On the Smoluchowski-Kramers approximation in the presence of a varying magnetic field

Sandra Cerrai*
University of Maryland, College Park, USA

Jan Wehr†
University of Arizona, Tucson, USA

Yichun Zhu
University of Maryland, College Park, USA

Abstract

We study the small mass limit for the equation that describes the planar motion of a charged particle of a small mass μ in a force field that has a deterministic as well as a stochastic component, combined with a magnetic field. 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 for $q_{\mu,\epsilon}$ gives the solution q_ϵ to a stochastic first order equation, where a noise-induced drift term is created. Then, by using a generalization of the classical averaging theorem for Hamiltonian systems by Freidlin and Wentcell, 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 intensity function of the magnetic field.

1 Introduction

We are dealing with the planar motion of a charged particle of a small mass μ in a force field that has a deterministic as well as a stochastic component combined with a magnetic field

$$\left\{\begin{array}{l}\mu\,\ddot{q}_\mu(t)=b(q_\mu(t))-\lambda(q_\mu(t))A\dot{q}_\mu(t)+\sigma(q_\mu(t))\,\dot{w}_t,\\[10pt] q_\mu(0)=q\in\mathbb{R}^2,\qquad \dot{q}_\mu(0)=p\in\mathbb{R}^2.\end{array}\right. \tag{1.1}$$

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, $\lambda:\mathbb{R}^2\to\mathbb{R}$ is some 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 here interested in understanding the limiting behavior of the solution q_μ to equation (1.1), as the mass μ vanishes. This is the so called Smoluchowski-Kramers approximation.

*Partially supported by the NSF grants DMS-1407615 and DMS-1712934.
†Partially supported by

It is well known (see [9] 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 coincides with the identity matrix), then $q_\mu(t)$ can be approximated with the solution of the first order equation

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

More precisely, for every fixed $T > 0$

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

Notice that here the case of an arbitrary number of degrees of freedom can be covered. 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 part, with the limiting equation (1.2) replaced by

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

The case of non constant friction has been widely studied recently (see [12] and [13] for example). They have considered the following system

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

for some h-dimensional Brownian motion $w(t)$. They have assumed that the coefficients $b : \mathbb{R}^k \to \mathbb{R}^k$, $\gamma : \mathbb{R}^k \to \mathbb{R}^{k\times k}$ and $\sigma : \mathbb{R}^k \to \mathbb{R}^{h\times k}$ are smooth and uniformly bounded and the smallest eigenvalue $\lambda_1(q)$ of the symmetric matrix $\gamma(q) + \gamma^\star(q)$ is strictly positive, uniformly with respect to $q \in \mathbb{R}^k$. Namely

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

They have proved that limit (1.3) is still valid, but now $q(t)$ is the solution of the modified equation

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

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

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

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

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

In [4], the case of a particle subject to a constant strength 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 semplicity 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 valid anymore.

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 if such continuous function is replaced by white noise, then there is no more convergence of q_μ to the solution of (1.4), as $\mu \downarrow 0$. Actually, while

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

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

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

so that

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

Because of this, in [4] the problem has been regularized, so that a suitable counterpart of the Smoluchowski-Kramers approximation has been proved. The first regularization consisted in introducing in equation (1.1) a small friction proportional to the velocity. Namely, the following equation has been considered

$$\left\{\begin{array}{l}\mu\,\ddot{q}_{\mu,\epsilon}(t) = b(q_{\mu,\epsilon}(t)) - A_\epsilon \dot{q}_{\mu,\epsilon}(t) + \sigma(q_{\mu,\epsilon}(t))\dot{w}(t),\\[10pt] q_{\mu,\epsilon}(0) = q \in \mathbb{R}^2,\qquad \dot{q}_{\mu,\epsilon}(0) = p \in \mathbb{R}^2,\end{array}\right.$$

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

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

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

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

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

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

Another approach to regularization (see also [15] for the case of non constant 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 the standard white noise $w(t)$, 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

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

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 has to be interpreted in Stratonovich sense.

In the present paper we are interested in the small mass limit in presence of a non-constant magnetic field. To this purpose we proceed by adding a small constant friction and we consider the regularized equation

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

We show that under suitable conditiuons on the coefficients b, σ and λ, the problem above 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 [12] and [13], so that we can take the small mass limit as μ goes to zero and we obtain that for every $\epsilon > 0$

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

where q_ϵ is the solution of the problem

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

After some computations, it turns out that q_ϵ solves the equation

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

for some mappings $\gamma : \mathbb{R}^2 \to \mathbb{R}$, $B, B_\epsilon : \mathbb{R}^2 \to \mathbb{R}^2$ and $\Sigma, \Sigma_\epsilon : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ that are explicitly given. This means that the motion of q_ϵ is made of a fast component on the level sets of λ and a slow transversal motion. Thus, by using a suitable generalization of the classical result of Freildin and Wentcell on averaging for Hamiltonian systems (see [11, Chapter 8] and [24]), we 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 suitable Markov process Y, whose generator is explicitly given.

2 Well-posedness of the regularized problem

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

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

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 in the equation above satisfy the following conditions. In fact, in Section 4 we will impose a more restrictive grown condition on λ.

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

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

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

Moreover

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

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

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

where

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

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 $k \geq 1$, equation (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

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

Notice that $\lambda_n : \mathbb{R}^2 \to \mathbb{R}$ 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 that

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

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

With these notations, we introduce the problem

$$\left\{\begin{array}{l} \mu\,\ddot{q}^n_{\mu,\epsilon}(t) = b(q^n_{\mu,\epsilon}(t)) - \Lambda_{\epsilon,n}(q^n_{\mu,\epsilon}(t))\beta_n(\dot{q}^n_{\mu,\epsilon}(t)) + \sigma(q^n_{\mu,\epsilon}(t))\,\dot{w}_t, \\[10pt] q^n_{\mu,\epsilon}(0) = q \in \mathbb{R}^2, \qquad \dot{q}^n_{\mu,\epsilon}(0) = p \in \mathbb{R}^2, \end{array}\right.$$

which can be rewritten as

$$\left\{\begin{array}{l} dq^n_{\mu,\epsilon}(t) = p^n_{\mu,\epsilon}(t)\,dt, \qquad q^n_{\mu,\epsilon}(0) = q \\[10pt] \mu dp^n_{\mu,\epsilon}(t) = \left[b(q^n_{\mu,\epsilon}(t)) - \Lambda_{\epsilon,n}(q^n_{\mu,\epsilon}(t))\beta_n p^n_{\mu,\epsilon}(t)\right] + \sigma(q^n_{\mu,\epsilon}(t))\,dw(t), \qquad p^n_{\mu,\epsilon}(t0) = p. \end{array}\right. \tag{2.7}$$

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 $(q^n_{\mu,\epsilon}, p^n_{\mu,\epsilon}) \in L^p(\Omega; C^1([0,T];\mathbb{R}^2) \times C([0,T];\mathbb{R}^2))$.

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

$$|q^n_{\mu,\epsilon}(t)|^{2k} + |p^n_{\mu,\epsilon}(t)|^{2k} = |q|^{2k} + |p|^{2k} + k\int_0^t |q^n_{\mu,\epsilon}(s)|^{2k-2}\langle q^n_{\mu,\epsilon}(s), p^n_{\mu,\epsilon}(s)\rangle\,ds$$

$$+\frac{k}{\mu}\int_0^t |p^n_{\mu,\epsilon}(s)|^{2k-2}\langle p^n_{\mu,\epsilon}(s), b(q^n_{\mu,\epsilon}(s) - \Lambda_{\epsilon,n}(q^n_{\mu,\epsilon}(s))\beta_n(p^n_{\mu,\epsilon}(s))\rangle\,ds$$

$$+\frac{k}{2\mu^2}\int_0^t |p^n_{\mu,\epsilon}(s)|^{2k-2}\mathrm{Tr}\left[\sigma\sigma^\star(q^n_{\mu,\epsilon}(s))\right]\,ds + \frac{k(k-1)}{2\mu^2}\int_0^t |p^n_{\mu,\epsilon}(s)|^{2k-4}|\sigma(q^n_{\mu,\epsilon}(s))p^n_{\mu,\epsilon}(s)|^2\,ds$$

$$+\frac{k}{\mu}\int_0^t |p^n_{\mu,\epsilon}(s)|^{2k-2}\langle p^n_{\mu,\epsilon}(s), \sigma(q^n_{\mu,\epsilon}(s))\,dw(s)\rangle.$$

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

$$|q^n_{\mu,\epsilon}(t)|^{2k} + |p^n_{\mu,\epsilon}(t)|^{2k} \leq |q|^{2k} + |p|^{2k} + c_{k,\mu}\int_0^t \left[|q^n_{\mu,\epsilon}(s)|^{2k} + |p^n_{\mu,\epsilon}(s)|^{2k}\right]\,ds$$

$$+\frac{k}{\mu}\int_0^t |p^n_{\mu,\epsilon}(s)|^{2k-2}\langle p^n_{\mu,\epsilon}(s), \sigma(q^n_{\mu,\epsilon}(s))\,dw(s)\rangle.$$

After we take expectation in both sides, due to the Gronwall lemma we obtain

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

Therefore, since

$$q^n_{\mu,\epsilon}(t) = q + \int_0^t p^n_{\mu,\epsilon}(s)\,ds,$$

and

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

due to (2.5), from (2.8) we obtain

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

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

$$\tau_n = \inf\left\{t \geq 0\ :\ |q^n_{\mu,\epsilon}(t)| \vee |p^n_{\mu,\epsilon}(t)| \geq n\right\},$$

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

$$(q^n_{\mu,\epsilon}(t), p^n_{\mu,\epsilon}(t)) = (q^m_{\mu,\epsilon}(t), p^m_{\mu,\epsilon}(t)), \qquad n < m, \quad t \leq \tau_n, \tag{2.10}$$

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

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

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

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

This implies that

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

and then, due to the arbitrariness of T, we conclude

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

In particular, if we set

$$(q_{\mu,\epsilon}(t), p_{\mu,\epsilon}(t)) = (q^n_{\mu,\epsilon}(t\wedge\tau_n), p^n_{\mu,\epsilon}(t\wedge\tau_n)), \qquad t \leq \tau,$$

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

It is immediate to check that for every $\epsilon > 0$ and $q \in \mathbb{R}^2$, the matrix $\Lambda_\epsilon(q)$ is invertible and

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

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

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

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

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

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

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

It is immediate to check that

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

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

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

and

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

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

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

Proof. Thanks to (3.3), we have

$$J^\epsilon_{11}(q) = \frac{\beta_0(q)}{\epsilon} + \beta_1(q)\int_0^\infty \cos(\lambda(q)r)e^{-\epsilon r}\,dr - \beta_2(q)\int_0^\infty \sin(\lambda(q)r)e^{-\epsilon r}\,dr$$

$$J^\epsilon_{22}(q) = \frac{\beta_0(q)}{\epsilon} - \beta_1(q)\int_0^\infty \cos(\lambda(q)r)e^{-\epsilon r}\,dr + \beta_2(q)\int_0^\infty \sin(\lambda(q)r)e^{-\epsilon r}\,dr$$

$$J^\epsilon_{12}(q) = J^\epsilon_{21}(q) = \beta_1(q)\int_0^\infty \sin(\lambda(q)r)e^{-\epsilon r}\,dr + \beta_2(q)\int_0^\infty \cos(\lambda(q)r)e^{-\epsilon r}\,dr.$$

Integrating by parts, we have

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

and

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

This allows to conclude that

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

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

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

Therefore, if we replace (3.6) and (3.7) in (3.2), we obtain

$$\begin{aligned} S^\epsilon_1(q) = &-\frac{2\epsilon\lambda(q)}{(\lambda^2(q)+\epsilon^2)^2}\left[\left(\frac{\beta_0(q)}{\epsilon} + \beta_1(q)\frac{\epsilon}{\lambda^2(q)+\epsilon^2} - \beta_2(q)\frac{\lambda(q)}{\lambda^2(q)+\epsilon^2}\right)\partial_1\lambda(q)\right. \\ &\left.+\left(\beta_1(q)\frac{\lambda(q)}{\lambda^2(q)+\epsilon^2} + \beta_2(q)\frac{\epsilon}{\lambda^2(q)+\epsilon^2}\right)\partial_2\lambda(q)\right] \\ &+\frac{\lambda^2(q)-\epsilon^2}{(\lambda^2(q)+\epsilon^2)^2}\left[\left(\beta_1(q)\frac{\lambda(q)}{\lambda^2(q)+\epsilon^2} + \beta_2(q)\frac{\epsilon}{\lambda^2(q)+\epsilon^2}\right)\partial_1\lambda(q)\right. \\ &\left.+\left(\frac{\beta_0(q)}{\epsilon} - \beta_1(q)\frac{\epsilon}{\lambda^2(q)+\epsilon^2} + \beta_2(q)\frac{\lambda(q)}{\lambda^2(q)+\epsilon^2}\right)\partial_2\lambda(q)\right], \end{aligned}$$

and

$$S_2^\epsilon(q) = -\frac{\lambda^2(q)-\epsilon^2}{(\lambda^2(q)+\epsilon^2)^2}\left[\left(\frac{\beta_0(q)}{\epsilon} + \beta_1(q)\frac{\epsilon}{\lambda^2(q)+\epsilon^2} - \beta_2(q)\frac{\lambda(q)}{\lambda^2(q)+\epsilon^2}\right)\partial_1\lambda(q)\right.$$

$$\left. + \left(\beta_1(q)\frac{\lambda(q)}{\lambda^2(q)+\epsilon^2} + \beta_2(q)\frac{\epsilon}{\lambda^2(q)+\epsilon^2}\right)\partial_2\lambda(q)\right]$$

$$-\frac{2\epsilon\lambda(q)}{(\lambda^2(q)+\epsilon^2)^2}\left[\left(\beta_1(q)\frac{\lambda(q)}{\lambda^2(q)+\epsilon^2} + \beta_2(q)\frac{\epsilon}{\lambda^2(q)+\epsilon^2}\right)\partial_1\lambda(q)\right.$$

$$\left. + \left(\frac{\beta_0(q)}{\epsilon} - \beta_1(q)\frac{\epsilon}{\lambda^2(q)+\epsilon^2} + \beta_2(q)\frac{\lambda(q)}{\lambda^2(q)+\epsilon^2}\right)\partial_2\lambda(q)\right],$$

Now, we define

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

and

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

With these notations, we have

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

$$+R_{11}^\epsilon(q)\partial_1\lambda(q) + R_{12}^\epsilon(q)\partial_2\lambda(q),$$

where

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

and

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

In a similar way, 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

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

and

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

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

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

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

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

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

Remark 3.2. 1. According to the 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 $\nabla\lambda : \mathbb{R}^2 \to \mathbb{R}^2$ Lipschitz-continuous.

2. In the same way, if we assume that $\nabla\lambda : \mathbb{R}^2 \to \mathbb{R}^2$ 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*

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

where q_ϵ is the solution of the problem

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

Proof. According to 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 thanks to [12, Theorem 2.4]. □

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

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

the quantity $\lambda(X_\epsilon(t))$ is not anymore conserved. However, for any fixed time interval $[0, T]$ and for every $k \geq 1$, we have

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

and, as an immediate consequence,

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

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

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

where q_ϵ is the solution of equation (3.14). As we 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 $\lambda : \mathbb{R}^2 \to \mathbb{R}$ is four times continuously differentiable, with bounded second derivative.*

3. *The mapping λ has only a finite number of critical points $x_1, \ldots, x_n$. The matrix of second derivatives $D^2\lambda(x_i)$ is non degenerate, for every $i = 1, \ldots, n$ and $\lambda(x_i) \neq \lambda(x_j)$, if $i \neq j$.*

4. *There exist three positive constants a_1, a_2, a_3 such that $\lambda(x) \geq a_1 |x|^2$, $|\nabla \lambda(x)| \geq a_2 |x|$ and $\Delta \lambda(x) \geq a_3$, for all $x \in \mathbb{R}^2$, with $|x|$ large enough.*

Remark 4.1. Remember that the function β_0 was defined as $[(\sigma\sigma^\star)^2_{11} + (\sigma\sigma^\star)^2_{22}]/4$. 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

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

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

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

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 $k = 1, \ldots, N(z)$, 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 intervals $I_1, \ldots I_n$ and vertices $O_1, \ldots, O_m$. 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 endpoint of the interval in the graph corresponding to the point at infinity.

In what follows, we shall denote by $\Pi : \mathbb{R}^2 \to \Gamma$ the *identification map*, that associates to every point $x \in \mathbb{R}^2$ the corresponding point $\Pi(x)$ on the graph Γ. We have $\Pi(x) = (\lambda(x), k(x))$, where $k(x)$ denotes the number of the interval 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) for 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

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

where the minimum is taken over all possible paths from y_1 to y_2, through every possible sequence of vertices $O_{i_1}, \ldots, O_{i_j}$, 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 rest 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 $t \to \pm\infty$.

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 we have 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

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

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

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

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

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

and

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

Therefore, recalling how $R^\epsilon(q)$ was defined in (3.8), (3.9), (3.10) and (3.11), we can conclude. □

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

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

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

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

This means that, as $\epsilon \downarrow 0$, part 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 $\epsilon \downarrow 0$.

If we apply 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 $q \in \mathbb{R}^2$

$$\mathcal{G}f(q) = \frac{1}{2}\mathrm{Tr}\left[\Sigma\Sigma^\star(q)D^2f(q)\right] + \langle Df(q), B(q)\rangle,$$

$$\mathcal{A}f(q) = \Sigma(q)^\star Df(q),$$

$$\mathcal{G}_\epsilon f(q) = \frac{1}{2}\mathrm{Tr}\left[\left(\epsilon\,\Sigma_\epsilon\Sigma_\epsilon^\star(q) + \Sigma\Sigma_\epsilon^\star(q) + \Sigma_\epsilon\Sigma^\star(q)\right)D^2f(q)\right] + \langle Df(q), B_\epsilon(q)\rangle,$$

and

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

We recall that the graph Γ is made of n intervals $I_1, \ldots, I_n$ and m verteces $O_1, \ldots, O_m$. For every $j = 1, \ldots, n$ 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

$$\alpha_j(z) = \oint_{C_j(z)} |\mathcal{A}\lambda(x)|^2 \, d\mu_{z,j}(x) = \oint_{C_j(z)} |\Sigma^\star(x)\nabla\lambda(x)|^2 \, d\mu_{z,j}(x),$$

$$\gamma_j(z) = \oint_{C_j(z)} \mathcal{G}\lambda(x) \, d\mu_{z,j}(x),$$

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

Definition 4.3. *For each interior vertex O_k and any segment I_j meeting at O_k (notation $I_j \sim O_k$), let ρ_{kj} be the positive constant defined by*

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

We denote by $D(L) \subset C(\Gamma)$ the set consisting all continuous 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

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

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 λ, along the edge I_j and the sign $\pm$ are taken if $\lambda > \lambda(O_k)$ or $\lambda < \lambda(O_k)$.

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

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

As proven in [11, 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 $(L, D(L))$, introducedd in Definition 4.3.*

Proof. In case 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 [11, 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 [24], under these more general assumptions, an averaging principle of the same type of the one described in [11, 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.

□

References

[1] J. Birrell, S. Hottovy, G. Volpe, J. Wehr, *Small mass limit of a Langevin equation on a manifold*, Annales Henri Poincaré 18 (2017), pp. 707–755.

[2] S. Cerrai, M. Freidlin, *On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom*, Probability Theory and Related Fields 135 (2006), pp. 363-394.

[3] S. Cerrai, M. Freidlin, *Smoluchowski-Kramers approximation for a general class of SPDE's*, Journal of Evolution Equations 6 (2006), pp. 657-689.

[4] S. Cerrai, M. Freidlin, *Small mass asymptotics for a charged particle in a magnetic field and longtime influence of small perturbations*, Journal of Statistical Physics 144 (2011), pp. 101–123.

[5] S. Cerrai, M. Freidlin, M. Salins, *On the Smoluchowski-Kramers approximation for SPDEs and its interplay with large deviations and long time behavior*, Discrete and Continuous Dynamical Systems, Series A, 37 (2017), pp. 33–76.

[6] S. Cerrai, M. Salins, *Smoluchowski-Kramers approximation and large deviations for infinite dimensional gradient systems*, Asymptotic Analysis 88 (2013), pp. 201-215.

[7] S. Cerrai, M. Salins, *Smoluchowski-Kramers approximation and large deviations for infinite dimensional non-gradient systems with applications to the exit problem*, Annals of Probability 44 (2016), pp. 2591–2642.

[8] S. Cerrai, M. Salins, *On the Smoluchowski-Kramers approximation for a system with an infinite number of degrees of freedom subject to a magnetic field*, Stochastic Processes and their Applications 127 (2017) pp. 273–303.

[9] M. Freidlin, *Some remarks on the Smoluchowski-Kramers approximation*, J. Statist. Phys. 117 (2004), pp. 617–634.

[10] M. Freidlin, W. Hu, *Smoluchowski–Kramers approximation in the case of variable friction*, Journal of Mathematical Sciences 179 (2011), pp. 184–207.

[11] M. Freidlin, A. Wentzell, RANDOM PERTURBATIONS OF DYNAMICAL SYSTEMS, third edition, Springer Verlag 2012.

[12] D. Herzog, S. Hottovy, G. Volpe, *The small-mass limit for Langevin dynamics with unbounded coefficients and positive friction*, Journal of Statistical Physics 163 (2016), pp. 659–673.

[13] S. Hottovy, A. McDaniel, G. Volpe, J. Wehr, *The Smoluchowski-Kramers limit of stochastic differential equations with arbitrary state-dependent friction*, Communications in Mathematical Physics 336 (2015), pp. 1259–1283.

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

[15] J. J. Lee, *Small mass asymptotics of a charged particle in a variable magnetic field*, Asymptotic Analysis 86 (2014), pp. 99–121.

[16] Y. Lv, A. Roberts, *Averaging approximation to singularly perturbed nonlinear stochastic wave equations*, Journal of Mathematical Physics 53 (2012), pp. 1–11.

[17] Y. Lv, A. Roberts, *Large deviation principle for singularly perturbed stochastic damped wave equations*, Stochastic Analysis and Applications 32 (2014), pp. 50-60.

[18] Y. Lv, A. Roberts, W. Wang, *Approximation of the random inertial manifold of singularly perturbed stochastic wave equations*, Stochastics and Dynamics 32, 2014.

[19] Y. Lv, W. Wang, *Limiting dynamics for stochastic wave equations*, Journal of Differential Equations 244, (2008), pp. 1–23.

[20] H. Nguyen, *The small-mass limit and white-noise limit of an infinite dimensional Generalized Langevin Equation*, arXiv:1804.09682 (2018).

[21] M. Salins, *Smoluchowski-Kramers approximation for the damped stochastic wave equation with multiplicative noise in any spatial dimension*, arXiv:1801.10538 (2018).

[22] M. Smoluchowski, *Drei Vortage über Diffusion Brownsche Bewegung und Koagulation von Kolloidteilchen*, Physik Zeit. 17 (1916), pp. 557-585.

[23] K. Spiliopoulos, *A note on the Smoluchowski-Kramers approximation for the Langevin equation with reflection*, Stochastics and Dynamics 7 (2007), pp. 141–152.

[24] Y. Zhu, *A generalization of the Freidlin-Wentcell theorem on averaging for Hamiltonian systems*, arXiv: