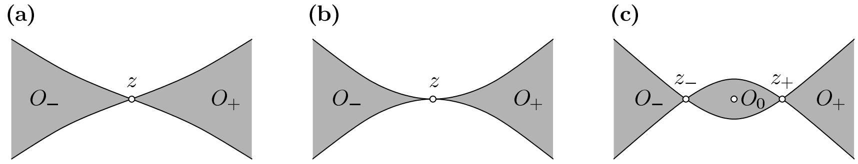

Figure 7. Saddles and open valleys of the normal-form potential (4.15), **(a)** for $\lambda_1 < 0$, **(b)** for $\lambda_1 = 0$ and **(c)** for $\lambda_1 > 0$. The system undergoes a longitudinal pitchfork bifurcation at $\lambda_1 = 0$.

where $\lambda_1(\gamma^\star) = 0$, while $C_4(\gamma^\star) > 0$ and $\lambda_j(\gamma^\star) > 0$ for all $j \geqslant 2$. We assume here that V is even in y_1, which is the most common situation in which pitchfork bifurcations are observed. As before, we no longer indicate the γ-dependence of eigenvalues, and all quantities except λ_1 are assumed to be bounded away from zero.

When $\lambda_1 < 0$, $z = 0$ is a quadratic saddle. When $\lambda_1 > 0$, $z = 0$ is a local minimum, but there exist two saddles $z_\pm$ with coordinates

$$z_\pm = \Big(\pm \sqrt{\lambda_1/4C_4} + \mathcal{O}(\lambda_1), 0, \dots, 0\Big) + \mathcal{O}(\lambda_1^2). \tag{4.16}$$

The open valleys of z_+ and z_- share a path-connected component which we denote by O_0, while we denote their other components by $O_\pm$, cf. Figure 7. Let us denote the eigenvalues of $\nabla^2 V(z_\pm)$ by $\mu_1, \dots, \mu_d$. For $\lambda_1 > 0$ we have

$$\begin{aligned} \mu_1 &= -2\lambda_1 + \mathcal{O}(\lambda_1^{3/2}), \\ \mu_j &= \lambda_j + \mathcal{O}(\lambda_1^{3/2}) \quad \text{for } j \in \{2, \dots, d\}. \end{aligned} \tag{4.17}$$

Finally, the value of the potential on the saddles $z_\pm$ satisfies

$$V(z_+) = V(z_-) = V(z) + \frac{\lambda_1^2}{16C_4} + \mathcal{O}(\lambda_1^{5/2}). \tag{4.18}$$

Such a bifurcation occurs, for instance, in Example 4.1 for $\gamma = 1/3$. Then both saddles z_- and z_+ have to be crossed on any minimal path between the global minima, as in Figure 1c.

We can now state a sharp estimate of the capacity in this situation, which is proved in Section 5.4.

Theorem 4.2. *Assume the normal form at $z = 0$ satisfies* (4.15)*. Let A and B belong to the pathconnected components O_- and O_+ respectively. Assume further that $G(A,B) = \{z\}$ (resp. that $\{z_-\}$ and $\{z_+\}$ both form a gate between A*

and B if $\lambda_1 > 0$). Then for $\lambda_1 < 0$,

$$\operatorname{cap}_A(B) = \sqrt{\frac{(2\pi)^{d-2}[|\lambda_1| + (2\varepsilon C_4)^{1/2}]}{\lambda_2 \dots \lambda_d}} \times \frac{\varepsilon^{d/2}}{\Psi_+\big(|\lambda_1|/(2\varepsilon C_4)^{1/2}\big)} \,\mathrm{e}^{-V(z)/\varepsilon} \big[1 + R_+(\varepsilon, |\lambda_1|)\big], \tag{4.19}$$

while for $\lambda_1 > 0$,

$$\operatorname{cap}_A(B) = \sqrt{\frac{(2\pi)^{d-2}[|\mu_1| + (2\varepsilon C_4)^{1/2}]}{\mu_2 \dots \mu_d}} \times \frac{\varepsilon^{d/2}}{\Psi_-\big(|\mu_1|/(2\varepsilon C_4)^{1/2}\big)} \,\mathrm{e}^{-V(z_\pm)/\varepsilon} \big[1 + R_-(\varepsilon, |\mu_1|)\big]. \tag{4.20}$$

The functions Ψ_+ and Ψ_- and the remainders $R_\pm$ are the same as in (4.9) *and* (4.12).

When $\lambda_1 \gg \varepsilon^{1/2}$, the function Ψ_- in (4.20) is close to 2, so that the total capacity equals half the capacity associated with each saddle z_- and z_+. As in electrostatics, when two capacitors are set up in series, the inverse of their equivalent capacity is thus equal to the sum of the inverses of the individual capacities.

Corollary 4.2. *Assume the normal form at $z = 0$ satisfies* (4.15).

- *If $\lambda_1 < 0$, assume that the minimum of V in one of the path-connected components of $\mathcal{OV}(z)$ is reached at a unique point x, which is quadratic. Let B belong to a different path-connected component of $\mathcal{OV}(z)$, with $G(\{x\}, B) = \{z\}$. Then the expected first-hitting time of B satisfies*

$$\mathbb{E}^x\{\tau_B\} = 2\pi \sqrt{\frac{\lambda_2 \dots \lambda_d}{[|\lambda_1| + (2\varepsilon C_4)^{1/2}] \det(\nabla^2 V(x))}} \,\Psi_+\Big(\frac{|\lambda_1|}{(2\varepsilon C_4)^{1/2}}\Big) \times \mathrm{e}^{[V(z)-V(x)]/\varepsilon} \big[1 + R_+(\varepsilon, |\lambda_1|)\big]. \tag{4.21}$$

- *If $\lambda_1 > 0$, assume that the minimum of V in the path-connected component O_- of $\mathcal{OV}(z_-)$ is reached at a unique point x, which is quadratic. Let B belong to the path-connected component O_+ of $\mathcal{OV}(z_+)$, with $G(\{x\}, B) = \{z_+\}$ or $\{z_-\}$. Assume finally that $V(x)$ is strictly smaller than the minimum of V in the common component O_0 of $\mathcal{OV}(z_-)$ and $\mathcal{OV}(z_+)$. Then the expected first-hitting time of B satisfies*

$$\mathbb{E}^x\{\tau_B\} = 2\pi \sqrt{\frac{\mu_2 \dots \mu_d}{[|\mu_1| + (2\varepsilon C_4)^{1/2}] \det(\nabla^2 V(x))}} \,\Psi_-\Big(\frac{|\mu_1|}{(2\varepsilon C_4)^{1/2}}\Big) \times \mathrm{e}^{[V(z_\pm)-V(x)]/\varepsilon} \big[1 + R_-(\varepsilon, |\mu_1|)\big]. \tag{4.22}$$

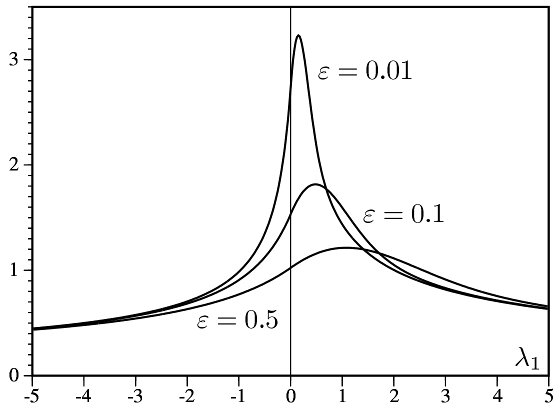

Figure 8. The prefactor of the expected transition time near a longitudinal pitchfork bifurcation, as a function of the bifurcation parameter λ_1, shown for three different values of ε. (To be precise, we show the function $\lambda_1 \mapsto \Psi_+(|\lambda_1|/\varepsilon^{1/2})/\sqrt{|\lambda_1| + \varepsilon^{1/2}}$ for $\lambda_1 < 0$ and the function $\lambda_1 \mapsto \Psi_-(2\lambda_1/\varepsilon^{1/2})/\sqrt{2\lambda_1 + \varepsilon^{1/2}}$ for $\lambda_1 > 0$.)

For $\lambda_1 = 0$, this result reduces to Corollary 3.1. For negative λ_1 of order 1, one recovers the usual Eyring – Kramers formula, while for positive λ_1 of order 1, one obtains the Eyring – Kramers formula with an extra factor 2, which is due to the fact that two saddles have to be crossed in a row. Note that when applying this result to Example 4.1 near $\gamma = 1/3$, all expected transition times have to be divided by 2, because there are two equivalent paths from one global minimum to the other one.

4.3. Bifurcations with double-zero eigenvalue

We consider finally the case where the γ-dependent potential V admits, for $\gamma = \gamma^\star$, $z = 0$ as a nonquadratic saddle with two vanishing eigenvalues. Then the normal form is given by (3.20) with $V_3 = 0$. This singularity having codimension 2, there are several different ways to perturb it, and actually two parameters are needed to describe all of them. We restrict our attention to the particular perturbation (which is generic, e.g., in cases with D_N-symmetry, $N \geqslant 3$)

$$V(y) = \frac{1}{2}\lambda_1(\gamma)y_1^2 + \frac{1}{2}\lambda_2(\gamma)(y_2^2 + y_3^2) + V_4(y_2, y_3; \gamma) + \frac{1}{2}\sum_{j=4}^{d}\lambda_j(\gamma)y_j^2 + \mathcal{O}(\|y\|_2^5), \tag{4.23}$$

where $V_4(y_2, y_3; \gamma)$ is of the form (2.20) with positive discriminant, satisfying $V_{1111} > 0$. Here $\lambda_2(\gamma^\star) = 0$, while $\lambda_1(\gamma^\star) < 0$, $\lambda_4(\gamma^\star) > 0$, and so on. Again, we no longer indicate the γ-dependence of the eigenvalues in the sequel. All quantities except λ_2 are assumed to be bounded away from zero.

Example 4.2. Consider the potential (4.1) for $N \geqslant 3$. In Fourier variables

$$z_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \mathrm{e}^{-ki2\pi/N}\, x_j, \quad -\lfloor N/2 \rfloor < k \leqslant \lfloor N/2 \rfloor, \tag{4.24}$$

the potential takes the form

$$\widehat{V}(z) = \frac{1}{2} \sum_k \eta_k z_k z_{-k} + \frac{1}{4N} \sum_{k_1+k_2+k_3+k_4=0 \ \mathrm{mod}(N)} z_{k_1} z_{k_2} z_{k_3} z_{k_4}, \tag{4.25}$$

where $\eta_k = -1 + 2\gamma \sin^2(k\pi/N)$. For $\gamma > (2\sin^2(\pi/N))^{-1}$, all η_k except $\eta_0 = -1$ are positive, and thus the origin is a quadratic saddle. It forms the gate between the two global minima of the potential, given (in original variables) by $I^\pm = \pm(1, \dots, 1)$. As γ approaches $(2\sin^2(\pi/N))^{-1}$ from above, the two eigenvalues $\eta_{\pm 1}$ go to zero, and the origin becomes a singular saddle of codimension 2. In [3] it is shown that as γ decreases further, a certain number (which depends on N) of quadratic saddles emerges from the origin, all of the same potential height. The set of all these saddles then forms the gate between I^- and I^+.

The normal form of (4.25) is equivalent to (4.23) with $d = N$, up to a linear transformation (one has to work with the real and imaginary part of each z_k instead of the pairs (z_k, z_{-k})), and to a relabelling of the coordinates, setting $\lambda_1 = \eta_0 = -1$ and $\lambda_{2k} = \lambda_{2k+1} = \eta_k$.

We start by considering saddles with normal form (4.23) for $\lambda_2 \geqslant 0$, when the origin is a saddle, and for slightly negative λ_2. Indeed, in these cases one can derive a general result which does not depend on the details of the nonlinear terms. We define $k(\varphi)$, as in (3.22), by $V_4(r\cos\varphi, r\sin\varphi; \gamma) = r^4 k(\varphi; \gamma)$.

Theorem 4.3. *Assume the normal form at $z = 0$ satisfies (4.23). Let A and B belong to different path-connected components of $\mathcal{OV}(z)$ (respectively of the newly created saddles if $\lambda_2 < 0$). Assume further that the gate $G(A, B)$ between A and B is formed by the saddle z in case $\lambda_2 \geqslant 0$, and by the newly created saddles, otherwise. Then for $\lambda_2 \geqslant 0$,*

$$\begin{aligned}\operatorname{cap}_A(B) = {} & \frac{1}{2\pi} \int_0^{2\pi} \sqrt{\frac{(2\pi)^{d-2}|\lambda_1|}{\bigl[\lambda_2 + (2\varepsilon k(\varphi))^{1/2}\bigr]^2 \lambda_4 \dots \lambda_d}}\, \Theta_+\Bigl(\frac{\lambda_2}{(2\varepsilon k(\varphi))^{1/2}}\Bigr)\, \mathrm{d}\varphi \\ & \times \varepsilon^{d/2}\, \mathrm{e}^{-V(z)/\varepsilon} \bigl[1 + R_+(\varepsilon, \lambda_2)\bigr],\end{aligned} \tag{4.26}$$

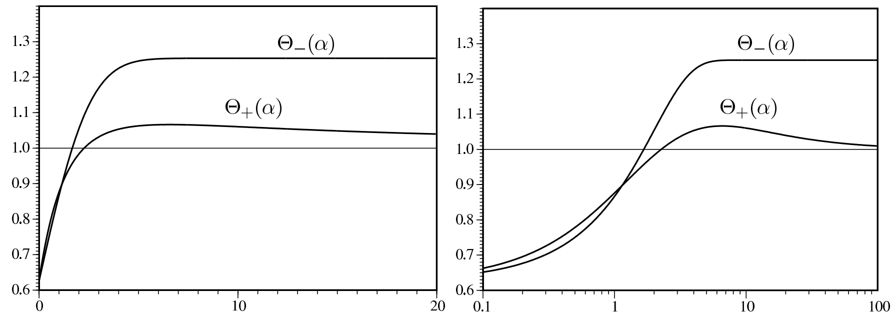

Figure 9. The functions $\Theta_\pm(\alpha)$, shown on a linear and on a logarithmic scale.

while for $-(\varepsilon|\log\varepsilon|)^{1/2} \leqslant \lambda_2 < 0$

$$\operatorname{cap}_A(B) = \frac{1}{2\pi}\int_0^{2\pi} \sqrt{\frac{(2\pi)^{d-2}|\lambda_1|}{\lambda_4\dots\lambda_d}}\, \frac{\Theta_-(-\lambda_2/(2\varepsilon k(\varphi))^{1/2})}{(2\varepsilon k(\varphi))^{1/2}}\, e^{\lambda_2^2/16\varepsilon k(\varphi)}\,d\varphi \times \varepsilon^{d/2}\, e^{-V(z)/\varepsilon}\left[1 + R_-(\varepsilon,\lambda_2)\right]. \tag{4.27}$$

The functions Θ_+ *and* Θ_- *are bounded above and below uniformly on* $\mathbb{R}_+$. *They are given by*

$$\Theta_+(\alpha) = \sqrt{\frac{\pi}{2}}(1+\alpha)\,e^{\alpha^2/8}\,\Phi\Big(-\frac{\alpha}{2}\Big), \tag{4.28}$$

$$\Theta_-(\alpha) = \sqrt{\frac{\pi}{2}}\Phi\Big(\frac{\alpha}{2}\Big), \tag{4.29}$$

where $\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} \exp\{-y^2/2\}\,dy$ *denotes the distribution function of the standard normal law. In particular,*

$$\lim_{\alpha\to+\infty}\Theta_+(\alpha) = 1, \qquad \lim_{\alpha\to+\infty}\Theta_-(\alpha) = \sqrt{\frac{\pi}{2}}, \tag{4.30}$$

and

$$\lim_{\alpha\to0}\Theta_+(\alpha) = \lim_{\alpha\to0}\Theta_-(\alpha) = \sqrt{\frac{\pi}{8}} \simeq 0.6267. \tag{4.31}$$

Finally, the error terms satisfy

$$|R_\pm(\varepsilon,\lambda)| \leqslant C\Big[\frac{\varepsilon|\log\varepsilon|^3}{\max\{|\lambda|,(\varepsilon|\log\varepsilon|)^{1/2}\}}\Big]^{1/2}. \tag{4.32}$$

The proof is given in Section (5.4). The functions $\Theta_\pm(\alpha)$, which are shown in Figure 9, are again universal in the sense that they will be the same for all bifurcations admitting the normal form (4.23).

Corollary 4.3. *Assume the normal form at $z = 0$ satisfies (4.23).*

- *If $\lambda_2 > 0$, assume that the minimum of V in one of the path-connected components of $\mathcal{OV}(z)$ is reached at a unique point x, which is quadratic. Let B belong to a different path-connected component of $\mathcal{OV}(z)$, with $G(\{x\}, B) = \{z\}$. Then the expected first-hitting time of B satisfies*

$$\mathbb{E}^x\{\tau_B\} = 2\pi\sqrt{\frac{\lambda_4 \dots \lambda_d}{|\lambda_1| \det(\nabla^2 V(x))}}\, \mathrm{e}^{[V(z)-V(x)]/\varepsilon} \times \left(\frac{1}{2\pi}\int_0^{2\pi} \frac{\Theta_+(\lambda_2/(2\varepsilon k(\varphi))^{1/2})}{\lambda_2 + (2\varepsilon k(\varphi))^{1/2}}\, \mathrm{d}\varphi\right)^{-1} \bigl[1 + R_+(\varepsilon, \lambda_2)\bigr]. \tag{4.33}$$

- *The above situation extends to slightly negative λ_2, that is, $-\sqrt{\varepsilon|\log\varepsilon|} \leqslant \lambda_2 \leqslant 0$, where*

$$\mathbb{E}^x\{\tau_B\} = 2\pi\sqrt{\frac{\lambda_4 \dots \lambda_d}{|\lambda_1| \det(\nabla^2 V(x))}}\, \mathrm{e}^{[V(z)-V(x)]/\varepsilon} \times \left(\frac{1}{2\pi}\int_0^{2\pi} \frac{\Theta_-(\lambda_2/(2\varepsilon k(\varphi))^{1/2}}{(2\varepsilon k(\varphi))^{1/2}}\, \mathrm{e}^{\lambda_2^2/16\varepsilon k(\varphi)}\, \mathrm{d}\varphi\right)^{-1} \bigl[1 + R_-(\varepsilon, \lambda_2)\bigr]. \tag{4.34}$$

For positive λ_2 of order 1, one recovers the usual Eyring – Kramers law (recall that $\lambda_2 = \lambda_3$), while for $\lambda_2 = 0$, one recovers Corollary 3.3. For negative λ_2, if the function $k(\varphi)$ is nonconstant, the integral over φ in (4.34) can be evaluated by the Laplace method. Extrapolating the result to λ_2 of order -1, the integral would yield an extra $\varepsilon^{1/2}$ which cancels with the $\varepsilon^{1/2}$ in the integral's denominator.

Example 4.3. Consider the potential (4.25) for $N = 3$ or $N \geqslant 5$. The resonant terms near the bifurcation occurring for $\lambda_1 = 0$ are the monomials proportional to $z_1 z_{-1}$ and to $z_1^2 z_{-1}^2$. In order to obtain the standard normal form (4.23), we set $z_{\pm 1} = (y_2 \pm i y_3)/\sqrt{2}$ (the factor $\sqrt{2}$ guarantees that the change of variables is isometric). This yields

$$V_4(y_2, y_3) = \frac{3}{8N}(y_2^2 + y_3^2)^2, \tag{4.35}$$

and thus a constant $k(\varphi) = 3/8N$. The integrals in (4.33) and (4.34) can thus easily be computed. For instance, for odd N and $\lambda_2 = -1 + 2\gamma\sin^2(\pi/N) > 0$, the prefactor of the transition time from I^- to a neighbourhood of I^+ is given by

$$2\pi\frac{\lambda_4\lambda_6 \dots \lambda_{N-1}}{\sqrt{\det \nabla^2 V(I^-)}}\, \frac{\lambda_2 + (3\varepsilon/4N)^{1/2}}{\Theta_+(\lambda_2/(3\varepsilon/4N)^{1/2})}\bigl[1 + R_+(\varepsilon, \lambda_2)\bigr], \tag{4.36}$$

which admits a limit as $\lambda_2 \to 0_+$. Note that the eigenvalues of $\nabla^2 V(I^-)$ are of the form $\nu_1 = 2$, $\nu_{2k} = \nu_{2k+1} = 2 + 2\gamma \sin^2(k\pi/N)$, so that all quantities in (4.36) are known.

Example 4.4. Consider now the potential (4.25) for $N = 4$. Then there are two additional resonant terms, namely the monomials proportional to z_1^4 and z_{-1}^4. Proceeding as in the previous example, this yields

$$V_4(y_2, y_3) = \frac{1}{8}(y_2^4 + y_3^4), \tag{4.37}$$

and thus $k(\varphi) = (3 + \cos(4\varphi))/32$. It is, however, more convenient to keep rectangular coordinates instead of using polar coordinates. Then the coordinates y_2 and y_3 separate, each one undergoing independently a pitchfork bifurcation.

The behaviour for negative $\lambda_2 < -\sqrt{\varepsilon|\log\varepsilon|}$ is determined by the shape of the potential in the (y_2, y_3)-plane. Near the bifurcation point, the potential is sombrero-shaped. As λ_2 decreases, however, the sombrero may develop "dips in the rim", and these dips will determine the value of the integral over φ. We conclude the discussion by an in-depth study of this phenomenon in the case of our model potential.

Example 4.5. Consider again the potential (4.25) for $N = 3$ or $N \geqslant 5$. As seen in Example 4.3 above, the quartic term then is rotation-invariant. This, however, is not sufficient to determine the behaviour for $\lambda_2 < -\sqrt{\varepsilon|\log\varepsilon|}$.

In fact, we know from symmetry arguments [3] that the normal form of the potential around the origin has the form

$$\begin{aligned} V(y) = {} & -\frac{1}{2}|\lambda_1| y_1^2 + \frac{1}{2}\lambda_2 r^2 + \sum_{q=2}^{M} C_{2q} r^{2q} \\ & + D_{2M} r^{2M} \cos(2M\varphi) + \frac{1}{2}\sum_{j=4}^{N} \lambda_j y_j^2 + \mathcal{O}(\|y\|_2^{2M+2}), \end{aligned} \tag{4.38}$$

with $M = N$ if N is odd, and $M = N/2$ if N is even. Here (r, φ) denote polar coordinates for (y_2, y_3). To be precise, the approach presented in [3] is based on a centre-manifold analysis, but the expression of the potential on the centre manifold is equal to the resonant part of the normal form (this is because after the normal-form transformation, the centre manifold is $\|y\|_2^{2M+2}$-close to the (y_2, y_3)-plane).

For $\lambda_2 < 0$, $V(y)$ has nontrivial stationary points of the form $z^\star = (0, r^\star, \varphi^\star, 0, \dots, 0) + \mathcal{O}(\lambda_2^2)$, where

$$r^\star = \sqrt{\frac{-\lambda_2}{4C_4}} + \mathcal{O}(|\lambda_2|^{3/2}), \quad \sin(2M\varphi^\star) = 0. \tag{4.39}$$

The eigenvalues of the Hessian around these points are of the form

$$\begin{aligned}
\mu_1 &= \lambda_1 + \mathcal{O}(|\lambda_2|), \\
\mu_2 &= \frac{1}{r^2}\frac{\partial^2 V}{\partial\varphi^2}(r^\star,\varphi^\star) = \pm(2M)^2 D_{2M}\Big(\frac{-\lambda_2}{4C_4}\Big)^{M-1} + \mathcal{O}(|\lambda_2|^M), \\
\mu_3 &= \frac{\partial^2 V}{\partial r^2}(r^\star,\varphi^\star) = -2\lambda_2 + \mathcal{O}(|\lambda_2|^2), \\
\mu_k &= \lambda_k + \mathcal{O}(|\lambda_2|), \quad k = 4,\dots,N.
\end{aligned} \tag{4.40}$$

The points $z^\star$ with $\mu_2 > 0$ are saddles. They correspond to the "dips on the rim of the sombrero". In the sequel we will therefore assume $\mu_2 > 0$.

Defining the sets A and B as usual, an analogous computation to the previous ones (see Section 5.4 for details) shows that

$$\begin{aligned}
\operatorname{cap}_A(B) = {}& 2M\sqrt{\frac{(2\pi)^{d-2}|\mu_1|}{(\mu_2\mu_3 + (2M)^2 8\varepsilon C_4)\mu_4\dots\mu_N}}\,\Theta_-\Big(\frac{\mu_3}{(8\varepsilon C_4)^{1/2}}\Big) \\
&\times \chi\Big(\frac{\mu_2\mu_3}{(2M)^2 8\varepsilon C_4}\Big)\varepsilon^{d/2}\,\mathrm{e}^{-V(z^\star)/\varepsilon}\big[1 + R_-(\varepsilon,\mu_2)\big],
\end{aligned} \tag{4.41}$$

with $R_-(\varepsilon,\mu_2)$ as in (4.32). Here Θ_- is the function defined in (4.29), and

$$\chi(\alpha) = 2\sqrt{1+\alpha}\,\mathrm{e}^{-\alpha}\,I_0(\alpha), \tag{4.42}$$

where I_0 is the modified Bessel function of the first kind. It satisfies

$$\lim_{\alpha\to 0}\chi(\alpha) = 2, \quad \lim_{\alpha\to\infty}\chi(\alpha) = \sqrt{\frac{2}{\pi}}. \tag{4.43}$$

It follows that if B is a small neighbourhood of I^+, then for $\lambda_2 < 0$,

$$\begin{aligned}
\mathbb{E}^{I^-}\{\tau_B\} = {}& \frac{2\pi}{2M}\sqrt{\frac{(\mu_2\mu_3 + (2M)^2 8\varepsilon C_4)\mu_4\dots\mu_N}{|\mu_1|\det(\nabla^2 V(I^-))}}\,\mathrm{e}^{[V(z^\star)-V(x)]/\varepsilon} \\
&\times \Big(\Theta_-\Big(\frac{\mu_3}{(8\varepsilon C_4)^{1/2}}\Big)\chi\Big(\frac{\mu_2\mu_3}{(2M)^2 8\varepsilon C_4}\Big)\Big)^{-1}\big[1 + R_-(\varepsilon,\mu_2)\big].
\end{aligned} \tag{4.44}$$

The normal-form coefficient C_4 is in fact equal to $3/8N$ (compare (4.35)), which allows to check the continuity of the prefactor at $\lambda_2 = 0$. One can identify three different regimes, depending on the value of $\mu_3 > 0$ (or $\lambda_2 < 0$):

- For $0 \leqslant \mu_3 \ll \sqrt{\varepsilon}$, the arguments of the functions Θ_- and χ are negligible, so that the denominator can be approximated by $\Theta_-(0)\chi(0) = \sqrt{\pi/2}$. Furthermore, the term $(2M)^2 8\varepsilon C_4$ under the square root dominates $\mu_2\mu_3$. This results in a prefactor of order $\varepsilon^{1/2}$, compatible with (4.34) and (4.36). In this regime, the potential in the (y_2,y_3)-plane is almost flat.

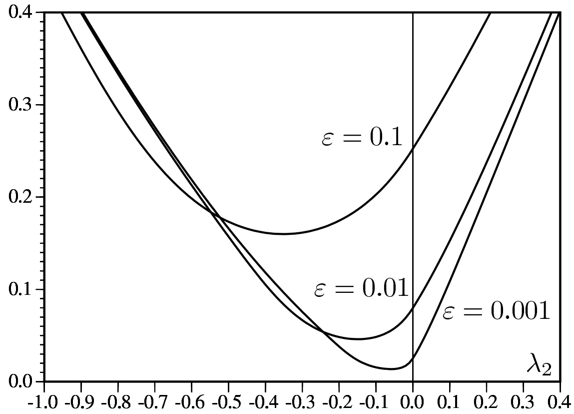

Figure 10. λ_2-dependence of the prefactor of the expected transition time for the model system (4.1) for $N = 3$.

- For $\sqrt{\varepsilon} \ll \mu_3 \ll \varepsilon^{1/M}$, the argument of Θ_- is very large, so that Θ_- can be approximated by $\sqrt{\pi/2}$, while the argument of χ is still negligible. The product of Θ_- and χ is thus close to $\sqrt{2\pi}$. In this regime, the potential in the (y_2, y_3)-plane is sombrero-shaped, and very close to rotation-invariant.
- Finally, for $\mu_3 \gg \varepsilon^{1/M}$, the arguments of both Θ_- and χ are very large, so that their product is close to 1. Furthermore, $\mu_2\mu_3$ under the square root dominates $(2M)^2 8\varepsilon C_4$. In this regime, the potential in the (y_2, y_3)-plane is sombrero-shaped, but with $2M$ dips of noticeable depth. One recovers the standard Eyring – Kramers formula, with an extra prefactor of $1/2M$ accounting for the fact that the gate consists of $2M$ saddles at equal height.

The λ_2-dependence of the prefactor is shown in Figure 10 in the case $N = 3$. In general, the prefactor behaves like λ_2 for $\lambda_2 \gg \sqrt{\varepsilon}$ and like $|\lambda_2|^{M/2}$ for $\lambda_2 \ll -\varepsilon^{1/M}$. Note that as N increases, the second crossover at $\varepsilon^{1/M}$ occurs later and later, as the system becomes closer and closer to being rotation-invariant.

5. Proofs

5.1. Upper bound on the capacity

We assume that the potential V is of class $\mathcal{C}^{r+1}$ for some $r \geqslant 2$, and that the origin 0 is a stationary point with eigenvalues satisfying $\lambda_1 \leqslant 0 < \lambda_{q+1} \leqslant \dots \leqslant \lambda_d$ for some $q \geqslant 2$. The eigenvalues $\lambda_2, \dots, \lambda_q$ lie between λ_1 and λ_{q+1}, but may have an arbitrary sign. We assume that $\lambda_{q+1} \leqslant \dots \leqslant \lambda_d$ are of order

1. In the vicinity of the origin, V admits a normal form

$$V(y) = -u_1(y_1) + u_2(y_2, \dots, y_q) + \frac{1}{2} \sum_{j=q+1}^{d} \lambda_j y_j^2 + \mathcal{O}(\|y\|_2^{r+1}), \tag{5.1}$$

where we may assume that u_1 and u_2 are polynomials of degree less or equal r, and increasing at infinity. Suppose that

- either the origin $z^\star = 0$ itself is a saddle,
- or that there exists a finite number of saddles $z_i^\star$ in the vicinity of $z = 0$.

Let A and B belong to two different path-connected components of $\mathcal{OV}(z^\star)$ or $\mathcal{OV}(z_i^\star)$, respectively; and assume for simplicity that the gate $G(A, B)$ consists of the above-mentioned saddles only. The assumptions below will guarantee that the saddles $z_i^\star$ are close to $z = 0$. However, the results can easily be extended to more complicated situations with several gates which are not necessarily close to the origin, simply by summing the contributions of all gates.

Proposition 5.1 (Upper bound). *Assume there exist strictly positive numbers $\delta_1 = \delta_1(\varepsilon)$, $\delta_2 = \delta_2(\varepsilon)$ and $c \geqslant 0$ (independent of ε) such that*

$$\begin{aligned} u_1(y_1) &\leqslant d\varepsilon|\log\varepsilon| && \textit{whenever } |y_1| \leqslant \delta_1, \\ u_2(y_2, \dots, y_q) &\geqslant -cd\varepsilon|\log\varepsilon| && \textit{whenever } \|(y_2, \dots, y_q)\|_2 \leqslant \delta_2, \\ u_2(y_2, \dots, y_q) &\geqslant 2d\varepsilon|\log\varepsilon| && \textit{whenever } \|(y_2, \dots, y_q)\|_2 \geqslant \delta_2, \end{aligned} \tag{5.2}$$

and such that

$$[\delta_1(\varepsilon) + \delta_2(\varepsilon)]^{r+1} = o(\varepsilon|\log\varepsilon|). \tag{5.3}$$

Then

$$\operatorname{cap}_A(B) \leqslant \varepsilon \frac{\int_{\mathcal{B}_{\delta_2}(0)} \mathrm{e}^{-u_2(y_2,\dots,y_q)/\varepsilon} \, \mathrm{d}y_2 \dots \mathrm{d}y_q}{\int_{-\delta_1}^{\delta_1} \mathrm{e}^{-u_1(y_1)/\varepsilon} \, \mathrm{d}y_1} \prod_{j=q+1}^{d} \sqrt{\frac{2\pi\varepsilon}{\lambda_j}} \big[1 + R_1(\varepsilon)\big] + R_2(\varepsilon), \tag{5.4}$$

where $\mathcal{B}_{\delta_2}(0) = \{(y_2, \dots, y_q) \colon y_2^2 + \dots + y_q^2 < \delta_2^2\}$ and the error terms satisfy

$$\begin{aligned} R_1(\varepsilon) &\leqslant \begin{cases} C\big[\varepsilon^{1/2}|\log\varepsilon|^{3/2} + \varepsilon^{-1}(\delta_1^{r+1} + \delta_2^{r+1}) + \delta_1 + \delta_2\big] & \textit{for } r = 2, \\ C\big[\varepsilon^{1/2}|\log\varepsilon|^{1/2} + \varepsilon^{-1}(\delta_1^{r+1} + \delta_2^{r+1}) + \delta_1 + \delta_2\big] & \textit{for } r > 2, \end{cases} \\ R_2(\varepsilon) &\leqslant C\varepsilon^{3d/4+1}\delta_1^{-2} \leqslant C\varepsilon^{d/2+3/2}\delta_1^{-2} \end{aligned} \tag{5.5}$$

for some constant $C > 0$.

Proof. The proof is adapted from the proof of [2, Theorem 5.1].

Recall that the capacity can be computed as the minimal value of the Dirichlet form $\Phi_{(A\cup B)^c}$, cf. (3.3), which involves an integration over x. In the vicinity of the saddle, we carry out the normal-form transformation of Proposition 2.6 or 2.7 in the integral. This can always be done locally, setting $x = y + \rho \circ g(y)$, where ρ is a smooth cut-off function which is the identity in a small ball of radius Δ, and identically zero outside a larger ball of radius 2Δ. Inside the smaller ball, we may thus assume that the potential is given by (5.1).

Let $\delta = 2\sqrt{(1+c)d\varepsilon|\log\varepsilon|}$. We introduce a set

$$C_\varepsilon = [-\delta_1, \delta_1] \times \mathcal{B}_{\delta_2}(0) \times \prod_{j=q+1}^{d} [-\delta_j, \delta_j], \tag{5.6}$$

where δ_1 and δ_2 satisfy (5.2) and (5.3) and we choose $\delta_j = \delta/\sqrt{\lambda_j}$, $j = q+1, \dots, d$. By assumption, making ε small enough we can construct a layer $\mathcal{S}_\varepsilon$ of width $2\delta_1$, separating the connected components of the open valley $\mathcal{OV}(0)$, and such that $V(y)$ is strictly positive for all $y \in \mathcal{S}_\varepsilon \setminus C_\varepsilon$. Let D_- and D_+ denote the connected components of $\mathbb{R}^d \setminus \mathcal{S}_\varepsilon$ containing A and B respectively. Note that we can choose a radius Δ for the ball $B_\Delta(0)$ in which we carry out the normal form transformation independently of ε in such a way that V is bounded away from zero on $\mathcal{S}_\varepsilon \cap B_\Delta(0)^{\mathrm{c}}$.

The variational principle (3.3) implies that it is sufficient to construct a function $h^+ \in \mathcal{H}_{A,B}$ such that $\Phi_{(A\cup B)^c}(h^+)$ satisfies the upper bound. We choose

$$h^+(y) = \begin{cases} 1 & \text{for } y \in D_-, \\ 0 & \text{for } y \in D_+, \\ f(y_1) & \text{for } y \in C_\varepsilon, \end{cases} \tag{5.7}$$

while $h^+(y)$ is arbitrary for $y \in \mathcal{S}_\varepsilon \setminus C_\varepsilon$, except that we require $\|\nabla h^+\|_2 \leqslant \mathit{const}/\delta_1$. The function $f(y_1)$ is chosen as the solution of the one-dimensional differential equation

$$\varepsilon f''(y_1) - \frac{\partial V}{\partial y_1}(y_1, 0, \dots, 0) f'(y_1) = 0 \tag{5.8}$$

with boundary conditions 1 in $-\delta_1$ and 0 in δ_1, that is,

$$f(y_1) = \frac{\int_{y_1}^{\delta_1} \mathrm{e}^{V(t,0,\dots,0)/\varepsilon}\, \mathrm{d}t}{\int_{-\delta_1}^{\delta_1} \mathrm{e}^{V(t,0,\dots,0)/\varepsilon}\, \mathrm{d}t}. \tag{5.9}$$

Inserting h^+ into the expression (3.2) of the capacity, we obtain two non-vanishing terms, namely the integrals over $\mathcal{S}_\varepsilon \setminus C_\varepsilon$ and over C_ε. The first of these can be bounded as follows. For sufficiently small Δ, Assumptions (5.2) and (5.3) imply that for all $y \in (\mathcal{S}_\varepsilon \setminus C_\varepsilon) \cap B_\Delta(0)$,

$$\begin{aligned}\frac{V(y)}{\varepsilon} &\geqslant -d|\log\varepsilon| - cd|\log\varepsilon| + 2(1+c)d|\log\varepsilon| + \mathcal{O}\big(\varepsilon^{-1}[\delta+\delta_1+\delta_2]^{r+1}\big) \\ &\geqslant \frac{3}{4}d|\log\varepsilon| \end{aligned} \tag{5.10}$$

for sufficiently small ε. On $(\mathcal{S}_\varepsilon \setminus C_\varepsilon) \cap B_\Delta(0)^{\mathrm{c}}$, V is bounded away from zero, and by the assumption of exponentially tight level sets, the contribution from this part of the integral is negligible. It follows that

$$\varepsilon \int_{\mathcal{S}_\varepsilon \setminus C_\varepsilon} \mathrm{e}^{-V(y)/\varepsilon} \frac{const}{\delta_1^2} \,\mathrm{d}y = \mathcal{O}\big(\varepsilon^{3d/4+1}\delta_1^{-2}\big) =: R_2(\varepsilon). \tag{5.11}$$

When calculating the second term which is given by

$$\Phi_{C_\varepsilon}(h^+) = \varepsilon \int_{C_\varepsilon} \mathrm{e}^{-V(y)/\varepsilon} |f'(y_1)|^2 \,\mathrm{d}y = \varepsilon \frac{\int_{C_\varepsilon} \mathrm{e}^{-V(y)/\varepsilon} \mathrm{e}^{2V(y_1,0,\dots,0)/\varepsilon} \,\mathrm{d}y}{\big(\int_{-\delta_1}^{\delta_1} \mathrm{e}^{V(y_1,0,\dots,0)/\varepsilon} \,\mathrm{d}y_1\big)^2}, \tag{5.12}$$

it is sufficient to carry out the normal form transformation in a smaller ball of radius $2\tilde{\Delta}$ with $\tilde{\Delta} = \sqrt{d}\max\{\delta, \delta_1, \delta_2\}$ as $B_{\tilde{\Delta}}(0)$ contains C_ε. The Jacobian of the transformation thus yields a multiplicative error term $1 + \mathcal{O}(\tilde{\Delta})$. By (5.1), we have for $y \in C_\varepsilon$

$$\begin{aligned} V(y) - 2V(y_1, 0, \dots, 0) = u_1(y_1) + u_2(y_2, \dots, y_q) + \frac{1}{2}\sum_{j=q+1}^{d} \lambda_j y_j^2 \\ + \mathcal{O}\big([\delta+\delta_1+\delta_2]^{r+1}\big). \end{aligned} \tag{5.13}$$

Hence the numerator in (5.12) is given by

$$\begin{aligned} \int_{-\delta_1}^{\delta_1} \mathrm{e}^{-u_1(y_1)/\varepsilon} \,\mathrm{d}y_1 \int_{\mathcal{B}_{\delta_2}(0)} \mathrm{e}^{-u_2(y_2,\dots,y_q)/\varepsilon} \,\mathrm{d}y_2 \cdots \mathrm{d}y_q \prod_{j=q+1}^{d} \int_{-\delta_j}^{\delta_j} \mathrm{e}^{-\lambda_j y_j^2/2\varepsilon} \,\mathrm{d}y_j \\ \times \Big[1 + \mathcal{O}\Big(\frac{[\delta+\delta_1+\delta_2]^{r+1}}{\varepsilon}\Big)\Big]. \end{aligned} \tag{5.14}$$

Substituting in (5.12) we get

$$\begin{aligned} \Phi_{C_\varepsilon}(h^+) = \varepsilon \frac{\int_{\mathcal{B}_{\delta_2}(0)} \mathrm{e}^{-u_2(y_2,\dots,y_q)/\varepsilon} \,\mathrm{d}y_2 \dots \mathrm{d}y_q}{\int_{-\delta_1}^{\delta_1} \mathrm{e}^{-u_1(y_1)/\varepsilon} \,\mathrm{d}y_1} \prod_{j=q+1}^{d} \int_{-\delta_j}^{\delta_j} \mathrm{e}^{-\lambda_j y_j^2/2\varepsilon} \,\mathrm{d}y_j \\ \times \Big[1 + \mathcal{O}\Big(\frac{[\delta+\delta_1+\delta_2]^{r+1}}{\varepsilon}\Big)\Big]. \end{aligned} \tag{5.15}$$

Using the fact that the Gaussian integrals over y_j, $j = q+1, \dots, d$, are bounded above by $\sqrt{2\pi\varepsilon/\lambda_j}$, the desired bound (5.4) follows. □

Remark 5.1. Using the conditions (5.2) in order to bound the integrals over y_1 and $(y_2, \dots, y_q)$, one obtains as a rough *a priori* bound

$$\operatorname{cap}_A(B) \leqslant \mathit{const}\, \frac{\delta_2^{q-1}}{\delta_1}\, \varepsilon^{1-q/2-(c+1/2)d}[1 + R_1(\varepsilon)] + R_2(\varepsilon). \tag{5.16}$$

In applications we will of course obtain much sharper bounds by using explicit expressions for $u_1(y_1)$ and $u_2(y_2, \dots, y_q)$, but the above rough bound will be sufficient to obtain a lower bound on the capacity, valid without further knowledge of the functions u_1 and u_2.

5.2. Lower bound on the capacity

Before we proceed to deriving a lower bound on the capacity, we need a crude bound on the equilibrium potential $h_{A,B}$. We obtain such a bound by adapting similar results from [2, Section 4] to the present situation.

Lemma 5.1. *Let A and B be disjoint sets, and let $x \in (A \cup B)^c$ be such that the ball $\mathcal{B}_\varepsilon(x)$ does not intersect $A \cup B$. Then there exists a constant C such that*

$$h_{A,B}(x) \leqslant C\varepsilon^{-d} \operatorname{cap}_{\mathcal{B}_\varepsilon(x)}(A)\, \mathrm{e}^{\overline{V}(\{x\},B)/\varepsilon}\,. \tag{5.17}$$

Proof. [2, Proposition 4.3] provides the upper bound

$$h_{A,B}(x) \leqslant C\frac{\operatorname{cap}_{\mathcal{B}_\varepsilon(x)}(A)}{\operatorname{cap}_{\mathcal{B}_\varepsilon(x)}(B)}, \tag{5.18}$$

so that it suffices to obtain a lower bound for the denominator. This is done as in [2, Proposition 4.7] with $\rho = \varepsilon$, cf. in particular equation (4.26) in that work, which provides a lower bound for the capacity in terms of an integral of $\exp\{V/\varepsilon\}$ over a critical path from x to B. Evaluating the integral by the Laplace method, one gets $\exp\{\overline{V}(\{x\}, B)/\varepsilon\}$ as leading term, with a multiplicative correction. The only difference is that while Bovier *et al.* assume quadratic saddles, which yields a correction of order $\sqrt{\varepsilon}$, here we do not assume anything on the saddles, so that in the worst case the prefactor is constant. This yields the bound (5.17). □

The capacity $\operatorname{cap}_{\mathcal{B}_\varepsilon(x)}(A)$ behaves roughly like $\exp\{-\overline{V}(\{x\}, A)/\varepsilon\}$, so that the bound (5.17) is useful whenever $\overline{V}(\{x\}, A) \gg \overline{V}(\{x\}, B)$. This is the case, in particular, when A and B belong to different path-connected components of the open valley of a saddle z, and x belongs to the same component as B. If, by contrast, x belongs to the same component as A, the symmetry $h_{A,B}(x) =$

$1 - h_{B,A}(x)$ yields a lower bound for the equilibrium potential which is close to 1.

We now consider the same situation as in Section 5.1. Let $\delta_1(\varepsilon)$, $\delta_2(\varepsilon)$ and c be the constants introduced in Proposition 5.1.

Proposition 5.2 (Lower bound). *Let $K > \max\{2d(2+c), d-q\}$, and assume that $\delta_1(\varepsilon) \geqslant \varepsilon$*[4]*. Furthermore, suppose there exist strictly positive numbers $\hat\delta_1 = \hat\delta_1(\varepsilon)$ and $\hat\delta_2 = \hat\delta_2(\varepsilon)$ such that*

$$\begin{aligned} u_1(\pm\hat\delta_1) &\geqslant 4K\varepsilon|\log\varepsilon|, \\ u_2(y_2,\dots,y_q) &\leqslant K\varepsilon|\log\varepsilon|, \quad \textit{whenever } \|(y_2,\dots,y_q)\|_2 \leqslant \hat\delta_2, \end{aligned}$$

and such that

$$\big[\hat\delta_1(\varepsilon) + \hat\delta_2(\varepsilon)\big]^{r+1} = \mathcal{O}(\varepsilon|\log\varepsilon|). \tag{5.19}$$

Then

$$\operatorname{cap}_A(B) \geqslant \varepsilon \frac{\int_{\mathcal{B}_{\hat\delta_2}(0)} \mathrm{e}^{-u_2(y_2,\dots,y_q)/\varepsilon}\, \mathrm{d}y_2\dots\mathrm{d}y_q}{\int_{-\hat\delta_1}^{\hat\delta_1} \mathrm{e}^{-u_1(y_1)/\varepsilon}\,\mathrm{d}y_1} \prod_{j=q+1}^{d} \sqrt{\frac{2\pi\varepsilon}{\lambda_j}}\,[1 - R_3(\varepsilon)], \tag{5.20}$$

where the remainder $R_3(\varepsilon)$ satisfies

$$R_3(\varepsilon) \leqslant \begin{cases} C\big[\varepsilon^{1/2}|\log\varepsilon|^{3/2} + \varepsilon^{-1}(\hat\delta_1^{r+1} + \hat\delta_2^{r+1}) + \hat\delta_1 + \hat\delta_2 + \sqrt{\varepsilon}\big] & \textit{for } r = 2, \\ C\big[\varepsilon^{1/2}|\log\varepsilon|^{1/2} + \varepsilon^{-1}(\hat\delta_1^{r+1} + \hat\delta_2^{r+1}) + \hat\delta_1 + \hat\delta_2 + \sqrt{\varepsilon}\big] & \textit{for } r > 2, \end{cases} \tag{5.21}$$

for some constant $C > 0$.

Proof. As in the proof of Proposition 5.1, we start by locally carrying out the normal form transformation in the integral defining the Dirichlet form. Next we define a slightly different neighbourhood of the saddle,

$$\widehat{C}_\varepsilon = \big[-\hat\delta_1, \hat\delta_1\big] \times \mathcal{B}_{\hat\delta_2}(0) \times \prod_{j=q+1}^{d} \big[-\hat\delta_j, \hat\delta_j\big] = \big[-\hat\delta_1, \hat\delta_1\big] \times \widehat{C}_\varepsilon^\perp, \tag{5.22}$$

where we now choose

$$\hat\delta_j = \frac{\delta}{\sqrt{(d-q)\lambda_j}}, \quad j = q+1,\dots,d, \tag{5.23}$$

with $\delta = \sqrt{K\varepsilon|\log\varepsilon|}$. The reason for this choice is that we want the potential to be smaller than $-\delta^2$ on the "sides" $\{\pm\hat\delta_1\} \times \widehat{C}_\varepsilon^\perp$ of the box. Indeed, we have

[4]It actually suffices to have $\delta_1 \geqslant \varepsilon^{K/2+1/4-d/8}$ which is a very weak condition satisfied whenever $\delta_1 \geqslant \varepsilon^\kappa$ for a $\kappa > 1$ depending on c, d and q.

for all $y_\perp \in \widehat{C}_\varepsilon^\perp$,

$$\begin{aligned}\frac{V(\pm\hat{\delta}_1, y_\perp)}{\varepsilon} &\leqslant -4K|\log\varepsilon| + K|\log\varepsilon| + (d-q)\frac{\delta^2}{2\varepsilon(d-q)} \\ &\quad + \mathcal{O}\big(\varepsilon^{-1}\big[\delta + \hat{\delta}_1 + \hat{\delta}_2\big]^{r+1}\big) \leqslant -K|\log\varepsilon| \end{aligned} \tag{5.24}$$

for sufficiently small ε. As a consequence, if $h^\star = h_{A,B}$ denotes the equilibrium potential, Lemma 5.1 and the *a priori* bound (5.16) yield

$$\begin{aligned} h^\star&(\hat{\delta}_1, y_\perp) \\ &= \mathcal{O}\Big(\varepsilon^{-d}\Big[\frac{\delta_2^{q-1}}{\delta_1}\varepsilon^{1-q/2-(c+1/2)d}[1+R_1(\varepsilon)] + R_2(\varepsilon)\Big]\exp\Big\{\frac{V(\hat{\delta}_1, y_\perp)}{\varepsilon}\Big\}\Big) \\ &= \mathcal{O}\Big(\varepsilon^{-d}\Big[\frac{\mathcal{O}\big((\varepsilon|\log\varepsilon|)^{(q-1)/(r+1)}\big)}{\varepsilon^{K/2+1/4-d/8}}\varepsilon^{1-q/2-(c+1/2)d}[1+|\log\varepsilon|]\Big]\varepsilon^K + \varepsilon^{1/2}\Big) \\ &= \mathcal{O}\big(\varepsilon^{1/2}\big) \end{aligned} \tag{5.25}$$

while

$$h^\star(-\hat{\delta}_1, y_\perp) = 1 - \mathcal{O}\big(\varepsilon^{1/2}\big). \tag{5.26}$$

We can now proceed to deriving the lower bound. Observe that

$$\operatorname{cap}_A(B) = \Phi_{(A\cup B)^c}(h^\star) \geqslant \Phi_{\widehat{C}_\varepsilon}(h^\star). \tag{5.27}$$

Now we can write, for any $h \in \mathcal{H}_{A,B}$,

$$\Phi_{\widehat{C}_\varepsilon}(h) \geqslant \varepsilon \int_{\widehat{C}_\varepsilon} \mathrm{e}^{-V(y)/\varepsilon}\Big(\frac{\partial h}{\partial y_1}\Big)^2 \,\mathrm{d}y = \varepsilon \int_{\widehat{C}_\varepsilon^\perp}\int_{-\hat{\delta}_1}^{\hat{\delta}_1} \mathrm{e}^{-V(y)/\varepsilon}\Big(\frac{\partial h(y_1, y_\perp)}{\partial y_1}\Big)^2 \,\mathrm{d}y_1\,\mathrm{d}y_\perp,$$

and thus

$$\Phi_{\widehat{C}_\varepsilon}(h^\star) \geqslant \varepsilon \int_{\widehat{C}_\varepsilon^\perp}\Big[\inf_{f\colon f(\pm\hat{\delta}_1) = h^\star(\pm\hat{\delta}_1, y_\perp)} \int_{-\hat{\delta}_1}^{\hat{\delta}_1} \mathrm{e}^{-V(y)/\varepsilon} f'(y_1)^2 \,\mathrm{d}y_1\Big]\,\mathrm{d}y_\perp. \tag{5.28}$$

The Euler – Lagrange equation for the variational problem is

$$\varepsilon f''(y_1) - \frac{\partial V}{\partial y_1}(y_1, y_\perp) f'(y_1) = 0 \tag{5.29}$$

with boundary conditions $h^\star(-\hat{\delta}_1, y_\perp)$ in $-\hat{\delta}_1$ and $h^\star(\hat{\delta}_1, y_\perp)$ in $\hat{\delta}_1$, and has the solution

$$f(y_1) = h^\star\big(\hat{\delta}_1, y_\perp\big) - \big[h^\star\big(\hat{\delta}_1, y_\perp\big) - h^\star\big(-\hat{\delta}_1, y_\perp\big)\big]\,\frac{\int_{y_1}^{\hat{\delta}_1}\mathrm{e}^{V(t,y_\perp)/\varepsilon}\,\mathrm{d}t}{\int_{-\hat{\delta}_1}^{\hat{\delta}_1}\mathrm{e}^{V(t,y_\perp)/\varepsilon}\,\mathrm{d}t}. \tag{5.30}$$

As a consequence,

$$f'(y_1) = \big[h^\star\big(\hat\delta_1, y_\perp\big) - h^\star\big(-\hat\delta_1, y_\perp\big)\big] \frac{\mathrm{e}^{V(y_1,y_\perp)/\varepsilon}}{\int_{-\hat\delta_1}^{\hat\delta_1} \mathrm{e}^{V(t,y_\perp)/\varepsilon}\,\mathrm{d}t}, \tag{5.31}$$

so that substitution in (5.28) yields

$$\Phi_{\widehat C_\varepsilon}(h^\star) \geqslant \varepsilon \int_{\widehat C_\varepsilon^\perp} \frac{[h^\star(\hat\delta_1, y_\perp) - h^\star(-\hat\delta_1, y_\perp)]^2}{\int_{-\hat\delta_1}^{\hat\delta_1} \mathrm{e}^{V(t,y_\perp)/\varepsilon}\,\mathrm{d}t}\,\mathrm{d}y_\perp. \tag{5.32}$$

The bounds (5.25) and (5.26) on $h^\star$ show that the numerator is of the form $1 - \mathcal{O}(\varepsilon^{1/2})$. It now suffices to use the normal form (5.1) of the potential, and to carry out integration with respect to $y_{q+1}, \dots y_d$. □

5.3. Non-quadratic saddles

Proof of Theorem 3.1. For the upper bound, it suffices to apply Proposition 5.1 in the case $r = 4$, $q = 2$, $u_1(y_1) = C_4 y_1^4$ and $u_2(y_2) = \lambda_2 y_2^2/2$. The conditions for the upper bound are fulfilled for $\delta_1 = (d\varepsilon|\log\varepsilon|/C_4)^{1/4}$, $\delta_2 = 2(d\varepsilon|\log\varepsilon|/\lambda_2)^{1/2}$ and $c = 0$. This yields error terms $R_1(\varepsilon) = \mathcal{O}(\varepsilon^{1/4}|\log\varepsilon|^{5/4})$ and $R_2(\varepsilon) = \mathcal{O}(\varepsilon^{d/2+1}/|\log\varepsilon|^{1/2})$. The integrals over y_1 and y_2 can be computed explicitly as extending their bounds to $\pm\infty$ only produces a negligible error, and we see that the contribution of $R_2(\varepsilon)$ is also negligible. A matching lower bound is obtained in a completely analogous way, using Proposition 5.2 with $\hat\delta_1$, $\hat\delta_2$ of the same order as δ_1, δ_2, respectively, yielding that $R_3(\varepsilon)$ is of the same order as $R_1(\varepsilon)$. □

Proof of Theorem 3.2. The proof is essentially the same as the previous one, only with the rôles of δ_1 and δ_2 interchanged. □

Proof of Theorem 3.3. We again apply Propositions 5.1 and 5.2, now with $q = 3$, $u_1(y_1) = |\lambda_1| y_1^2/2$ and $u_2(y_2, y_3) = V_4(y_2, y_3)$. The conditions for the upper bound are fulfilled for $\delta_1 = (2d\varepsilon|\log\varepsilon|/|\lambda_1|)^{1/2}$, $\delta_2 = (2d\varepsilon|\log\varepsilon|/K_-)^{1/4}$ and $c = 0$, and similarly for the lower bound. This yields error terms $R_1(\varepsilon) = \mathcal{O}(\varepsilon^{1/4}|\log\varepsilon|^{5/4})$ and $R_2(\varepsilon) = \mathcal{O}(\varepsilon^{(d+1)/2}/|\log\varepsilon|)$. The integral over y_1 can be computed explicitly. Writing the integral over y_2 and y_3 in polar coordinates, and performing the integration over r yields the stated expression. □

5.4. Bifurcations

We decompose the proof of Theorem 4.1 into several steps, dealing with positive and negative λ_2 separately.

Proposition 5.3. *Under the assumptions of Theorem* 4.1, *and for* $\lambda_2 > 0$,

$$\operatorname{cap}_A(B) = \frac{I_{a,\varepsilon}}{(2C_4)^{1/4}} \sqrt{\frac{(2\pi)^{d-3}|\lambda_1|}{\lambda_3 \dots \lambda_d}} \varepsilon^{d/2-1/2} \, \mathrm{e}^{-V(z)/\varepsilon} \big[1 + R_+(\varepsilon)\big], \tag{5.33}$$

where $R_+(\varepsilon)$ *is defined in* (4.12), $I_{a,\varepsilon}$ *is the integral*

$$I_{a,\varepsilon} = \int_{-\infty}^{\infty} \mathrm{e}^{-(x^4+ax^2)/2\varepsilon} \, \mathrm{d}x, \tag{5.34}$$

and $a = \lambda_2/\sqrt{2C_4}$.

Proof. It suffices to apply Propositions 5.1 and 5.2 with $r = 4$, $q = 2$, u_1 a quadratic function of y_1 and u_2 a polynomial of degree 4 in y_2 and $c = 0$. We only need to take some care in the choice of the δ_i. The conditions yield $\delta_1 = (2d\varepsilon|\log\varepsilon|/|\lambda_1|)^{1/2}$ and

$$\delta_2^2 = \frac{-\lambda_2 + \sqrt{\lambda_2^2 + 32dC_4\varepsilon|\log\varepsilon|}}{4C_4}. \tag{5.35}$$

For $\lambda_2 > (\varepsilon|\log\varepsilon|)^{1/2}$, this implies that δ_2 is of order $(\varepsilon|\log\varepsilon|/\lambda_2)^{1/2}$, while for $0 < \lambda_2 \leqslant (\varepsilon|\log\varepsilon|)^{1/2}$, it yields δ_2 of order $(\varepsilon|\log\varepsilon|)^{1/4}$. The expressions of $\hat\delta_1$ and $\hat\delta_2$ are similar. This yields the stated error terms. The integral over y_1 is carried out explicitly, while the integral over y_2 equals $I_{a,\varepsilon}/(2C_4)^{1/4}$, up to a negligible error term. □

Proposition 5.4. *Under the assumptions of Theorem* 4.1, *and for* $\lambda_2 < 0$,

$$\operatorname{cap}_A(B) = \frac{J_{b,\varepsilon}}{(2C_4)^{1/4}} \sqrt{\frac{(2\pi)^{d-3}|\mu_1|}{\mu_3 \dots \mu_d}} \, \varepsilon^{d/2-1/2} \, \mathrm{e}^{-V(z_\pm)/\varepsilon} [1 + R_-(\varepsilon)], \tag{5.36}$$

where $R_-(\varepsilon)$ *is defined in* (4.12), $J_{b,\varepsilon}$ *is the integral*

$$J_{b,\varepsilon} = \int_{-\infty}^{\infty} \mathrm{e}^{-(x^2-b/4)^2/2\varepsilon} \, \mathrm{d}x, \tag{5.37}$$

and $b = \mu_2/\sqrt{2C_4}$.

Proof. First note that for small negative λ_2,

$$u_2(y_2) = C_4\Big(y_2^2 - \frac{\mu_2}{8C_4}\Big)^2 - \frac{\mu_2^2}{64C_4} + \mathcal{O}(|\lambda_2|^{3/2}y_2^2), \tag{5.38}$$

where the constant term corresponds to $V(z_\pm) - V(z)$ (recall $z = 0$). The situation is more difficult than before, because u_2 is not increasing on $\mathbb{R}_+$. When applying Proposition 5.1, we distinguish two regimes.

- For $\mu_2 < (\varepsilon|\log\varepsilon|)^{1/2}$, it is sufficient to choose δ_2 of order $(\varepsilon|\log\varepsilon|)^{1/4}$.
- For $\mu_2 \geqslant (\varepsilon|\log\varepsilon|)^{1/2}$, we cannot apply Proposition 5.1 as is, but first split the integral over y_2 into the integrals over $\mathbb{R}_+$ and over $\mathbb{R}_-$. Each integral is in fact dominated by the integral over an interval of order $(\varepsilon|\log\varepsilon|/\mu_2)^{1/2}$ around the minimum $y_2 = \pm(\mu_2/8C_4)^{1/2}$, so that one can choose δ_2 of that order.

We make a similar distinction between regimes when choosing $\hat{\delta}_2$ in order to apply Proposition 5.2. This yields the stated error terms, and the integrals are treated as before. □

Proof of Theorem 4.1. In order to complete the proof of Theorem 4.1, it remains to examine the integrals $I_{a,\varepsilon}$ and $J_{b,\varepsilon}$. First note that

$$I_{a,\varepsilon} = \sqrt{\frac{2\pi\varepsilon^{1/2}}{1+\alpha}}\,\Psi_+(\alpha), \tag{5.39}$$

where $\alpha = a/\sqrt{\varepsilon}$ and

$$\Psi_+(\alpha) = \sqrt{\frac{1+\alpha}{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-(y^4+\alpha y^2)/2}\,\mathrm{d}y. \tag{5.40}$$

The change of variables $y = z/\sqrt{1+\alpha}$ yields

$$\Psi_+(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\Big\{-\frac{1}{2}\Big[\frac{z^4}{(1+\alpha)^2} + \frac{\alpha z^2}{1+\alpha}\Big]\Big\}\,\mathrm{d}z, \tag{5.41}$$

which allows to show that Ψ_+ is bounded above and below by positive constants, and to compute the limits as $\alpha \to 0$ and $\alpha \to \infty$. The expressions in terms of Bessel functions are obtained by observing that

$$f(\delta) := \int_{-\infty}^{\infty} \exp\Big\{-\frac{1}{2}\Big[y^4 + 2\delta y^2 + \frac{\delta^2}{2}\Big]\Big\}\,\mathrm{d}y = \sqrt{\frac{\delta}{2}}\,K_{1/4}\Big(\frac{\delta^2}{4}\Big), \tag{5.42}$$

because it satisfies the equation $f''(\delta) = (\delta^2/4)f(\delta)$. The other integral is treated in a similar way. □

Proof of Theorem 4.2. For $\lambda_1 < 0$, the proof is analogous to the proof of Theorem 4.1, with the rôles of y_1 and y_2 interchanged. The same applies for positive λ_1 up to order $\sqrt{\varepsilon|\log\varepsilon|}$.

For larger λ_1, Propositions 5.3 and 5.4 have to be slightly adapted:

- For the upper bound, we define a neighbourhood C_ε^+ of z_+ and a neighbourhood C_ε^- of z_- in the usual way. Instead of two regions $D_\pm$, we construct three regions D_-, D_0 and D_+, intersecting respectively O_-, O_0 and O_+ and contained in the corresponding basins of attraction (see Figure 7c). They are separated by layers $S_\varepsilon^\pm$. The function h_+ is then defined to be equal to 1 in D_-, to 1/2 in D_0 and to 0 in D_+. Inside the boxes $C_\varepsilon^\pm$, h_+ is constructed in a similar way as before, only with different boundary conditions. This yields a factor 1/2 in the capacity.
- For the lower bound, we first construct boxes $\widehat{C}_\varepsilon^\pm$ around the saddles $z^\pm$ in the same way as before. Then we connect $\widehat{C}_\varepsilon^+$ and $\widehat{C}_\varepsilon^-$ by a tube staying inside O_0 (whose cross-section is of the same size as the sides of the boxes). One can define coordinates $(y_1, \dots, y_d)$, given by the normal-form transformation inside the boxes, and such that y_1 runs along the length of the tube, in such a way that the Jacobian of the transformation $x \mapsto y$ is close to 1. Lemma 5.1 is still applicable, and yields *a priori* bounds on the equilibrium potential on the sides of the boxes not touching the tube (that is, contained in $O_\pm$). The Dirichlet form is then bounded below by restricting the domain of integration to the union of the boxes and the connecting tube. The remainder of the proof is similar, except that the range of y_1 is larger. The value of the integral is dominated by the contributions of the two boxes.

□

Proof of Theorem 4.3. We first consider the case $\lambda_2 \geqslant 0$. It suffices to apply Propositions 5.1 and 5.2, taking some care in the choice of the δ_i. The conditions yield $\delta_1 = (2d\varepsilon|\log\varepsilon|/|\lambda_1|)^{1/2}$ and

$$\delta_2^2 = \frac{-\lambda_2 + \sqrt{\lambda_2^2 + 32dK_-\varepsilon|\log\varepsilon|}}{4K_-}. \tag{5.43}$$

For $\lambda_2 > (\varepsilon|\log\varepsilon|)^{1/2}$, this implies that δ_2 has order $(\varepsilon|\log\varepsilon|/\lambda_2)^{1/2}$, while for $0 \leqslant \lambda_2 \leqslant (\varepsilon|\log\varepsilon|)^{1/2}$, it yields δ_2 of order $(\varepsilon|\log\varepsilon|)^{1/4}$. The expressions of $\hat\delta_1$ and $\hat\delta_2$ are similar. This yields the stated error terms. The integral over y_1 is carried out explicitly, while the integral over y_2 and y_3 gives, using polar coordinates,

$$\begin{aligned}\int_{\mathcal{B}_{\delta_2}(0)} \mathrm{e}^{-u_2(y_2,y_3)/\varepsilon}\,\mathrm{d}y_2\,\mathrm{d}y_3 &= \int_0^{2\pi}\int_0^{\delta_2} \mathrm{e}^{-(\lambda_2 r^2 + 2k(\varphi)r^4)/2\varepsilon}\, r\,\mathrm{d}r\,\mathrm{d}\varphi \\ &= \int_0^{2\pi} \frac{\sqrt{\varepsilon}}{\sqrt{2k(\varphi)}} \int_0^{(2k(\varphi)/\varepsilon)^{1/4}\delta_2} \mathrm{e}^{-(y^4+\alpha(\varphi)y^2)/2}\, y\,\mathrm{d}y\,\mathrm{d}\varphi,\end{aligned} \tag{5.44}$$

where $\alpha(\varphi) = \lambda_2/\sqrt{2\varepsilon k(\varphi)}$. Using the fact that

$$\int_0^\infty \mathrm{e}^{-(y^2+d)^2/2}\, y \,\mathrm{d}y = \frac{1}{2}\int_d^\infty \mathrm{e}^{-z^2/2}\,\mathrm{d}z = \sqrt{\frac{\pi}{2}}\Phi(-d), \tag{5.45}$$

a straightforward computation shows that the integral over y is approximated by

$$\frac{\Theta_+(\alpha(\varphi))}{1+\alpha(\varphi)}. \tag{5.46}$$

For small negative λ_2, we can write

$$u_2(y_2, y_3) = k(\varphi)\Big(r^2 - \frac{|\lambda_2|}{4k(\varphi)}\Big)^2 - \frac{\lambda_2^2}{16k(\varphi)}, \tag{5.47}$$

where the constant term corresponds to the actual minimum of the potential.

If $-(\varepsilon|\log\varepsilon|)^{1/2} < \lambda_2 < 0$, it suffices to apply Propositions 5.1 and 5.2 with δ_2 of order $(\varepsilon|\log\varepsilon|)^{1/4}$. □

Proof of (4.41). The main task is to compute the integral related to u_2, namely

$$\begin{aligned}\mathcal{J} := \int \exp\Big\{ -\frac{\lambda_2 r^2 + 2\sum_{q=2}^{M} C_{2q} r^{2q}}{2\varepsilon}\Big\} \\ \times \int_0^{2\pi} \exp\Big\{ -\frac{D_{2M} r^{2M}\cos(2M\varphi)}{\varepsilon}\Big\}\,\mathrm{d}\varphi r\,\mathrm{d}r.\end{aligned} \tag{5.48}$$

We first carry out the integral over φ, yielding

$$\int_0^{2\pi} \exp\Big\{ -\frac{D_{2M} r^{2M}\cos(2M\varphi)}{\varepsilon}\Big\}\,\mathrm{d}\varphi = 2\pi I_0\Big(\frac{D_{2M} r^{2M}}{\varepsilon}\Big), \tag{5.49}$$

where I_0 is the modified Bessel function

$$I_0(\alpha) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{\alpha\cos\varphi}\,\mathrm{d}\varphi. \tag{5.50}$$

The Laplace method shows that for large α, $I_0(\alpha)$ behaves like $\mathrm{e}^{\alpha}/\sqrt{2\pi\alpha}$. We thus introduce the bounded function

$$\chi(\alpha) = 2\sqrt{1+\alpha}\,\mathrm{e}^{-\alpha} I_0(\alpha) = \frac{\sqrt{1+\alpha}}{\pi}\int_0^{2\pi} \mathrm{e}^{-\alpha(1-\cos\varphi)}\,\mathrm{d}\varphi. \tag{5.51}$$

Inserting in (5.48) and performing the change of variable $r^2 = (\varepsilon/2C_4)^{1/2}u$ yields

$$\mathcal{J} = \frac{\pi}{\sqrt{8C_4}} \int \exp\Big\{ -\frac{u^2 - 2(|\lambda_2|/\sqrt{8C_4})u + \mathcal{O}(u^3)}{2\varepsilon} \Big\} \times \frac{\chi(D_{2M}\varepsilon^{-1}(u/(2C_4)^{1/2})^M)}{\sqrt{1 + D_{2M}\varepsilon^{-1}(u/(2C_4)^{1/2})^M}} \, \mathrm{d}u. \tag{5.52}$$

Applying the Laplace method shows that the integral is dominated by u close to $u^\star = |\lambda_2|/(8C_4)^{1/2}$. Relating the obtained expression to the eigenvalues at the new saddles via the relation

$$\frac{D_{2M}}{\varepsilon}\Big(\frac{u^\star}{(2C_4)^{1/2}}\Big)^M = \frac{\mu_2\mu_3}{(2M)^2}\frac{1+\mathcal{O}(\lambda_2)}{8\varepsilon C_4} \tag{5.53}$$

yields the necessary control on $\mathcal{J}$. The stated formula for the capacity follows. □

A. Normal forms

Proof of Proposition 2.6. Let us denote by $\mathcal{G}_k(n,m)$ the vector space of functions $g\colon \mathbb{R}^n \to \mathbb{R}^m$ which are homogeneous of degree k (i.e., $g(tx) = t^k g(x)\ \forall x$). We write the Taylor series of V in the form

$$V(x) = V_2(x) + V_3(x) + V_4(x) + \mathcal{O}(\|x\|_2^4), \tag{A.1}$$

where $V_k \in \mathcal{G}_k(d,1)$ for $k = 2,3,4$. We first look for a function $g_2 \in \mathcal{G}_2(d,d)$ such that $V \circ [\mathrm{id} + g_2]$ contains as few terms of order 3 as possible. The Taylor series of $V \circ [\mathrm{id} + g_2]$ can be written

$$V(x + g_2(x)) = V_2(x) + \underbrace{\nabla V_2(x)\cdot g_2(x) + V_3(x)}_{\text{order 3}} + \underbrace{V_2(g_2(x)) + \nabla V_3(x)\cdot g_2(x) + V_4(x)}_{\text{order 4}} + \mathcal{O}(\|x\|_2^4). \tag{A.2}$$

Now consider the so-called adjoint map $T\colon \mathcal{G}_2(d,d) \to \mathcal{G}_3(d,1)$, $g_2 \mapsto \nabla V_2(\cdot)\cdot g_2$, seen as a linear map between vector spaces. All terms of $V_3(x)$ in the image of T can be eliminated by a suitable choice of g_2. Let e_l denote the lth vector in the canonical basis of $\mathbb{R}^d$. We see that $T(x_jx_ke_l) = \lambda_l x_jx_kx_l \neq 0$ for $l = 2,\dots,d$. Thus all monomials except x_1^3 are in the image of T. Since T involves multiplication by x_2 or x_3 or $\dots$ or x_d, however, x_1^3 is not in the image of T. Hence this term is resonant. We can thus choose g_2 in such a way that

$$V(x + g_2(x)) = V_2(x) + \underbrace{V_{111}x_1^3}_{\text{order 3}} + \underbrace{V_2(g_2(x)) + \nabla V_3(x)\cdot g_2(x) + V_4(x)}_{\text{order 4}} + \mathcal{O}(\|x\|_2^4). \tag{A.3}$$

Now a completely analogous argument shows that we can construct a function $g_3 \in \mathcal{G}_3(d,d)$ such that $V \circ [\mathrm{id} + g_2] \circ [\mathrm{id} + g_3]$ has some constant times x_1^4 as the only term of order 4. It remains to determine this constant. From (A.3) we deduce that it has the expression

$$C_4 = \frac{1}{2}\sum_{j=2}^{d} \lambda_j (g_{11}^j)^2 + \sum_{j=1}^{d} V_{11j} g_{11}^j + V_{1111}, \tag{A.4}$$

where g_{11}^j denotes the coefficient of $x_1^2 e_j$ in g_2. The expression of T shows that necessarily $g_{11}^j = -V_{11j}/\lambda_j$ for $j = 2, \dots, d$, while we may choose $g_{11}^1 = 0$. This yields (2.14). □

Acknowledgements

We would like to thank Bastien Fernandez for helpful discussions and Anton Bovier for providing a preliminary version of [1]. BG thanks the MAPMO, Orléans, and NB the CRC 701 *Spectral Structures and Topological Methods in Mathematics* at the University of Bielefeld, for kind hospitality. Financial support by the French Ministry of Research, by way of the *Action Concertée Incitative* (*ACI*) *Jeunes Chercheurs, Modélisation stochastique de systèmes hors équilibre*, and the German Research Council (DFG), by way of the CRC 701 *Spectral Structures and Topological Methods in Mathematics*, is gratefully acknowledged.

References

[1] F. Barret, A. Bovier and S. Méléard (2010) Uniform estimates for metastable transition times in a coupled bistable system. *Electronic Journal of Probability* **15**, 12, 323–345.

[2] A. Bovier, M. Eckhoff, V. Gayrard and M. Klein (2004) Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Euro. Math. Soc.* **6** (4), 399–424.

[3] N. Berglund, B. Fernandez and B. Gentz (2007) Metastability in interacting nonlinear stochastic differential equations: I. From weak coupling to synchronization. *Nonlinearity* **20** (11), 2551–2581.

[4] N. Berglund, B. Fernandez and B. Gentz (2007) Metastability in interacting nonlinear stochastic differential equations II: Large-N behaviour. *Nonlinearity* **20** (11), 2583–2614.

[5] A. Bovier, V. Gayrard and M. Klein (2005) Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Euro. Math. Soc.* **7** (1), 69–99.

[6] H. Eyring (1935) The activated complex in chemical reactions. *J. Chem. Phys.* **3**, 107–115.

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

[8] B. Helffer, M. Klein and F. Nier (2004) Quantitative analysis of metastability in reversible diffusion processes via a Witten complex approach. *Mat. Contemp.* **26**, 41–85.

[9] B. Helffer and F. Nier (2005) *Hypoelliptic Estimates and Spectral Theory for Fokker – Planck operators and Witten Laplacians.* Lect. Notes Math. **1862**, Springer-Verlag, Berlin.

[10] B. Helffer and J. Sjöstrand (1984) Multiple wells in the semiclassical limit. I. *Commun. Partial Diff. Eq.* **9** (4), 337–408.

[11] B. Helffer and J. Sjöstrand (1985) Multiple wells in the semiclassical limit. III. Interaction through nonresonant wells. *Math. Nachr.* **124**, 263–313.

[12] B. Helffer and J. Sjöstrand (1985) Puits multiples en limite semi-classique. II. Interaction moléculaire. Symétries. Perturbation. *Ann. Inst. H. Poincaré, Phys. Theorique* **42** (2), 127–212.

[13] B. Helffer and J. Sjöstrand (1985) Puits multiples en mécanique semi-classique. IV. Étude du complexe de Witten. *Commun. Partial Diff. Eq.* **10** (3), 245–340.

[14] V.N. Kolokoltsov (2000) *Semiclassical Analysis for Diffusions and Stochastic Processes.* Lect. Notes Math. **1724**, Springer-Verlag, Berlin.

[15] H.A. Kramers (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–304.

[16] D.L. Stein (2005) Large fluctuations, classical activation, quantum tunneling, and phase transitions. *Braz. J. Phys.* **35**, 242–252.

[17] A.D. Ventcel′ and M.I. Freĭdlin (1969) Small random perturbations of a dynamical system with stable equilibrium position. *Dokl. Akad. Nauk SSSR* **187**, 506–509.

[18] A.D. Ventcel′ and M.I. Freĭdlin (1970) Small random perturbations of dynamical systems. *Usp. Mat. Nauk* **25** (1) (151), 3–55.