

*An Averaging Approach to the
Smoluchowski–Kramers Approximation in
the Presence of a Varying Magnetic Field*

Sandra Cerrai, Jan Wehr & Yichun Zhu

Journal of Statistical Physics

ISSN 0022-4715

Volume 181

Number 1

J Stat Phys (2020) 181:132–148

DOI 10.1007/s10955-020-02570-8

Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".



An Averaging Approach to the Smoluchowski–Kramers Approximation in the Presence of a Varying Magnetic Field

Sandra Cerrai¹ · Jan Wehr² · Yichun Zhu¹

Received: 22 March 2020 / Accepted: 15 May 2020 / Published online: 9 June 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

We study the small mass limit of the equation describing planar motion of a charged particle of a small mass μ in a force field, containing a magnetic component, perturbed by a stochastic term. We regularize the problem by adding a small friction of intensity $\epsilon > 0$. We show that for all small but fixed frictions the small mass limit of $q_{\mu, \epsilon}$ gives the solution q_ϵ to a stochastic first order equation, containing a noise-induced drift term. Then, by using a generalization of the classical averaging theorem for Hamiltonian systems by Freidlin and Wentzell, we take the limit of the slow component of the motion q_ϵ and we prove that it converges weakly to a Markov process on the graph obtained by identifying all points in the same connected components of the level sets of the magnetic field intensity function.

Keywords Stochastic differential equations · Smoluchowski–Kramers approximation · Averaging principle · Hamiltonian systems · Stochastic equations on graphs

1 Introduction

We are dealing with planar motion of a charged particle of a small mass μ in a force field, containing a magnetic component, perturbed by a stochastic term

Communicated by Eric A. Carlen.

S. Cerrai: Partially supported by the NSF Grants DMS 1407615 and DMS 1712934. J. Wehr: Partially supported by the NSF Grants DMS 1615045 and DMS 1911358.

✉ Sandra Cerrai
cerrai@math.umd.edu

Jan Wehr
wehr@math.arizona.edu

Yichun Zhu
yzhu00@math.umd.edu

¹ University of Maryland, College Park, USA

² University of Arizona, Tucson, USA

$$\begin{cases} \mu \ddot{q}_\mu(t) = b(q_\mu(t)) - \lambda(q_\mu(t))A\dot{q}_\mu(t) + \sigma(q_\mu(t))\dot{w}_t, \\ q_\mu(0) = q \in \mathbb{R}^2, \quad \dot{q}_\mu(0) = p \in \mathbb{R}^2. \end{cases} \tag{1.1}$$

Here b is a vector field in \mathbb{R}^2 , σ is 2×2 -matrix-valued mapping defined on \mathbb{R}^2 and $w(t)$ is a standard two-dimensional Brownian motion. Moreover, $\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a mapping, such that $\lambda(x) \geq \lambda_0 > 0$, for every $x \in \mathbb{R}^2$, and

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We are interested in understanding the limiting behavior of the solution q_μ to the equation (1.1), as the mass μ converges to 0. This is the so called Smoluchowski-Kramers limit.

It is well known (see [10] for all details) that when the variable magnetic field considered in the present paper is replaced by a constant friction (that is λ is constant and the matrix A is the identity matrix), then $q_\mu(t)$ converges to the solution of the first order equation

$$dq(t) = b(q(t))dt + \sigma(q(t))dw(t), \quad q(0) = q. \tag{1.2}$$

More precisely, for every fixed $T > 0$

$$\lim_{\mu \rightarrow 0} \mathbb{E} \max_{t \in [0, T]} |q_\mu(t) - q(t)|^2 = 0. \tag{1.3}$$

Here q_μ can be a vector of any finite dimension. The same result can be obtained also if λ is still constant, but A is a more general matrix, whose eigenvalues have strictly positive real parts, with the limiting equation (1.2) replaced by

$$dq(t) = A^{-1}b(q(t))dt + A^{-1}\sigma(q(t))dw(t), \quad q(0) = q. \tag{1.4}$$

The case of non constant friction has been widely studied recently (see [2] and [14] for example). To summarize the relevant results, consider the system

$$\begin{cases} \mu \ddot{q}_\mu(t) = b(q_\mu(t)) - \gamma(q_\mu(t))\dot{q}_\mu(t) + \sigma(q_\mu(t))\dot{w}_t, \\ q_\mu(0) = q \in \mathbb{R}^k, \quad \dot{q}_\mu(0) = p \in \mathbb{R}^k, \end{cases} \tag{1.5}$$

for an h -dimensional Brownian motion $w(t)$. Assume that the coefficients $b : \mathbb{R}^k \rightarrow \mathbb{R}^k$, $\gamma : \mathbb{R}^k \rightarrow \mathbb{R}^{k \times k}$ and $\sigma : \mathbb{R}^k \rightarrow \mathbb{R}^{h \times k}$ are smooth and uniformly bounded and the smallest eigenvalue $\lambda_1(q)$ of the symmetric matrix $\gamma(q) + \gamma^*(q)$ is strictly positive, uniformly in $q \in \mathbb{R}^k$,

$$\inf_{q \in \mathbb{R}^k} \lambda_1(q) =: \bar{\lambda} > 0.$$

In this case the relation (1.3) still holds, but now $q(t)$ is the solution of the modified equation

$$dq(t) = [\gamma^{-1}(q(t))b(q(t)) + S(q(t))]dt + \gamma^{-1}(q(t))\sigma(q(t))dw(t), \quad q(0) = q. \tag{1.6}$$

Here $S(q)$ is the *noise-induced drift* whose j -th component equals

$$S_j(q) = \sum_{i, l=1}^k \frac{\partial}{\partial q_i} (\gamma^{-1})_{jl}(q) J_{li}(q), \quad j = 1, \dots, k,$$

where J is the matrix-valued function solving the Lyapunov equation

$$J(q)\gamma^*(q) + \gamma(q)J(q) = \sigma(q)\sigma^*(q), \quad q \in \mathbb{R}^k.$$

In [5], the case of a particle subject to a constant magnetic field orthogonal to the plane where the particle moves has been considered. In this case, the motion of the particle is governed by equation (1.1), with $\lambda(q) \equiv \bar{\lambda}$, for every $q \in \mathbb{R}^2$ (for simplicity of notation in what follows we shall take $\bar{\lambda} = 1$). In particular, since the eigenvalues of A are purely imaginary, the methods and results described above are not applicable in this case.

It is not difficult to check that if the stochastic term in (1.1) is replaced by a continuous function, then q_μ converges uniformly in $[0, T]$ to the solution of (1.4). But for the white noise the result is different: while

$$\lim_{\mu \rightarrow 0} \int_0^t \sin \frac{s}{\mu} \varphi(s) ds = 0,$$

for every continuous function φ , when $w(t)$ is a Brownian motion we have

$$\text{Var} \left(\int_0^t \sin \frac{s}{\mu} dw(s) \right) = \int_0^t \sin^2 \frac{s}{\mu} ds \rightarrow \frac{t}{2}, \quad \text{as } \mu \downarrow 0,$$

so that

$$\lim_{\mu \rightarrow 0} \int_0^t \sin \frac{s}{\mu} dw(s) \neq 0.$$

Because of this, in [5] the problem has been regularized, in order to prove a suitable analog of the Smoluchowski-Kramers approximation. The first regularization consisted in introducing in equation (1.1) a small friction proportional to the velocity. Namely, the following equation has been considered

$$\begin{cases} \mu \ddot{q}_{\mu,\epsilon}(t) = b(q_{\mu,\epsilon}(t)) - A_\epsilon \dot{q}_{\mu,\epsilon}(t) + \sigma(q_{\mu,\epsilon}(t))\dot{w}(t), \\ q_{\mu,\epsilon}(0) = q \in \mathbb{R}^2, \quad \dot{q}_{\mu,\epsilon}(0) = p \in \mathbb{R}^2, \end{cases}$$

where $A_\epsilon = A + \epsilon I$ and $\epsilon > 0$ is a small parameter. It has been shown that for any $T > 0$

$$\lim_{\mu \rightarrow 0} \mathbb{E} \max_{t \in [0, T]} |q_{\mu,\epsilon}(t) - q_\epsilon(t)|^2 = 0, \tag{1.7}$$

where $q_\epsilon(t)$ is the solution of the problem

$$dq(t) = A_\epsilon^{-1} b(q(t)) dt + A_\epsilon^{-1} \sigma(q(t)) dw(t), \quad q(0) = q.$$

Next, it has been shown that

$$\lim_{\epsilon \rightarrow 0} \mathbb{E} \max_{t \in [0, T]} |q_\epsilon(t) - q(t)|^2 = 0,$$

where $q(t)$ is the solution of the problem

$$dq(t) = -A b(q(t)) dt - A \sigma(q(t)) dw(t), \quad q(0) = q. \tag{1.8}$$

Another approach to regularization (see also [16] for the case of nonconstant magnetic field) used the fact that the white noise $\dot{w}(t)$ can be considered as an idealization of an isotropic δ -correlated smooth mean-zero Gaussian process $\dot{w}^\delta(t)$, with $0 < \delta < 1$, which converges to $\dot{w}(t)$, as $\delta \downarrow 0$. In this case, it has been proven that if $q_{\mu,\delta}(t)$ is the solution of equation (1.1), with $\dot{w}(t)$ replaced by $\dot{w}^\delta(t)$, then

$$\lim_{\mu \rightarrow 0} \mathbb{E} \max_{t \in [0, T]} |q_{\mu,\delta}(t) - q_\delta(t)| = 0,$$

where $q_\delta(t)$ solves the equation

$$\dot{q}(t) = -Ab(q(t)) - A\sigma(q(t)) \dot{w}^\delta(t), \quad q(0) = q.$$

Next, by taking the limit as $\delta \downarrow 0$, it has been proven that $q_\delta(t)$ converges to the solution $\hat{q}(t)$ of the problem

$$d\hat{q}(t) = -Ab(\hat{q}(t)) dt - A\sigma(\hat{q}(t)) \circ dw(t), \quad \hat{q}(0) = q,$$

where the stochastic term is now interpreted in Stratonovich sense.

In the present paper we are interested in the small mass limit in the presence of a nonconstant magnetic field. To this purpose we add a small constant friction and we consider the regularized equation

$$\begin{cases} \mu \ddot{q}_{\mu,\epsilon}(t) = b(q_{\mu,\epsilon}(t)) - [\lambda(q_{\mu,\epsilon}(t))A + \epsilon I] \dot{q}_{\mu,\epsilon}(t) + \sigma(q_{\mu,\epsilon}(t)) \dot{w}_t, \\ q_{\mu,\epsilon}(0) = q \in \mathbb{R}^2, \quad \dot{q}_{\mu,\epsilon}(0) = p \in \mathbb{R}^2. \end{cases} \tag{1.9}$$

We show that under suitable conditions on the coefficients b , σ and λ , the above problem is well posed in $L^k(\Omega; C([0, T]; \mathbb{R}^2))$, for every $T > 0$ and $k \geq 1$.

For every fixed $\epsilon > 0$, equation (1.9) is of the same type as those considered in [13] and [14], so that we can take the small mass limit as μ goes to zero, obtaining

$$\lim_{\mu \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} |q_{\mu,\epsilon}(t) - q_\epsilon(t)| = 0,$$

where q_ϵ is the solution of the problem

$$\begin{cases} dq_\epsilon(t) = [(\lambda(q_\epsilon(t))A + \epsilon I)^{-1} b(q_\epsilon(t)) + S_\epsilon(q_\epsilon(t))] dt + (\lambda(q_\epsilon(t))A + \epsilon I)^{-1} \sigma(q_\epsilon(t)) dw(t), \\ q_\epsilon(0) = q. \end{cases}$$

After some computations, we obtain

$$\begin{aligned} dq_\epsilon(q) &= \frac{1}{\epsilon} \gamma(q_\epsilon(t)) \nabla^\perp \lambda(q_\epsilon(t)) dt + B(q_\epsilon(t)) dt + \Sigma(q_\epsilon(t)) dw(t), \\ &+ \epsilon [B_\epsilon(q_\epsilon(t)) dt + \Sigma_\epsilon(q_\epsilon(t)) dw(t)], \quad q_\epsilon(0) = q, \end{aligned}$$

where the mappings $\gamma : \mathbb{R}^2 \rightarrow \mathbb{R}$, $B, B_\epsilon : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ and $\Sigma, \Sigma_\epsilon : \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$ can be explicitly computed. This means that the motion of q_ϵ is a superposition of a fast component on the level sets of λ and a slow transversal motion. Using a suitable generalization of the classical result of Freildin and Wentzell on averaging for Hamiltonian systems (see [12, Chapter 8] and [26]), we shall prove that the projection of q_ϵ over the graph Γ , obtained by identifying all points on the same connected component of each level set of λ , converges to a Markov process Y , whose generator is explicitly given.

2 Well-Posedness of the Regularized Problem

As mentioned in the Introduction, we are dealing here with the equation

$$\begin{cases} \mu \ddot{q}_\mu(t) = b(q_\mu(t)) - \lambda(q_\mu(t))A \dot{q}_\mu(t) + \sigma(q_\mu(t)) \dot{w}_t, \\ q_\mu(0) = q \in \mathbb{R}^2, \quad \dot{q}_\mu(0) = p \in \mathbb{R}^2, \end{cases} \tag{2.1}$$

where μ is a small positive constant and $w(t)$ is a standard Brownian motion in \mathbb{R}^2 .

In this section, we shall assume that the coefficients of the equation above satisfy the following assumptions. In Sect. 4 we will impose a more restrictive growth condition on λ .

Hypothesis 1 1. The mappings $b : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ and $\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$ are Lipschitz-continuous.

2. The mapping $\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ is locally Lipschitz-continuous and there exist $\gamma \geq 0$ and $c > 0$ such that

$$|\lambda(q)| \leq c(1 + |q|^\gamma), \quad \lambda \in \mathbb{R}^2. \tag{2.2}$$

Moreover

$$\inf_{q \in \mathbb{R}^2} \lambda(q) =: \lambda_0 > 0.$$

Next, for every $\epsilon \geq 0$ we introduce the regularized problem

$$\begin{cases} \mu \ddot{q}_{\mu,\epsilon}(t) = b(q_{\mu,\epsilon}(t)) - \Lambda_\epsilon(q_{\mu,\epsilon}(t))\dot{q}_{\mu,\epsilon}(t) + \sigma(q_{\mu,\epsilon}(t)) \dot{w}_t, \\ q_{\mu,\epsilon}(0) = q \in \mathbb{R}^2, \quad \dot{q}_{\mu,\epsilon}(0) = p \in \mathbb{R}^2, \end{cases} \tag{2.3}$$

where

$$\Lambda_\epsilon(q) = \lambda(q)A + \epsilon I = \begin{pmatrix} \epsilon & \lambda(q) \\ -\lambda(q) & \epsilon \end{pmatrix}, \quad q \in \mathbb{R}^2.$$

Notice that for every $\epsilon > 0$ the matrix $\Lambda_\epsilon(q)$ is uniformly non-degenerate, as

$$\langle \Lambda_\epsilon(q)p, p \rangle = \epsilon |p|^2. \tag{2.4}$$

Moreover, when $\epsilon = 0$, equation (2.3) coincides with equation (2.1).

Theorem 2.1 Under Hypothesis 1, for every $\mu > 0$ and $\epsilon \geq 0$ and for every $T > 0$ and $k \geq 1$, Eq. (2.3) admits a unique adapted solution $q_{\mu,\epsilon} \in L^k(\Omega; C([0, T]; \mathbb{R}^2))$.

Proof For every $q, p \in \mathbb{R}^2$ and $n \in \mathbb{N}$, we define

$$\beta_n(p) = \begin{cases} p, & \text{if } |p| \leq n, \\ np/|p|, & \text{if } |p| \geq n, \end{cases}$$

and

$$\Lambda_{\epsilon,n}(q) = \lambda_n(q)A + \epsilon I, \quad \text{where } \lambda_n(q) = \begin{cases} \lambda(q), & \text{if } |q| \leq n, \\ \lambda((n+1)q/|q|), & \text{if } |q| \geq n+1, \end{cases}.$$

Notice that $\lambda_n : \mathbb{R}^2 \rightarrow \mathbb{R}$ is Lipschitz-continuous and

$$|\lambda_n(q)| \leq c(1 + |q|^\gamma), \quad |\beta_n(p)| \leq |p|, \tag{2.5}$$

for some constant c independent of n . Moreover, since $\langle A\beta_n(p), p \rangle = 0$, and $\langle \beta_n(p), p \rangle \leq |p|^2$, for every $p \in \mathbb{R}^2$ and $n \in \mathbb{N}$, we have

$$\langle \Lambda_{\epsilon,n}(q)\beta_n(p), p \rangle = \epsilon |p|^2, \tag{2.6}$$

for every $p, q \in \mathbb{R}^2, n \in \mathbb{N}$ and $\epsilon > 0$.

With these notations, we introduce the problem

$$\begin{cases} \mu \ddot{q}_{\mu,\epsilon}^n(t) = b(q_{\mu,\epsilon}^n(t)) - \Lambda_{\epsilon,n}(q_{\mu,\epsilon}^n(t))\beta_n(\dot{q}_{\mu,\epsilon}^n(t)) + \sigma(q_{\mu,\epsilon}^n(t)) \dot{w}_t, \\ q_{\mu,\epsilon}^n(0) = q \in \mathbb{R}^2, \quad \dot{q}_{\mu,\epsilon}^n(0) = p \in \mathbb{R}^2, \end{cases}$$

which can be rewritten as

$$\begin{cases} dq_{\mu,\epsilon}^n(t) = p_{\mu,\epsilon}^n(t) dt, & q_{\mu,\epsilon}^n(0) = q \\ \mu dp_{\mu,\epsilon}^n(t) = [b(q_{\mu,\epsilon}^n(t)) - \Lambda_{\epsilon,n}(q_{\mu,\epsilon}^n(t))\beta_n(p_{\mu,\epsilon}^n(t))] + \sigma(q_{\mu,\epsilon}^n(t)) dw(t), & p_{\mu,\epsilon}^n(t0) = p. \end{cases} \tag{2.7}$$

It is immediate to check that, for every fixed $n \in \mathbb{N}$ and $\epsilon > 0$, the mapping

$$(q, p) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \Lambda_{\epsilon,n}(q)\beta_n(p) \in \mathbb{R}^2,$$

is Lipschitz-continuous, so that equation (2.7) admits a unique adapted solution $(q_{\mu,\epsilon}^n, p_{\mu,\epsilon}^n) \in L^p(\Omega; C^1([0, T]; \mathbb{R}^2) \times C([0, T]; \mathbb{R}^2))$.

Now, applying Itô's formula to the function $\Phi(q, p) = |q|^{2k} + |p|^{2k}$, for $k \geq 2$, we obtain

$$\begin{aligned} & |q_{\mu,\epsilon}^n(t)|^{2k} + |p_{\mu,\epsilon}^n(t)|^{2k} \\ &= |q|^{2k} + |p|^{2k} + k \int_0^t |q_{\mu,\epsilon}^n(s)|^{2k-2} \langle q_{\mu,\epsilon}^n(s), p_{\mu,\epsilon}^n(s) \rangle ds \\ & \quad + \frac{k}{\mu} \int_0^t |p_{\mu,\epsilon}^n(s)|^{2k-2} \langle p_{\mu,\epsilon}^n(s), b(q_{\mu,\epsilon}^n(s) - \Lambda_{\epsilon,n}(q_{\mu,\epsilon}^n(s))\beta_n(p_{\mu,\epsilon}^n(s))) \rangle ds \\ & \quad + \frac{k}{2\mu^2} \int_0^t |p_{\mu,\epsilon}^n(s)|^{2k-2} \text{Tr} [\sigma \sigma^*(q_{\mu,\epsilon}^n(s))] ds \\ & \quad + \frac{k(k-1)}{2\mu^2} \int_0^t |p_{\mu,\epsilon}^n(s)|^{2k-4} |\sigma(q_{\mu,\epsilon}^n(s))p_{\mu,\epsilon}^n(s)|^2 ds \\ & \quad + \frac{k}{\mu} \int_0^t |p_{\mu,\epsilon}^n(s)|^{2k-2} \langle p_{\mu,\epsilon}^n(s), \sigma(q_{\mu,\epsilon}^n(s)) dw(s) \rangle. \end{aligned}$$

Therefore, thanks to (2.6) and to the Young inequality, we have for every $\epsilon > 0$

$$\begin{aligned} |q_{\mu,\epsilon}^n(t)|^{2k} + |p_{\mu,\epsilon}^n(t)|^{2k} &\leq |q|^{2k} + |p|^{2k} + c_{k,\mu} \int_0^t [|q_{\mu,\epsilon}^n(s)|^{2k} + |p_{\mu,\epsilon}^n(s)|^{2k}] ds \\ & \quad + \frac{k}{\mu} \int_0^t |p_{\mu,\epsilon}^n(s)|^{2k-2} \langle p_{\mu,\epsilon}^n(s), \sigma(q_{\mu,\epsilon}^n(s)) dw(s) \rangle. \end{aligned}$$

Taking expected in values of both sides and using the Gronwall lemma, we obtain

$$\mathbb{E}|q_{\mu,\epsilon}^n(t)|^{2k} + \mathbb{E}|p_{\mu,\epsilon}^n(t)|^{2k} \leq c_{k,\mu}(T) \left(1 + |q|^{2k} + |p|^{2k} \right), \quad t \in [0, T]. \tag{2.8}$$

Therefore, since

$$q_{\mu,\epsilon}^n(t) = q + \int_0^t p_{\mu,\epsilon}^n(s) ds,$$

and

$$p_{\mu,\epsilon}^n(t) = p + \frac{1}{\mu} \int_0^t [b(q_{\mu,\epsilon}^n(s)) - \Lambda_{\epsilon,n}(q_{\mu,\epsilon}^n(s))\beta_n(p_{\mu,\epsilon}^n(s))] ds + \frac{1}{\mu} \int_0^t \sigma(q_{\mu,\epsilon}^n(s)) dw(s),$$

using (2.5), from (2.8) we obtain

$$\sup_{n \in \mathbb{N}} \mathbb{E} \sup_{t \in [0, T]} \left(|q_{\mu, \epsilon}^n(t)|^{2k} + |p_{\mu, \epsilon}^n(t)|^{2k} \right) \leq c_{k, \mu}(T, |q|, |p|). \tag{2.9}$$

Now, for any $n \in \mathbb{N}$ we define

$$\tau_n = \inf \{ t \geq 0 : |q_{\mu, \epsilon}^n(t)| \vee |p_{\mu, \epsilon}^n(t)| \geq n \},$$

with the usual convention that $\inf \emptyset = +\infty$. Since

$$(q_{\mu, \epsilon}^n(t), p_{\mu, \epsilon}^n(t)) = (q_{\mu, \epsilon}^m(t), p_{\mu, \epsilon}^m(t)), \quad n < m, \quad t \leq \tau_n, \tag{2.10}$$

it follows that the sequence $\{\tau_n\}_{n \in \mathbb{N}}$ is non-decreasing, \mathbb{P} -a.s., so that we can define

$$\tau = \lim_{n \rightarrow \infty} \tau_n.$$

Due to (2.9), for every fixed $T > 0$ we have

$$\begin{aligned} & \mathbb{P} \left(\sup_{t \in [0, T]} |q_{\mu, \epsilon}^n(t)| \leq n, \quad \sup_{t \in [0, T]} |p_{\mu, \epsilon}^n(t)| \leq n \right) \\ & \geq 1 - \mathbb{P} \left(\sup_{t \in [0, T]} |q_{\mu, \epsilon}^n(t)| > n \right) - \mathbb{P} \left(\sup_{t \in [0, T]} |p_{\mu, \epsilon}^n(t)| > n \right) \\ & \geq 1 - \frac{2c_{1, \mu}(T, |q|, |p|)}{n}. \end{aligned}$$

This implies

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tau_n > T) = 1,$$

and then, since T is arbitrary, we conclude

$$\mathbb{P}(\tau = +\infty) = 1.$$

In particular, if we set

$$(q_{\mu, \epsilon}(t), p_{\mu, \epsilon}(t)) = (q_{\mu, \epsilon}^n(t \wedge \tau_n), p_{\mu, \epsilon}^n(t \wedge \tau_n)), \quad t \leq \tau,$$

due to (2.10) we can conclude that there exists a unique solution $(q_{\mu, \epsilon}, p_{\mu, \epsilon})$ to problem (2.3), belonging to $L^k(\Omega; C^1([0, T]; \mathbb{R}^2) \times C([0, T]; \mathbb{R}^2))$, for every $k \geq 1$ and $T > 0$. \square

3 The Smoluchowski–Kramers Approximation for the Regularized Problem

For every $\epsilon > 0$ and $q \in \mathbb{R}^2$, the matrix $\Lambda_\epsilon(q)$ is invertible and

$$\Lambda_\epsilon^{-1}(q) = \frac{1}{\lambda^2(q) + \epsilon^2} \begin{pmatrix} \epsilon & -\lambda(q) \\ \lambda(q) & \epsilon \end{pmatrix}. \tag{3.1}$$

We introduce the vector field $S^\epsilon(q)$, whose j -th component is defined by

$$S_j^\epsilon(q) = \sum_{i, l=1}^2 \partial_i (\Lambda_\epsilon^{-1})_{jl}(q) J_{li}^\epsilon(q), \quad j = 1, 2, \tag{3.2}$$

where $\partial_i = \partial/\partial q_i$ and J^ϵ is the matrix-valued function solving the Lyapunov equation

$$J^\epsilon(q)\Lambda_\epsilon^*(q) + \Lambda_\epsilon(q)J^\epsilon(q) = \sigma(q)\sigma^*(q), \quad q \in \mathbb{R}^2.$$

Thanks to (2.4), the equation above has a unique solution J^ϵ which can be explicitly written as

$$\begin{aligned} J^\epsilon(q) &= \int_0^\infty e^{-\Lambda_\epsilon(q)r} \sigma \sigma^*(q) e^{-\Lambda_\epsilon^*(q)r} dr \\ &= \int_0^\infty e^{-\lambda(q)Ar} \sigma \sigma^*(q) e^{\lambda(q)Ar} e^{-2\epsilon r} dr, \quad q \in \mathbb{R}^2. \end{aligned} \tag{3.3}$$

It is immediate to check that

$$e^{-\lambda(q)Ar} = \begin{pmatrix} \cos(\lambda(q)r) & -\sin(\lambda(q)r) \\ \sin(\lambda(q)r) & \cos(\lambda(q)r) \end{pmatrix}, \quad r \geq 0.$$

In what follows, for every $q \in \mathbb{R}^2$ we denote

$$\begin{pmatrix} a_1(q) & a_0(q) \\ a_0(q) & a_2(q) \end{pmatrix} =: \sigma \sigma^*(q),$$

and

$$\beta_0(q) := \frac{a_1(q) + a_2(q)}{4}, \quad \beta_1(q) := \frac{a_1(q) - a_2(q)}{4} \quad \beta_2(q) := \frac{a_0(q)}{2}. \tag{3.4}$$

Lemma 3.1 Assume that $\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ is differentiable. Then, there exist $M : \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$ and $R^\epsilon : \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$ such that for every $\epsilon > 0$

$$S^\epsilon(q) = \frac{1}{\epsilon} \frac{\beta_0(q)}{\lambda^2(q)} \nabla^\perp \lambda(q) - M(q) \nabla \lambda(q) + R^\epsilon(q) \nabla \lambda(q), \quad q \in \mathbb{R}^2. \tag{3.5}$$

Proof Thanks to (3.3), we have

$$\begin{aligned} J_{11}^\epsilon(q) &= \frac{\beta_0(q)}{\epsilon} + \beta_1(q) \int_0^\infty \cos(\lambda(q)r) e^{-\epsilon r} dr - \beta_2(q) \int_0^\infty \sin(\lambda(q)r) e^{-\epsilon r} dr \\ J_{22}^\epsilon(q) &= \frac{\beta_0(q)}{\epsilon} - \beta_1(q) \int_0^\infty \cos(\lambda(q)r) e^{-\epsilon r} dr + \beta_2(q) \int_0^\infty \sin(\lambda(q)r) e^{-\epsilon r} dr \\ J_{12}^\epsilon(q) &= J_{21}^\epsilon(q) = \beta_1(q) \int_0^\infty \sin(\lambda(q)r) e^{-\epsilon r} dr + \beta_2(q) \int_0^\infty \cos(\lambda(q)r) e^{-\epsilon r} dr. \end{aligned}$$

Integrating by parts, we obtain

$$\int_0^\infty \cos(\lambda(q)r) e^{-\epsilon r} dr = \frac{\epsilon}{\lambda^2(q) + \epsilon^2},$$

and

$$\int_0^\infty \sin(\lambda(q)r) e^{-\epsilon r} dr = \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2}.$$

This implies

$$\begin{aligned}
 J_{11}^\epsilon(q) &= \frac{\beta_0(q)}{\epsilon} + \beta_1(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} - \beta_2(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} \\
 J_{22}^\epsilon(q) &= \frac{\beta_0(q)}{\epsilon} - \beta_1(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} \\
 J_{12}^\epsilon(q) &= J_{21}^\epsilon(q) = \beta_1(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2}.
 \end{aligned}
 \tag{3.6}$$

Now, due to (3.1), for every $\epsilon > 0$ and $q \in \mathbb{R}^2$ we have

$$\begin{aligned}
 \partial_i (\Lambda_\epsilon^{-1})_{11}(q) &= \partial_i (\Lambda_\epsilon^{-1})_{22}(q) = -\frac{2\epsilon\lambda(q)}{(\lambda^2(q) + \epsilon^2)^2} \partial_i \lambda(q), \quad i = 1, 2, \\
 \partial_i (\Lambda_\epsilon^{-1})_{12}(q) &= -\partial_i (\Lambda_\epsilon^{-1})_{21}(q) = \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^2} \partial_i \lambda(q), \quad i = 1, 2.
 \end{aligned}
 \tag{3.7}$$

Substituting (3.6) and (3.7) in (3.2), we obtain

$$\begin{aligned}
 S_1^\epsilon(q) &= -\frac{2\epsilon\lambda(q)}{(\lambda^2(q) + \epsilon^2)^2} \left[\left(\frac{\beta_0(q)}{\epsilon} + \beta_1(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} - \beta_2(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} \right) \partial_1 \lambda(q) \right. \\
 &\quad \left. + \left(\beta_1(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} \right) \partial_2 \lambda(q) \right] \\
 &\quad + \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^2} \left[\left(\beta_1(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} \right) \partial_1 \lambda(q) \right. \\
 &\quad \left. + \left(\frac{\beta_0(q)}{\epsilon} - \beta_1(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} \right) \partial_2 \lambda(q) \right],
 \end{aligned}$$

and

$$\begin{aligned}
 S_2^\epsilon(q) &= -\frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^2} \left[\left(\frac{\beta_0(q)}{\epsilon} + \beta_1(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} - \beta_2(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} \right) \partial_1 \lambda(q) \right. \\
 &\quad \left. + \left(\beta_1(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} \right) \partial_2 \lambda(q) \right] \\
 &\quad - \frac{2\epsilon\lambda(q)}{(\lambda^2(q) + \epsilon^2)^2} \left[\left(\beta_1(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} \right) \partial_1 \lambda(q) \right. \\
 &\quad \left. + \left(\frac{\beta_0(q)}{\epsilon} - \beta_1(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} \right) \partial_2 \lambda(q) \right],
 \end{aligned}$$

Now, we define

$$\Gamma_1^\epsilon(q) := \beta_1(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2} + \beta_2(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2}, \quad \Gamma_1(q) := \frac{\beta_1(q)}{\lambda(q)},$$

and

$$\Gamma_2^\epsilon(q) := \beta_1(q) \frac{\epsilon}{\lambda^2(q) + \epsilon^2} - \beta_2(q) \frac{\lambda(q)}{\lambda^2(q) + \epsilon^2}, \quad \Gamma_2(q) := -\frac{\beta_2(q)}{\lambda(q)}.$$

With these notations

$$S_1^\epsilon(q) = \frac{1}{\epsilon} \frac{\beta_0(q)}{\lambda^2(q)} \partial_2 \lambda(q) + \left[-\frac{2\beta_0(q)}{\lambda^3(q)} + \frac{\Gamma_1(q)}{\lambda^2(q)} \right] \partial_1 \lambda(q) - \frac{\Gamma_2(q)}{\lambda^2(q)} \partial_2 \lambda(q) + R_{11}^\epsilon(q) \partial_1 \lambda(q) + R_{12}^\epsilon(q) \partial_2 \lambda(q),$$

where

$$R_{11}^\epsilon(q) := -\frac{2\epsilon\lambda(q)}{(\lambda^2(q) + \epsilon^2)^2} \Gamma_2^\epsilon(q) + \epsilon\beta_2(q) \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^3} + 2\beta_0(q) \left[\frac{1}{\lambda^3(q)} - \frac{\lambda(q)}{(\lambda^2(q) + \epsilon^2)^2} \right] - \beta_1(q) \left[\frac{1}{\lambda^3(q)} - \frac{\lambda(q)(\lambda^2(q) - \epsilon^2)}{(\lambda^2(q) + \epsilon^2)^3} \right], \tag{3.8}$$

and

$$R_{12}^\epsilon(q) := -\frac{2\lambda(q)\epsilon}{(\lambda^2(q) + \epsilon^2)^2} \Gamma_1^\epsilon(q) - \epsilon\beta_2(q) \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^3} - \frac{\beta_0(q)}{\epsilon} \left[\frac{1}{\lambda^2(q)} - \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^2} \right] + \beta_1(q) \left[\frac{1}{\lambda^3(q)} - \frac{\lambda(q)(\lambda^2(q) - \epsilon^2)}{(\lambda^2(q) + \epsilon^2)^3} \right]. \tag{3.9}$$

Similarly, we have

$$S_2^\epsilon(q) = -\frac{1}{\epsilon} \frac{\beta_0(q)}{\lambda^2(q)} \partial_1 \lambda(q) - \frac{\Gamma_2(q)}{\lambda^2(q)} \partial_1 \lambda(q) - \left[\frac{2\beta_0(q)}{\lambda^3(q)} + \frac{\Gamma_1(q)}{\lambda^2(q)} \right] \partial_2 \lambda(q) + R_{21}^\epsilon(q) \partial_1 \lambda(q) + R_{22}^\epsilon(q) \partial_2 \lambda(q),$$

where

$$R_{21}^\epsilon(q) := -\frac{2\lambda(q)\epsilon}{(\lambda^2(q) + \epsilon^2)^2} \Gamma_1^\epsilon(q) - \epsilon\beta_1(q) \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^3} + \frac{\beta_0(q)}{\epsilon} \left[\frac{1}{\lambda^2(q)} - \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^2} \right] - \beta_2(q) \left[\frac{1}{\lambda^3(q)} - \frac{\lambda(q)(\lambda^2(q) - \epsilon^2)}{(\lambda^2(q) + \epsilon^2)^3} \right], \tag{3.10}$$

and

$$R_{22}^\epsilon(q) := \frac{2\epsilon\lambda(q)}{(\lambda^2(q) + \epsilon^2)^2} \Gamma_2^\epsilon(q) - \epsilon\beta_2(q) \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^3} + 2\beta_0(q) \left[\frac{1}{\lambda^3(q)} - \frac{\lambda(q)}{(\lambda^2(q) + \epsilon^2)^2} \right] + \beta_1(q) \left[\frac{1}{\lambda^3(q)} - \frac{\lambda(q)(\lambda^2(q) - \epsilon^2)}{(\lambda^2(q) + \epsilon^2)^3} \right]. \tag{3.11}$$

Therefore, recalling that $\Gamma_1(q) = \beta_1(q)/\lambda(q)$ and $\Gamma_2(q) = -\beta_2(q)/\lambda(q)$, if we define

$$M(q) = \frac{1}{\lambda^3(q)} \begin{pmatrix} 2\beta_0(q) - \beta_1(q) & -\beta_2(q) \\ -\beta_2(q) & 2\beta_0(q) + \beta_1(q) \end{pmatrix}, \tag{3.12}$$

and $R^\epsilon(q) = (R_{ij}^\epsilon(q))_{i,j=1,2}$, where the components $R_{ij}^\epsilon(q)$ are defined in (3.8), (3.9), (3.10) and (3.11), we obtain (3.5). \square

In what follows we shall assume that the following condition is satisfied.

Hypothesis 2 1. The mapping $\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuously differentiable.

2. For every $\epsilon > 0$, the mapping $S_\epsilon : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ introduced in (3.2) is locally Lipschitz-continuous and has linear growth.

3. For every $\epsilon > 0$ the mappings $\Lambda_\epsilon^{-1}b : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ and $\Lambda_\epsilon^{-1}\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$ are locally Lipschitz-continuous and have linear growth.

Remark 3.2 1. AUsing the explicit expression of $M(q)$ given in (3.12) and the expressions for the coefficients of $R_\epsilon(q)$ given in (3.8), (3.9), (3.10) and (3.11), thanks to what we have already assumed in Hypothesis 1, we can check easily that Hypothesis 2 is satisfied if we assume σ to be bounded and λ to be bounded and differentiable, with $\nabla\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ Lipschitz-continuous.

2. On the other hand, if we assume that $\nabla\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is locally Lipschitz-continuous and has linear growth and there exists $c > 0$ such that for $|q|$ large enough

$$|\lambda(q)| \geq c |q|^2,$$

then Hypothesis 2 is satisfied, without assuming σ to be bounded.

Theorem 3.3 For every $\mu, \epsilon > 0$, let $q_{\mu,\epsilon}$ be the solution of problem (2.7). Then, under Hypotheses 1 and 2, for every $\epsilon > 0$ we have

$$\lim_{\mu \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} |q_{\mu,\epsilon}(t) - q_\epsilon(t)| = 0, \tag{3.13}$$

where q_ϵ is the solution of the problem

$$dq_\epsilon(t) = [\Lambda_\epsilon^{-1}b(q_\epsilon(t)) + S_\epsilon(q_\epsilon(t))] dt + \Lambda_\epsilon^{-1}\sigma(q_\epsilon(t)) dw(t), \quad q_\epsilon(0) = q. \tag{3.14}$$

Proof Assuming Hypotheses 1 and 2, we have that for every $\epsilon > 0$ and for every $k \geq 1$ and $T > 0$ problem (3.14) admits a unique solution $q_\epsilon \in L^k(\Omega; C([0, T]; \mathbb{R}^2))$. As $\langle \Lambda_\epsilon(q)p, p \rangle = \epsilon |p|^2$, this allows to conclude the proof thanks to [13, Theorem 2.4]. \square

4 The Averaging Limit

In this section we want to investigate the limiting behavior of the slow component of q_ϵ , as ϵ goes to zero. To this purpose, we need to introduce some preliminary material.

4.1 Some Notations and Further Assumptions

We consider here the system

$$\dot{X}(t) = \frac{\beta_0(X(t))}{\lambda^2(X(t))} \nabla^\perp \lambda(X(t)). \tag{4.1}$$

Clearly, for every $t \geq 0$, we have $\lambda(X(t)) = \lambda(X(0))$. Now, if we consider the perturbed system

$$\begin{aligned} dX_\epsilon(t) &= \frac{\beta_0(X_\epsilon(t))}{\lambda^2(X_\epsilon(t))} \nabla^\perp \lambda(X_\epsilon(t)) dt \\ &+ \epsilon \left[\frac{1}{\lambda(X_\epsilon(t))} Ab(X_\epsilon(t)) - M(X_\epsilon(t)) \nabla \lambda(X_\epsilon(t)) \right] dt \\ &+ \frac{\sqrt{\epsilon}}{\lambda(X_\epsilon(t))} A\sigma(X_\epsilon(t)) dw(t) \\ &+ \epsilon^2 \left[H^\epsilon(X_\epsilon(t))b(X_\epsilon(t)) + \hat{R}^\epsilon(X_\epsilon(t)) \nabla \lambda(X_\epsilon(t)) \right] dt \\ &+ \epsilon H^\epsilon(X_\epsilon(t))\sigma(X_\epsilon(t)) dw(t), \end{aligned}$$

the quantity $\lambda(X_\epsilon(t))$ is no longer conserved. However, for any fixed time interval $[0, T]$ and for every $k \geq 1$, we have

$$\lim_{\epsilon \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} |X_\epsilon(t) - X(t)|^k = 0,$$

and, as an immediate consequence,

$$\lim_{\epsilon \rightarrow 0} \mathbb{E} \sup_{t \in [0, T]} |\lambda(X_\epsilon(t)) - \lambda(X(0))|^k = 0.$$

Now, with the change of time $t \mapsto t/\epsilon$, we can check that

$$\mathcal{L}(X_\epsilon(\cdot/\epsilon)) = \mathcal{L}(q_\epsilon(\cdot)),$$

where q_ϵ is the solution of equation (3.14). As mentioned above, our aim is to identify the non trivial limit for the distribution of the process $\lambda(q_\epsilon(\cdot))$, as $\epsilon \downarrow 0$. To this purpose, in addition to Hypotheses 1 and 2, we assume that λ satisfies the following conditions.

Hypothesis 3 1. If β_0 is the function defined in (3.4), we have

$$\inf_{x \in \mathbb{R}^2} \beta_0(x) > 0. \tag{4.2}$$

2. The mapping $\lambda : \mathbb{R}^2 \rightarrow \mathbb{R}$ is four times continuously differentiable, with bounded second derivative.
3. The mapping λ has only a finite number of critical points x_1, \dots, x_n . The matrix of second derivative $D^2\lambda(x_i)$ is non degenerate, for every $i = 1, \dots, n$ and $\lambda(x_i) \neq \lambda(x_j)$, if $i \neq j$.
4. There exists a positive constant a such that $\lambda(x) \geq a|x|^2$, $|\nabla \lambda(x)| \geq a|x|$ and $\Delta \lambda(x) \geq a$, for all $x \in \mathbb{R}^2$, with $|x|$ large enough.

Remark 4.1 Remember that the function β_0 was defined as $[(\sigma\sigma^*)_{11}^2 + (\sigma\sigma^*)_{22}^2]/4$. Therefore, condition (4.2) is a non-degeneracy condition on the noisy perturbation.

Next, for every $z \geq \lambda_0$, we denote by $C(z)$ the z -level set

$$C(z) = \{x \in \mathbb{R}^2 : \lambda(x) = z\}.$$

The set $C(z)$ may consist of several connected components

$$C(z) = \bigcup_{k=1}^{N(z)} C_k(z),$$

and for every $x \in \mathbb{R}^2$ we have

$$X(0) = x \implies X(t) \in C_{k(x)}(\lambda(x)), \quad t \geq 0,$$

where $C_{k(x)}(x)$ is the connected component of the level set $C(\lambda(x))$, to which the point x belongs. For every $z \geq 0$ and $k = 1, \dots, N(z)$, we shall denote by $G_k(z)$ the domain of \mathbb{R}^2 bounded by the level set component $C_k(z)$.

If we identify all points in \mathbb{R}^2 belonging to the same connected component of a given level set $C(z)$ of the Hamiltonian λ , we obtain a graph Γ , given by several edges I_1, \dots, I_n and vertices O_1, \dots, O_m . The vertices will be of two different types, external and internal vertices. External vertices correspond to local extrema of λ , while internal vertices correspond to saddle points of λ . Among external vertices, we will also include O_∞ , the vertex of the graph corresponding to the point at infinity.

In what follows, we shall denote by $\Pi : \mathbb{R}^2 \rightarrow \Gamma$ the *identification map*, that associates to every point $x \in \mathbb{R}^2$ the corresponding point $\Pi(x)$ of the graph Γ . We have $\Pi(x) = (\lambda(x), k(x))$, where $k(x)$ denotes the number of the edge on the graph Γ , containing the point $\Pi(x)$. If O_i is one of the interior vertices, the second coordinate cannot be chosen in a unique way, as there are three edges having O_i as their endpoint. Notice that both $k(x)$ and $H(x)$ are first integrals (a discrete and a continuous one, respectively) of system (4.1).

On the graph Γ , a distance can be introduced in the following way. If $y_1 = (z_1, k)$ and $y_2 = (z_2, k)$ belong to the same edge I_k , then $d(y_1, y_2) = |z_1 - z_2|$. In the case y_1 and y_2 belong to different edges, then

$$d(y_1, y_2) = \min \{d(y_1, O_{i_1}) + d(O_{i_1}, O_{i_2}) + \dots + d(O_{i_j}, y_2)\},$$

where the minimum is taken over all possible paths from y_1 to y_2 , through every possible sequence of vertices O_{i_1}, \dots, O_{i_j} , connecting y_1 to y_2 .

If z is not a critical value, then each $C_k(z)$ consists of one periodic trajectory of the vector field $\nabla^\perp \lambda(x)$. If z is a local extremum of $\lambda(x)$, then, among the components of $C(z)$ there is a set consisting of one point, the equilibrium point of the flow. If $\lambda(x)$ has a saddle point at some point x_0 and $\lambda(x_0) = z$, then $C(z)$ consists of three trajectories: the equilibrium point x_0 and the two trajectories that have x_0 as their limiting point, as $t \rightarrow \pm\infty$.

Now, for every $(z, k) \in \Gamma$, we define

$$T_k(z) = \oint_{C_k(z)} \frac{\lambda^2(x)}{\beta_0(x)|\nabla\lambda(x)|} dl_{z,k}, \tag{4.3}$$

where $dl_{z,k}$ is the length element on $C_k(z)$. Notice that $T_k(z)$ is the period of the motion along the level set $C_k(z)$.

As is seen above, if $X(0) = x \in C_k(z)$, then $X(t) \in C_k(z)$, for every $t \geq 0$. As known, for every $(z, k) \in \Gamma$ the probability measure

$$d\mu_{z,k} := \frac{1}{T_k(z)} \frac{\lambda^2(x)}{\beta_0(x)|\nabla\lambda(x)|} dl_{z,k} \tag{4.4}$$

is invariant for system (4.1) on the level set $C_k(z)$.

4.2 The Limit of $\Pi(q_\epsilon)$

Due to (3.1), for every $\epsilon > 0$ we have

$$\Lambda_\epsilon^{-1}(q) = \frac{1}{\lambda(q)} A + \epsilon H^\epsilon(q), \tag{4.5}$$

where

$$H^\epsilon(q) := \frac{1}{\lambda^2(q) + \epsilon^2} \left(I - \frac{\epsilon}{\lambda(q)} A \right).$$

Notice that

$$\sup_{\epsilon > 0} |H^\epsilon(q)| < \infty, \quad q \in \mathbb{R}^2. \tag{4.6}$$

Lemma 4.2 *Let $R^\epsilon : \mathbb{R}^2 \rightarrow \mathbb{R}^{2 \times 2}$ be the mapping introduced in Lemma 3.1. Then*

$$\sup_{\epsilon > 0} \frac{1}{\epsilon} |\hat{R}^\epsilon(q)| < \infty, \quad q \in \mathbb{R}^2. \tag{4.7}$$

Proof We have

$$\frac{1}{\lambda^3(q)} - \frac{\lambda(q)}{(\lambda^2(q) + \epsilon^2)^2} = \epsilon^2 \left[\frac{2\lambda^2(q) + \epsilon^2}{\lambda^3(q)(\lambda^2(q) + \epsilon^2)^2} \right],$$

and

$$\frac{1}{\lambda^3(q)} - \frac{\lambda(q)(\lambda^2(q) - \epsilon^2)}{(\lambda^2(q) + \epsilon^2)^3} = \epsilon^2 \left[\frac{4\lambda^4(q) + 3\epsilon^2\lambda^2(q) + \epsilon^4}{\lambda^3(q)(\lambda^2(q) + \epsilon^2)^3} \right]$$

and

$$\frac{1}{\lambda^2(q)} - \frac{\lambda^2(q) - \epsilon^2}{(\lambda^2(q) + \epsilon^2)^2} = \epsilon^2 \left[\frac{3\lambda^2(q) + \epsilon^2}{\lambda^2(q)(\lambda^2(q) + \epsilon^2)^2} \right].$$

Therefore, in view of the definition of $R^\epsilon(q)$ in (3.8), (3.9), (3.10) and (3.11), this concludes the proof. □

According to (3.5), (4.5), (4.6) and (4.7), equation (3.14) can be rewritten as

$$dq_\epsilon(q) = \frac{1}{\epsilon} \frac{\beta_0(q_\epsilon(t))}{\lambda^2(q_\epsilon(t))} \nabla^\perp \lambda(q_\epsilon(t)) dt + B(q_\epsilon(t)) dt + \Sigma(q_\epsilon(t)) dw(t), \\ + \epsilon [B_\epsilon(q_\epsilon(t)) dt + \Sigma_\epsilon(q_\epsilon(t)) dw(t)], \quad q_\epsilon(0) = q, \tag{4.8}$$

where

$$B(q) = \frac{1}{\lambda(q)} Ab(q) - M(q)\nabla\lambda(q), \quad \Sigma(q) = \frac{1}{\lambda(q)} A\sigma(q),$$

and

$$B_\epsilon(q) = H^\epsilon(q)b(q) + \frac{1}{\epsilon} R^\epsilon(q)\nabla\lambda(q), \quad \Sigma_\epsilon(q) = H^\epsilon(q)\sigma(q).$$

This means that, as $\epsilon \downarrow 0$, some of the coefficients are of order $O(\epsilon^{-1})$, part of order $O(1)$ and part of order $O(\epsilon)$.

With the notations introduced in the previous section, in what follows, we want to investigate the limiting behavior of the Γ -valued process $\Pi(q_\epsilon(\cdot)) = (\lambda(q_\epsilon(\cdot)), k(q_\epsilon(\cdot)))$, as $\epsilon \downarrow 0$.

Applying Itô's formula to $\lambda(q_\epsilon(t))$, we get

$$d\lambda(q_\epsilon(t)) = \mathcal{G}\lambda(q_\epsilon(t)) dt + \mathcal{A}\lambda(q_\epsilon(t)) dw(t) + \epsilon \mathcal{G}_\epsilon\lambda(q_\epsilon(t)) dt + \epsilon \mathcal{A}_\epsilon\lambda(q_\epsilon(t)) dw(t),$$

where for every $f \in C^2(\mathbb{R}^2)$ and $q \in \mathbb{R}^2$

$$\begin{aligned} \mathcal{G}f(q) &= \frac{1}{2} \text{Tr} [\Sigma \Sigma^*(q) D^2 f(q)] + \langle Df(q), B(q) \rangle, \\ \mathcal{A}f(q) &= \Sigma(q)^* Df(q), \\ \mathcal{G}_\epsilon f(q) &= \frac{1}{2} \text{Tr} [(\epsilon \Sigma_\epsilon \Sigma_\epsilon^*(q) + \Sigma \Sigma_\epsilon^*(q) + \Sigma_\epsilon \Sigma^*(q)) D^2 f(q)] + \langle Df(q), B_\epsilon(q) \rangle, \end{aligned}$$

and

$$\mathcal{A}_\epsilon f(q) = \Sigma_\epsilon^*(q) Df(q).$$

We recall that the graph Γ is made of n edges I_1, \dots, I_n and m vertices O_1, \dots, O_m . For every $j = 1, \dots, n$ and for every f that is twice differentiable in the interior of the edge I_j , we denote

$$\mathcal{L}_j f(z) = \frac{1}{2} \alpha_j(z) f''(x) + \gamma_j(z) f'(z), \tag{4.9}$$

where

$$\begin{aligned} \alpha_j(z) &= \oint_{C_j(z)} |\mathcal{A}\lambda(x)|^2 d\mu_{z,j}(x) = \oint_{C_j(z)} |\Sigma^*(x) \nabla \lambda(x)|^2 d\mu_{z,j}(x), \\ \gamma_j(z) &= \oint_{C_j(z)} \mathcal{G}\lambda(x) d\mu_{z,j}(x), \end{aligned}$$

and $d\mu_{z,j}$ is the probability measure introduced in (4.4).

Definition 4.3 For each interior vertex O_k and any edge I_j adjacent to O_k (notation $I_j \sim O_k$), let ρ_{kj} be the positive constant defined by

$$\rho_{kj} = \oint_{C_{kj}} \frac{\lambda^2(x)}{\beta_0(x) |\nabla \lambda(x)|} |\Sigma^*(x) \nabla \lambda(x)|^2 dl(x).$$

We denote by $D(L) \subset C(\Gamma)$ the set consisting all continuously differentiable functions f defined on the graph Γ such that $\mathcal{L}_j f$ is well defined in the interior of the edge I_j and for every $I_j \sim O_k$ there exists finite

$$\lim_{x \rightarrow O_k} \mathcal{L}_j f(x)$$

and the limit is independent of the edge I_j . Moreover, for each interior vertex O_k

$$\sum_{j: I_j \sim O_k} \pm \rho_{kj} f'_j(\lambda(O_k)) = 0,$$

where f'_j denotes the derivative of f with respect to the local coordinate z , along the edge I_j and the sign \pm are taken if on the edge I_j it holds $z > \lambda(O_k)$ or $z < \lambda(O_k)$.

Next, for every $f \in D(L)$, we define

$$Lf(x) = \begin{cases} \mathcal{L}_j f(x), & \text{if } x \text{ is an interior point of } I_j, \\ \lim_{x \rightarrow O_k} \mathcal{L}_j f(x), & \text{if } x \text{ is the vertex } O_k \text{ and } I_j \sim O_k. \end{cases}$$

As proven in [12, Theorem 8.2.1], in case $\Sigma(q) = I$ the operator L defined on the domain $D(L)$, as described in Definition 4.3, is the generator of a strong Markov process Y_t on Γ with continuous trajectories. Here the same result holds, because of the non-degeneracy condition (4.2) satisfied by the diffusion coefficient $\Sigma(q)$.

In fact, as shown in the next theorem, the Markov process Y is the weak limit in $C([0, T]; \Gamma)$ of the slow motion $\Pi(q_\epsilon(\cdot))$ on Γ .

Theorem 4.4 *Under Hypotheses 1, 2 and 3, for every fixed $T > 0$ the Γ -valued process $\Pi(q_\epsilon(\cdot))$ converges weakly in $C([0, T]; \Gamma)$ to the Markov process Y generated by the operator $(L, D(L))$, introduced in Definition 4.3.*

Proof If in equation (4.8) we have $B(q) = B_\epsilon(q) = \Sigma_\epsilon(q) = 0$ and $\Sigma(q) = I$, the result above is what is proven in [12, Theorem 8.2.2]. In the present situation we are dealing with the more general situation in which we have a coefficient $B(q)$ of order $O(1)$ and coefficients $B_\epsilon(q)$ of order $O(\epsilon)$. Moreover we allow a non-constant diffusion coefficient $\Sigma(q) + \Sigma_\epsilon(q)$, where $\Sigma(q)$ is of order $O(1)$ and $\Sigma_\epsilon(q)$ is of order $O(\epsilon)$. As shown in [26], under these more general assumptions, an averaging principle of the same type of the one described in [12, Theorem 8.2.2] is still valid. This of course has required to introduce a suitable generalization of the operator $(L, D(L))$, that takes into account the coefficients B and Σ , and to extend the limiting result in presence of the vanishing terms B_ϵ and Σ_ϵ . \square

References

- Birrell, J., Hottovy, S., Volpe, G., Wehr, J.: Small mass limit of a Langevin equation on a manifold. *Ann. Henri Poincaré* **18**, 707–755 (2017)
- Birrell, J., Wehr, J.: Homogenization of dissipative, noisy Hamiltonian dynamics. *Stoch. Process. Appl.* **128**, 2367–2403 (2018)
- Cerrai, S., Freidlin, M.: On the Smoluchowski–Kramers approximation for a system with an infinite number of degrees of freedom. *Probab. Theory Relat. Fields* **135**, 363–394 (2006)
- Cerrai, S., Freidlin, M.: Smoluchowski–Kramers approximation for a general class of SPDE's. *J. Evol. Equ.* **6**, 657–689 (2006)
- Cerrai, S., Freidlin, M.: Small mass asymptotics for a charged particle in a magnetic field and longtime influence of small perturbations. *J. Stat. Phys.* **144**, 101–123 (2011)
- Cerrai, S., Freidlin, M., Salins, M.: On the Smoluchowski-Kramers approximation for SPDEs and its interplay with large deviations and long time behavior. *Discret. Contin. Dyn. Syst. A* **37**, 33–76 (2017)
- Cerrai, S., Salins, M.: Smoluchowski–Kramers approximation and large deviations for infinite dimensional gradient systems. *Asymptot. Anal.* **88**, 201–215 (2013)
- Cerrai, S., Salins, M.: Smoluchowski–Kramers approximation and large deviations for infinite dimensional non-gradient systems with applications to the exit problem. *Ann. Probab.* **44**, 2591–2642 (2016)
- Cerrai, S., Salins, M.: On the Smoluchowski–Kramers approximation for a system with an infinite number of degrees of freedom subject to a magnetic field. *Stoch. Process. Appl.* **127**, 273–303 (2017)
- Freidlin, M.: Some remarks on the Smoluchowski–Kramers approximation. *J. Stat. Phys.* **117**, 617–634 (2004)
- Freidlin, M., Hu, W.: Smoluchowski–Kramers approximation in the case of variable friction. *J. Math. Sci.* **179**, 184–207 (2011)
- Freidlin, M., Wentzell, A.: *Random Perturbations of Dynamical Systems*, 3rd edn. Springer Verlag, Berlin (2012)
- Herzog, D., Hottovy, S., Volpe, G.: The small-mass limit for Langevin dynamics with unbounded coefficients and positive friction. *J. Stat. Phys.* **163**, 659–673 (2016)
- Hottovy, S., McDaniel, A., Volpe, G., Wehr, J.: The Smoluchowski–Kramers limit of stochastic differential equations with arbitrary state-dependent friction. *Commun. Math. Phys.* **336**, 1259–1283 (2015)
- Kramers, H.: Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–304 (1940)
- Lee, J.J.: Small mass asymptotics of a charged particle in a variable magnetic field. *Asymptot. Anal.* **86**, 99–121 (2014)
- Lv, Y., Roberts, A.: Averaging approximation to singularly perturbed nonlinear stochastic wave equations. *J. Math. Phys.* **53**, 1–11 (2012)
- Lv, Y., Roberts, A.: Large deviation principle for singularly perturbed stochastic damped wave equations. *Stoch. Anal. Appl.* **32**, 50–60 (2014)

19. Lv, Y., Roberts, A., Wang, W.: Approximation of the random inertial manifold of singularly perturbed stochastic wave equations. *Stoch. Dyn.* **14**, 1350018 (2014)
20. Lv, Y., Wang, W.: Limiting dynamics for stochastic wave equations. *J. Differ. Equ.* **244**, 1–23 (2008)
21. Nguyen, H.: The small-mass limit and white-noise limit of an infinite dimensional Generalized Langevin Equation. [arXiv:1804.09682](https://arxiv.org/abs/1804.09682) (2018)
22. Roberts, A.J.: Normal form transforms separate slow and fast modes in stochastic dynamical systems. *Physica A* **387**, 12–38 (2008)
23. Salins, M.: Smoluchowski-Kramers approximation for the damped stochastic wave equation with multiplicative noise in any spatial dimension. [arXiv:1801.10538](https://arxiv.org/abs/1801.10538) (2018)
24. Smoluchowski, M.: Drei Vortage über Diffusion Brownsche Bewegung und Koagulation von Kolloidteilchen. *Physik Zeit.* **17**, 557–585 (1916)
25. Spiliopoulos, K.: A note on the Smoluchowski–Kramers approximation for the Langevin equation with reflection. *Stoch. Dyn.* **7**, 141–152 (2007)
26. Zhu, Y.: A generalization of the Freidlin–Wentzell theorem on averaging for Hamiltonian systems. [arXiv:2002.01851](https://arxiv.org/abs/2002.01851) (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.