Journal of Mathematical Sciences, Vol. 179, No. 1, November, 2011

SMOLUCHOWSKI–KRAMERS APPROXIMATION IN THE CASE OF VARIABLE FRICTION

M. Freidlin *

University of Maryland at College Park, MD 20742, USA
mif@math.umd.edu

W. Hu

University of Maryland at College Park, MD 20742, USA
huwenqing@math.umd.edu

Dedicated to N. V. Krylov
on the occasion of his 70th birthday

We consider the small mass asymptotics (Smoluchowski–Kramers approximation) for the Langevin equation with a variable friction coefficient. The limit of the solution in the classical sense does not exist in this case. We study a modification of the Smoluchowski–Kramers approximation. Some applications of the Smoluchowski–Kramers approximation to problems with fast oscillating or discontinuous coefficients are considered. Bibliography: 15 *titles.*

1 Introduction

The Langevin equation

$$\mu\ddot{\boldsymbol{q}}_t^\mu = \boldsymbol{b}(\boldsymbol{q}_t^\mu) - \lambda\dot{\boldsymbol{q}}_t^\mu + \sigma(\boldsymbol{q}_t^\mu)\dot{\boldsymbol{W}}_t, \quad \boldsymbol{q}_0^\mu = \boldsymbol{q} \in \mathbb{R}^n, \quad \dot{\boldsymbol{q}}_0^\mu = \boldsymbol{p} \in \mathbb{R}^n, \tag{1.1}$$

describes the motion of a particle of mass μ in a force field $\boldsymbol{b}(\boldsymbol{q})$, $\boldsymbol{q} \in \mathbb{R}^n$, subject to random fluctuations and to a friction proportional to the velocity. Here, $\boldsymbol{W}_t$ is the standard Wiener process in $\mathbb{R}^n$, $\lambda > 0$ is the friction coefficient. The vector field $\boldsymbol{b}(\boldsymbol{q})$ and the matrix function $\sigma(\boldsymbol{q})$ are assumed to be continuously differentiable and bounded together with their first order derivatives. The matrix $a(\boldsymbol{q}) = (a_{ij}(\boldsymbol{q})) = \sigma(\boldsymbol{q})\sigma^*(\boldsymbol{q})$ is assumed to be nonsingular.

Put $\boldsymbol{p}_t^\mu = \dot{\boldsymbol{q}}_t^\mu$. Then (1.1) can be written as the first order system:

$$\begin{cases} \dot{\boldsymbol{q}}_t^\mu = \boldsymbol{p}_t^\mu, \\ \dot{\boldsymbol{p}}_t^\mu = \dfrac{1}{\mu}\boldsymbol{b}(\boldsymbol{q}_t^\mu) - \dfrac{\lambda}{\mu}\boldsymbol{p}_t^\mu + \dfrac{1}{\mu}\sigma(\boldsymbol{q}_t^\mu)\dot{\boldsymbol{W}}_t. \end{cases} \tag{1.2}$$

* To whom the correspondence should be addressed.

Translated from *Problems in Mathematical Analysis* **61**, October 2011, pp. 155–174.

The diffusion process $(\boldsymbol{p}_t^\mu, \boldsymbol{q}_t^\mu) = \boldsymbol{X}_t^\mu$ in $\mathbb{R}^{2n}$ is governed by the generator L:

$$Lu(\boldsymbol{p},\boldsymbol{q}) = \frac{1}{2\mu^2}\sum_{i,j=1}^{n} a_{ij}(\boldsymbol{q})\frac{\partial^2 u}{\partial p_i \partial p_j} + \frac{1}{\mu}(\boldsymbol{b}(\boldsymbol{q}) - \lambda \boldsymbol{p})\cdot \nabla_{\boldsymbol{p}} u + \boldsymbol{p}\cdot\nabla_{\boldsymbol{q}} u.$$

Note that, since the functions $\boldsymbol{q}_t^\mu$ are continuously differentiable with probability one,

$$\int_0^t \sigma_{ij}(\boldsymbol{q}_s^\mu)dW_s^j = \sigma_{ij}(\boldsymbol{q}_t^\mu)W_t^j - \int_0^t W_s^j\left(\nabla_{\boldsymbol{q}}\sigma_{ij}(\boldsymbol{q}_s^\mu)\cdot \boldsymbol{p}_s^\mu\right)ds.$$

This allows us to consider Equations (1.2) for each trajectory $\boldsymbol{W}_t$ individually, and there is no necessity in the introduction of a stochastic integral. In particular, if (1.2) is considered as a stochastic differential equation, stochastic integrals in the Itô and Stratonovich senses coincide:

$$\int_0^t \sigma(\boldsymbol{q}_s^\mu)d\boldsymbol{W}_s = \int_0^t \sigma(\boldsymbol{q}_s^\mu)\circ d\boldsymbol{W}_s.$$

It is usually assumed that the friction coefficient λ is constant. Under this assumption, one can prove that $\boldsymbol{q}_t^\mu$ converges in probability as $\mu \downarrow 0$ uniformly on each finite time interval $[0,T]$ to an n-dimensional diffusion process $\boldsymbol{q}_t$: for any $\kappa, T > 0$ and any $\boldsymbol{p}_0^\mu = \boldsymbol{p} \in \mathbb{R}^n$ fixed,

$$\lim_{\mu\downarrow 0}\mathbf{P}\big(\max_{0\leqslant t\leqslant T}|\boldsymbol{q}_t^\mu - \boldsymbol{q}_t|_{\mathbb{R}^d} > \kappa\big) = 0.$$

Here, $\boldsymbol{q}_t$ is the solution of the equation

$$\dot{\boldsymbol{q}}_t = \frac{1}{\lambda}\boldsymbol{b}(\boldsymbol{q}_t) + \frac{1}{\lambda}\sigma(\boldsymbol{q}_t)\dot{\boldsymbol{W}}_t, \quad \boldsymbol{q}_0 = \boldsymbol{q}_0^\mu = \boldsymbol{q}\in\mathbb{R}^n. \tag{1.3}$$

The stochastic term in (1.3) should be understood in the Itô sense.

The approximation of $\boldsymbol{q}_t^\mu$ by $\boldsymbol{q}_t$ for $0 < \mu \ll 1$ called the *Smoluchowski–Kramers approximation*. This is the main justification for replacement of the second order equation (1.1) by the first order equation (1.3). The price for such a simplification, in particular, consists of certain nonuniversality of Equation (1.3): the white noise in (1.1) is an idealization of a more regular stochastic process $\dot{\boldsymbol{W}}_t^\delta$ with correlation radius $\delta \ll 1$ converging to $\dot{\boldsymbol{W}}_t$ as $\delta\downarrow 0$. Let $\boldsymbol{q}_t^{\mu,\delta}$ be the solution of Equation (1.1) with $\dot{\boldsymbol{W}}_t$ replaced by $\dot{\boldsymbol{W}}_t^\delta$. Then the limit of $\boldsymbol{q}_t^{\mu,\delta}$ as $\mu,\delta\downarrow 0$ depends on the relation between μ and δ. Say, if first $\delta\downarrow 0$ and then $\mu\downarrow 0$, the stochastic integral in (1.3) should be understood in the Itô sense; if first $\mu\downarrow 0$ and then $\delta \downarrow 0$, $\boldsymbol{q}_t^{\mu,\delta}$ converges to the solution of (1.3) with stochastic integral in the Stratonovich sense (cf., for example, [1].)

Consider the case of a variable friction coefficient $\lambda = \lambda(\boldsymbol{q})$. We assume that $\lambda(\boldsymbol{q})$ has continuous bounded derivatives and $0 < \lambda_0 \leqslant \lambda(\boldsymbol{q}) \leqslant \Lambda < \infty$. It turns out, as we will see in the next section, that, in this case, the solution $\boldsymbol{q}_t^\mu$ of (1.1) does not converge, in general, to the solution of (1.3) with $\lambda = \lambda(\boldsymbol{q})$, so that the Smoluchowski–Kramers approximation should be modified. In order to do this, we consider Equation (1.1) with $\dot{\boldsymbol{W}}_t$ replaced by $\dot{\boldsymbol{W}}_t^\delta$ described above:

$$\mu\ddot{\boldsymbol{q}}_t^{\mu,\delta} = \boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta}) - \lambda(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{q}}_t^{\mu,\delta} + \sigma(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{W}}_t^\delta, \quad \boldsymbol{q}_0^{\mu,\delta} = \boldsymbol{q}, \quad \dot{\boldsymbol{q}}_0^{\mu,\delta} = \boldsymbol{p}. \tag{1.4}$$

We prove that after such a regularization, the solution of (1.4) has a limit $\widetilde{\boldsymbol{q}}_t^\delta$ as $\mu \downarrow 0$, and $\widetilde{\boldsymbol{q}}_t^\delta$ is a unique solution of the equation obtained from (1.4) as $\mu = 0$:

$$\dot{\widetilde{\boldsymbol{q}}}_t^\delta = \frac{1}{\lambda(\widetilde{\boldsymbol{q}}_t^\delta)}\boldsymbol{b}(\widetilde{\boldsymbol{q}}_t^\delta) + \frac{1}{\lambda(\widetilde{\boldsymbol{q}}_t^\delta)}\sigma(\widetilde{\boldsymbol{q}}_t^\delta)\dot{\boldsymbol{W}}_t^\delta, \quad \widetilde{\boldsymbol{q}}_0^\delta = \boldsymbol{q}. \tag{1.5}$$

Now we can take $\delta \downarrow 0$ in (1.5). As a result, we get the equation

$$\dot{\widehat{\boldsymbol{q}}}_t = \frac{1}{\lambda(\widehat{\boldsymbol{q}}_t)}\boldsymbol{b}(\widehat{\boldsymbol{q}}_t) + \frac{1}{\lambda(\widehat{\boldsymbol{q}}_t)}\sigma(\widehat{\boldsymbol{q}}_t) \circ \dot{\boldsymbol{W}}_t, \quad \widehat{\boldsymbol{q}}_0 = \boldsymbol{q}, \tag{1.6}$$

where the stochastic term should be understood in the Stratonovich sense. So the regularization leads to a modified Smoluchowski–Kramers equation (1.6). We prove this in Section 3.

Some applications of the Smoluchowski–Kramers approximation are considered in Sections 4 and 5: the case of fast oscillating in the spatial variable, periodic or stochastic, friction coefficient is studied; gluing condition at the discontinuity points of the friction coefficient are considered. In Section 6, we briefly consider some remarks and generalizations.

Notation. We use $|\bullet|_{\mathbb{R}^d}$ to denote the standard Euclidean norm in $\mathbb{R}^d$. When $d = 1$, we set $|\bullet|_{\mathbb{R}^1} = |\bullet|$. For a vector-valued function $\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_d(\boldsymbol{x}))$, $\boldsymbol{x} \in \mathbb{R}^d$, we set $\|\boldsymbol{f}\|_\infty = \max\limits_{1 \leqslant i \leqslant d} \|f_i\|_\infty = \max\limits_{1 \leqslant i \leqslant d} \sup\limits_{\boldsymbol{x} \in \mathbb{R}^d} |f_i(\boldsymbol{x})|$. All the vectors are marked with either bold letters or with an arrow on it.

2 Some Estimates. The Classical Smoluchowski–Kramers Approximation Does not Work for Variable Friction Coefficients

We consider the system

$$\mu \ddot{\boldsymbol{q}}_t^\mu = \boldsymbol{b}(\boldsymbol{q}_t^\mu) - \lambda(\boldsymbol{q}_t^\mu)\dot{\boldsymbol{q}}_t^\mu + \dot{\boldsymbol{W}}_t, \quad \boldsymbol{q}_0^\mu = \boldsymbol{q} \in \mathbb{R}^d, \quad \dot{\boldsymbol{q}}_0^\mu = \boldsymbol{p} \in \mathbb{R}^d. \tag{2.1}$$

Here, $\infty > \Lambda \geqslant \lambda(\bullet) \geqslant \lambda_0 > 0$ is a function of $\boldsymbol{q}_t^\mu$. We assume that the function $\lambda(\bullet)$ and vector field $\boldsymbol{b}(\bullet)$ are continuously differentiable and bounded together with their first derivatives. The process $\boldsymbol{W}_t$ is the standard Wiener process in $\mathbb{R}^d$. For simplicity of calculations, we consider here the case where the diffusion matrix $a(\bullet)$ is the identity (cf. (1.1)). The case of general diffusion matrix can be considered in a similar way, and we will briefly mention it in Section 6.

Let $\boldsymbol{p}_t^\mu = \dot{\boldsymbol{q}}_t^\mu$. Then (2.1) is equivalent to the system

$$\begin{cases} \dot{\boldsymbol{q}}_t^\mu = \boldsymbol{p}_t^\mu, \\ \dot{\boldsymbol{p}}_t^\mu = \dfrac{1}{\mu}\boldsymbol{b}(\boldsymbol{q}_t^\mu) - \dfrac{\lambda(\boldsymbol{q}_t^\mu)}{\mu}\boldsymbol{p}_t^\mu + \dfrac{1}{\mu}\dot{\boldsymbol{W}}_t. \end{cases} \tag{2.2}$$

Then

$$\frac{d}{dt}\left(e^{\frac{1}{\mu}\int\limits_0^t \lambda(\boldsymbol{q}_s^\mu)ds}\boldsymbol{p}_t^\mu\right) = e^{\frac{1}{\mu}\int\limits_0^t \lambda(\boldsymbol{q}_s^\mu)ds}\left(\dot{\boldsymbol{p}}_t^\mu + \frac{1}{\mu}\lambda(\boldsymbol{q}_t^\mu)\boldsymbol{p}_t^\mu\right) = e^{\frac{1}{\mu}\int\limits_0^t \lambda(\boldsymbol{q}_s^\mu)ds}\left(\frac{1}{\mu}\boldsymbol{b}(\boldsymbol{q}_t^\mu) + \frac{1}{\mu}\dot{\boldsymbol{W}}_t\right)$$

and

$$e^{\frac{1}{\mu}\int\limits_0^t \lambda(\boldsymbol{q}_s^\mu)ds}\boldsymbol{p}_t^\mu - \boldsymbol{p} = \frac{1}{\mu}\int\limits_0^t e^{\frac{1}{\mu}\int\limits_0^s \lambda(\boldsymbol{q}_r^\mu)dr}\boldsymbol{b}(\boldsymbol{q}_s^\mu)ds + \frac{1}{\mu}\int\limits_0^t e^{\frac{1}{\mu}\int\limits_0^s \lambda(\boldsymbol{q}_r^\mu)dr}d\boldsymbol{W}_s. \tag{2.3}$$

For notational convenience we introduce the function

$$A(\mu, t) = \int\limits_0^t \lambda(\boldsymbol{q}_s^\mu)ds.$$

It is clear that $t\Lambda \geqslant A(\mu, t) \geqslant t\lambda_0$. Using (2.3), we find

$$\boldsymbol{p}_t^\mu = e^{-\frac{1}{\mu}A(\mu,t)}\left(\boldsymbol{p} + \frac{1}{\mu}\int\limits_0^t e^{\frac{1}{\mu}A(\mu,s)}\boldsymbol{b}(\boldsymbol{q}_s^\mu)ds + \frac{1}{\mu}\int\limits_0^t e^{\frac{1}{\mu}A(\mu,s)}d\boldsymbol{W}_s\right).$$

Therefore,

$$\boldsymbol{q}_t^\mu = \boldsymbol{q} + \int\limits_0^t \boldsymbol{p}_s^\mu ds = \boldsymbol{q} + \boldsymbol{p}\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}ds + \frac{1}{\mu}\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int\limits_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds$$

$$+\frac{1}{\mu}\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int\limits_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds = \boldsymbol{q} + \boldsymbol{\alpha}(\mu) + \boldsymbol{\beta}(\mu) + \boldsymbol{\gamma}(\mu). \tag{2.4}$$

Here, $\boldsymbol{\alpha}(\mu), \boldsymbol{\beta}(\mu), \boldsymbol{\gamma}(\mu)$ are three (vector) functions on the right-hand side of (2.4):

$$\boldsymbol{\alpha}(\mu) = \boldsymbol{p}\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}ds,$$

$$\boldsymbol{\beta}(\mu) = \frac{1}{\mu}\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int\limits_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds,$$

$$\boldsymbol{\gamma}(\mu) = \frac{1}{\mu}\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int\limits_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds.$$

In the following, we use the relation

$$\frac{d}{dt}\left(e^{-\frac{1}{\mu}A(\mu,t)}\right) = -\frac{1}{\mu}e^{-\frac{1}{\mu}A(\mu,t)}\frac{dA(\mu,t)}{dt} = -\frac{1}{\mu}e^{-\frac{1}{\mu}A(\mu,t)}\lambda(\boldsymbol{q}_t^\mu). \tag{2.5}$$

We also use the estimates

$$\frac{\mu}{c\Lambda}(1 - e^{-\frac{c\Lambda t}{\mu}}) = \int\limits_0^t e^{-\frac{c\Lambda s}{\mu}}ds \leqslant \int\limits_0^t e^{-\frac{c}{\mu}A(\mu,s)}ds \leqslant \int\limits_0^t e^{-\frac{c\lambda_0 s}{\mu}}ds = \frac{\mu}{c\lambda_0}(1 - e^{-\frac{c\lambda_0 t}{\mu}}) \leqslant \frac{\mu}{c\lambda_0}, \tag{2.6}$$

$$\frac{\mu}{c\Lambda}(1-e^{-\frac{c\Lambda t}{\mu}}) = \int\limits_0^t e^{-\frac{c\Lambda(t-s)}{\mu}}ds \leqslant \int\limits_0^t e^{-\frac{c}{\mu}(A(\mu,t)-A(\mu,s))}ds$$

$$\leqslant \int\limits_0^t e^{-\frac{c\lambda_0(t-s)}{\mu}}ds = \frac{\mu}{c\lambda_0}(1-e^{-\frac{c\lambda_0 t}{\mu}}) \leqslant \frac{\mu}{c\lambda_0}. \tag{2.7}$$

Here, c is a positive constant.

In this section, we get some bounds for $\boldsymbol{\alpha}(\mu)$, $\boldsymbol{\beta}(\mu)$, $\boldsymbol{\gamma}(\mu)$ which show, in particular, that the classical Smoluchowski–Kramers approximation does not hold in the case of variable friction. These bounds will be used to obtain a modified Smoluchowski–Kramers approximation.

2.1 Estimates of $\boldsymbol{\alpha}(\mu)$

By (2.5), we have

$$\boldsymbol{\alpha}(\mu) = \boldsymbol{p}\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}ds = \boldsymbol{p}\int\limits_0^t (-\mu)\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d(e^{-\frac{1}{\mu}A(\mu,s)})$$

$$= -\boldsymbol{p}\mu\left[\frac{e^{-\frac{1}{\mu}A(\mu,t)}}{\lambda(\boldsymbol{q}_t^\mu)} - \frac{1}{\lambda(\boldsymbol{q})} - \int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}d(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)})\right].$$

Let

$$R_{\boldsymbol{\alpha}}(\mu) = \mu\left[\frac{e^{-\frac{1}{\mu}A(\mu,t)}}{\lambda(\boldsymbol{q}_t^\mu)} - \frac{1}{\lambda(\boldsymbol{q})}\right].$$

It is easy to estimate $|R_{\boldsymbol{\alpha}}(\mu)| \leqslant \mu/\lambda_0$. Therefore, $|R_{\boldsymbol{\alpha}}(\mu)| \to 0$ as $\mu \downarrow 0$.

Let

$$(I) = \int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}d(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}).$$

We have

$$(I) = -\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}\frac{1}{\lambda^2(q_s^\mu)}\nabla\lambda(q_s^\mu)\cdot\boldsymbol{p}_s^\mu ds = -\int\limits_0^t e^{-\frac{1}{\mu}A(\mu,s)}\frac{1}{\lambda^2(q_s^\mu)}e^{-\frac{1}{\mu}A(\mu,s)}\nabla\lambda(q_s^\mu)$$

$$\times\left(\boldsymbol{p} + \frac{1}{\mu}\int\limits_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr + \frac{1}{\mu}\int\limits_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds = (I_1)+(I_2)+(I_3).$$

Here,

$$(I_1) = -\boldsymbol{p}\cdot\int\limits_0^t e^{-\frac{2}{\mu}A(\mu,s)}\frac{\nabla\lambda(\boldsymbol{q}_s^\mu)}{\lambda^2(\boldsymbol{q}_s^\mu)}ds,$$

$$(I_2) = -\frac{1}{\mu}\int\limits_0^t e^{-\frac{2}{\mu}A(\mu,s)}\frac{1}{\lambda^2(\boldsymbol{q}_s^\mu)}\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\left(\int\limits_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds,$$

$$(I_3) = -\frac{1}{\mu}\int_0^t e^{-\frac{2}{\mu}A(\mu,s)}\frac{1}{\lambda^2(q_s^\mu)}\nabla\lambda(q_s^\mu)\cdot\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds.$$

Using (2.6) and (2.7), we derive

$$|(I_1)| \leqslant \frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}|\boldsymbol{p}|_{\mathbb{R}^d}\int_0^t e^{-\frac{2}{\mu}\lambda_0 s}ds \leqslant \frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}|\boldsymbol{p}|_{\mathbb{R}^d}\frac{\mu}{2\lambda_0},$$

$$\begin{aligned}
|(I_2)| &\leqslant \frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\frac{1}{\mu}\int_0^t e^{-\frac{2}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}dr\right)ds \\
&\leqslant \frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\frac{1}{\mu}\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(s-r)\lambda_0}dr\right)e^{-\frac{1}{\mu}\lambda_0 s}ds \\
&= \frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\frac{1}{\mu}\int_0^t \frac{\mu}{\lambda_0}(1-e^{-\frac{\lambda_0 s}{\mu}})e^{-\frac{\lambda_0 s}{\mu}}ds \\
&\leqslant \frac{\|\nabla\lambda\|_\infty}{\lambda_0^3}\|\boldsymbol{b}\|_\infty\int_0^t e^{-\frac{\lambda_0 s}{\mu}}ds \leqslant \frac{\|\nabla\lambda\|_\infty}{\lambda_0^3}\|\boldsymbol{b}\|_\infty\frac{\mu}{\lambda_0}.
\end{aligned}$$

Since

$$|(I_3)| \leqslant \frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\frac{1}{\mu}\left|\int_0^t e^{-\frac{1}{2\mu}A(\mu,s)}\left(\int_0^s e^{-\frac{1}{2\mu}A(\mu,s)}e^{-\frac{1}{\mu}A(\mu,s)+\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds\right|_{\mathbb{R}^d},$$

we can estimate, by the Cauchy–Schwarz inequality and (2.6), (2.7),

$$\begin{aligned}
\mathbf{E}|(I_3)|^2 &\leqslant \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2\frac{1}{\mu^2}\mathbf{E}\left|\int_0^t e^{-\frac{1}{2\mu}A(\mu,s)}\left(\int_0^s e^{-\frac{1}{2\mu}A(\mu,s)}e^{-\frac{1}{\mu}A(\mu,s)+\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds\right|_{\mathbb{R}^d}^2 \\
&\leqslant \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2\frac{1}{\mu^2}\mathbf{E}\left(\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}ds\right)\left(\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left|\int_0^s e^{-\frac{1}{\mu}A(\mu,s)+\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right|_{\mathbb{R}^d}^2 ds\right) \\
&\leqslant \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2\frac{1}{\mu^2}\left(\int_0^t e^{-\frac{\lambda_0 s}{\mu}}ds\right)\left(\int_0^t e^{-\frac{\lambda_0 s}{\mu}}\mathbf{E}\left|\int_0^s e^{-\frac{1}{\mu}A(\mu,s)+\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right|_{\mathbb{R}^d}^2 ds\right) \\
&= \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2\frac{1}{\mu^2}\left(\int_0^t e^{-\frac{\lambda_0 s}{\mu}}ds\right)\left(\int_0^t e^{-\frac{\lambda_0 s}{\mu}}\left(\int_0^s \mathbf{E}e^{-\frac{2}{\mu}A(\mu,s)+\frac{2}{\mu}A(\mu,r)}dr\right)ds\right) \\
&\leqslant \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2\frac{1}{\mu^2}\left(\int_0^t e^{-\frac{\lambda_0 s}{\mu}}ds\right)\left(\int_0^t e^{-\frac{\lambda_0 s}{\mu}}\left(\int_0^s e^{-\frac{2\lambda_0 s}{\mu}+\frac{2\lambda_0 r}{\mu}}dr\right)ds\right)
\end{aligned}$$

$$\leqslant \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2 \frac{1}{\mu^2}\frac{\mu}{\lambda_0}\left(\int_0^t e^{-\frac{\lambda_0 s}{\mu}}\frac{\mu}{2\lambda_0}ds\right) \leqslant \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2 \frac{\mu}{2\lambda_0^3}.$$

Combining these estimates, we see that $\mathbf{E}|(I)|^2 \to 0$ as $\mu \downarrow 0$. So, $\mathbf{E}|\boldsymbol{\alpha}(\mu)|^2_{\mathbb{R}^d} \to 0$ as $\mu \downarrow 0$ for any $|\boldsymbol{p}|_{\mathbb{R}^d} < \infty$.

2.2 Estimates of $\boldsymbol{\beta}(\mu)$

By (2.5), we have

$$\begin{aligned}
\boldsymbol{\beta}(\mu) &= \frac{1}{\mu}\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds \\
&= \frac{1}{\mu}\int_0^t\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\left(-\frac{\mu}{\lambda(\boldsymbol{q}_s^\mu)}\right)d(e^{-\frac{1}{\mu}A(\mu,s)}) \\
&= \int_0^t\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\left(-\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\right)d(e^{-\frac{1}{\mu}A(\mu,s)}) \\
&= -\frac{e^{-\frac{1}{\mu}A(\mu,s)}}{\lambda(\boldsymbol{q}_s^\mu)}\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\Bigg|_{s=0}^{s=t} + \int_0^t e^{-\frac{1}{\mu}A(\mu,s)}d\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\right) \\
&= -\frac{e^{-\frac{1}{\mu}A(\mu,t)}}{\lambda(\boldsymbol{q}_t^\mu)}\int_0^t e^{\frac{1}{\mu}A(\mu,s)}\boldsymbol{b}(\boldsymbol{q}_s^\mu)ds + \int_0^t\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds \\
&\quad + \int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)d\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\right) \\
&= \boldsymbol{R}_{\boldsymbol{\beta}}(\mu) + \int_0^t\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds + (\vec{II}).
\end{aligned}$$

It is easy to see that

$$|\boldsymbol{R}_{\boldsymbol{\beta}}(\mu)|_{\mathbb{R}^d} \leqslant \frac{\|\boldsymbol{b}\|_\infty}{\lambda_0}\int_0^t e^{-\frac{\lambda_0}{\mu}(t-s)}ds = \frac{\|\boldsymbol{b}\|_\infty}{\lambda_0}\frac{\mu}{\lambda_0}.$$

We also have

$$\begin{aligned}
(\vec{II}) &= -\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\frac{1}{\lambda^2(\boldsymbol{q}_s^\mu)}\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\boldsymbol{p}_s^\mu ds \\
&= -\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\frac{1}{\lambda^2(\boldsymbol{q}_s^\mu)}\nabla\lambda(\boldsymbol{q}_s^\mu)
\end{aligned}$$

$$
e^{-\frac{1}{\mu}A(\mu,s)}\left(\boldsymbol{p}+\frac{1}{\mu}\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr+\frac{1}{\mu}\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds
$$
$$
=(\vec{II_1})+(\vec{II_2})+(\vec{II_3}).
$$

Here,

$$
(\vec{II_1})=-\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\boldsymbol{p}ds,
$$

$$
(\vec{II_2})=-\frac{1}{\mu}\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds,
$$

$$
(\vec{II_3})=-\frac{1}{\mu}\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds.
$$

We conclude that

$$
|(\vec{II_1})|_{\mathbb{R}^d}\leqslant\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}|\boldsymbol{p}|_{\mathbb{R}^d}\|\boldsymbol{b}\|_\infty\int_0^t e^{-\frac{\lambda_0 s}{\mu}}\left(\int_0^s e^{-\frac{(s-r)\lambda_0}{\mu}}dr\right)ds\leqslant\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}|\boldsymbol{p}|_{\mathbb{R}^d}\|\boldsymbol{b}\|_\infty\frac{\mu^2}{\lambda_0^2},
$$

$$
|(\vec{II_2})|_{\mathbb{R}^d}\leqslant\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty^2\int_0^t\left(\int_0^s e^{-\frac{(s-r)\lambda_0}{\mu}}dr\right)^2 ds\leqslant\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty^2\frac{\mu t}{\lambda_0^2};
$$

$$
\mathbf{E}|(\vec{II_3})|^2_{\mathbb{R}^d}\leqslant\left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\mathbf{E}\left|\int_0^t e^{-\frac{2}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}dr\right)\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds\right|^2_{\mathbb{R}^d}
$$
$$
=\left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\mathbf{E}\left|\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dr\right)\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right)ds\right|^2_{\mathbb{R}^d}
$$
$$
\leqslant\left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\mathbf{E}\left(\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dr\right)^2 ds\right)
$$
$$
\times\left(\int_0^t\left|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right|^2_{\mathbb{R}^d}ds\right)
$$
$$
\leqslant\left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\left(\int_0^t\left(\int_0^s e^{-\frac{(s-r)\lambda_0}{\mu}}dr\right)^2 ds\right)\left(\int_0^t\mathbf{E}\left|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right|^2_{\mathbb{R}^d}ds\right)
$$

$$\leqslant \left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2 \left(\int_0^t \left(\int_0^s e^{-\frac{(s-r)\lambda_0}{\mu}} dr\right)^2 ds\right)\left(\int_0^t \left(\int_0^s e^{-\frac{2(s-r)\lambda_0}{\mu}} dr\right) ds\right)$$

$$\leqslant \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2 \left(\frac{t}{\lambda_0}\right)^2 \left(\frac{\mu t}{2\lambda_0}\right).$$

Combining these estimates we see that $\mathbf{E}|(\vec{II})|^2_{\mathbb{R}^d} \to 0$ as $\mu \downarrow 0$. This implies that

$$\mathbf{E}\left|\boldsymbol{\beta}(\mu) - \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)} ds\right|^2_{\mathbb{R}^d} \to 0 \quad \text{as } \mu \downarrow 0.$$

2.3 Estimates of $\boldsymbol{\gamma}(\mu)$ - the reason why the classical Smoluchowski–Kramers approximation does not work

We show that

$$\mathbf{E}\left|\boldsymbol{\gamma}(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s\right|^2_{\mathbb{R}^d},$$

in general, does not tend to 0 as $\mu \downarrow 0$. Therefore, the Smoluchowski–Kramers approximation does not work in the case of purely white noise perturbation.

By (2.5), we have

$$\begin{aligned}
\boldsymbol{\gamma}(\mu) &= \frac{1}{\mu}\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r\right) ds \\
&= \frac{1}{\mu}\int_0^t \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r\right)\left(-\frac{\mu}{\lambda(\boldsymbol{q}_s^\mu)}\right) d(e^{-\frac{1}{\mu}A(\mu,s)}) \\
&= -\left[\frac{\int_0^t e^{\frac{1}{\mu}A(\mu,s)} d\boldsymbol{W}_s}{\lambda(\boldsymbol{q}_t^\mu)} e^{-\frac{1}{\mu}A(\mu,t)} - \int_0^t e^{-\frac{1}{\mu}A(\mu,s)} d\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r\right)\right] \\
&= -\frac{\int_0^t e^{\frac{1}{\mu}A(\mu,s)} d\boldsymbol{W}_s}{\lambda(\boldsymbol{q}_t^\mu)} e^{-\frac{1}{\mu}A(\mu,t)} + \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s \int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r\right) d\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\right) \\
&= \boldsymbol{R}_{\boldsymbol{\gamma}}(\mu) + \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s + (\vec{III}).
\end{aligned}$$

It is easy to check that

$$\mathbf{E}|\boldsymbol{R}_{\boldsymbol{\gamma}}(\mu)|^2_{\mathbb{R}^d} \leqslant \frac{1}{\lambda_0^2}\int_0^t e^{-\frac{2\lambda_0(t-s)}{\mu}} ds \leqslant \frac{\mu}{2\lambda_0^3}.$$

We have

$$
\begin{aligned}
(I\vec{I}I) &= \int_0^t e^{-\frac{1}{\mu}A(\mu,s)} \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r \right) \left(-\frac{1}{\lambda^2(\boldsymbol{q}_s^\mu)} \right) \nabla\lambda(\boldsymbol{q}_s^\mu) \cdot \boldsymbol{p}_s^\mu ds \\
&= (I\vec{I}I_1) + (I\vec{I}I_2) + (I\vec{I}I_3)
\end{aligned}
$$

where

$$
(I\vec{I}I_1) = -\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)} \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r \right) \nabla\lambda(\boldsymbol{q}_s^\mu) \cdot \boldsymbol{p} ds,
$$

$$
(I\vec{I}I_2) = -\frac{1}{\mu}\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)} \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r \right) \nabla\lambda(\boldsymbol{q}_s^\mu) \cdot \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} \boldsymbol{b}(\boldsymbol{q}_r^\mu) dr \right) ds,
$$

$$
(I\vec{I}I_3) = -\frac{1}{\mu}\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)} \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r \right) \nabla\lambda(\boldsymbol{q}_s^\mu) \cdot \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r \right) ds.
$$

We can estimate

$$
\begin{aligned}
\mathbf{E}|(I\vec{I}I_1)|_{\mathbb{R}^d}^2 &\leqslant \left(\frac{|\boldsymbol{p}|_{\mathbb{R}^d} \|\nabla\lambda\|_\infty}{\lambda_0^2} \right)^2 \mathbf{E} \left| \int_0^t e^{-\frac{1}{\mu}A(\mu,s)} \left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))} d\boldsymbol{W}_r \right) ds \right|_{\mathbb{R}^d}^2 \\
&\leqslant \left(\frac{|\boldsymbol{p}|_{\mathbb{R}^d} \|\nabla\lambda\|_\infty}{\lambda_0^2} \right)^2 \mathbf{E} \left(\int_0^t e^{-\frac{2}{\mu}A(\mu,s)} ds \right) \left(\int_0^t \left| \int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))} d\boldsymbol{W}_r \right|_{\mathbb{R}^d}^2 ds \right) \\
&\leqslant \left(\frac{|\boldsymbol{p}|_{\mathbb{R}^d} \|\nabla\lambda\|_\infty}{\lambda_0^2} \right)^2 \left(\int_0^t e^{-\frac{2\lambda_0 s}{\mu}} ds \right) \left(\int_0^t \left(\int_0^s e^{-\frac{2\lambda_0(s-r)}{\mu}} dr \right) ds \right) \\
&\leqslant \left(\frac{|\boldsymbol{p}|_{\mathbb{R}^d} \|\nabla\lambda\|_\infty}{\lambda_0^2} \right)^2 \left(\frac{\mu}{2\lambda_0} \right) \left(\frac{\mu t}{2\lambda_0} \right).
\end{aligned}
$$

The term $(I\vec{I}I_2)$ can be estimated in the same way as $(I\vec{I}_3)$:

$$
\mathbf{E}|(I\vec{I}I_2)|_{\mathbb{R}^d}^2 \leqslant \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2} \|\boldsymbol{b}\|_\infty \right)^2 \left(\frac{t}{\lambda_0} \right)^2 \left(\frac{\mu t}{2\lambda_0} \right).
$$

But, in general, one cannot estimate $\mathbf{E}|(I\vec{I}I_3)|^2$ up to a term which goes to 0 as $\mu \downarrow 0$. As an example, suppose that $\Lambda = \|\lambda\|_\infty$ and for $0 \leqslant t \leqslant T < \infty$ we have $\nabla\lambda(\boldsymbol{q}_t^\mu) = \mathbf{e}_1$. Here $\mathbf{e}_1$ is the unit basis vector $\mathbf{e}_1 = (1, 0, \ldots, 0)$ in $\mathbb{R}^d$. Let W_r^k be the kth $(1 \leqslant k \leqslant d)$ component of the

Wiener process $\boldsymbol{W}_r$. For $0 < t \leqslant T$ we have

$$\mathbf{E}|(I\vec{I}I_3)|_{\mathbb{R}^d} \geqslant \frac{1}{\mu\Lambda^2}\mathbf{E}\bigg|\int_0^t \bigg(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))} d\boldsymbol{W}_r\bigg)\bigg(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))} dW_r^1\bigg) ds\bigg|_{\mathbb{R}^d}$$

$$= \frac{1}{\mu\Lambda^2}\mathbf{E}\Bigg[\bigg(\int_0^t \bigg(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))} dW_r^1\bigg)^2 ds\bigg)^2$$

$$+ \sum_{k=2}^d \bigg(\int_0^t \bigg(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))} dW_r^k\bigg)\bigg(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))} dW_r^1\bigg) ds\bigg)^2\Bigg]^{\frac{1}{2}}$$

$$\geqslant \frac{1}{\mu\Lambda^2}\mathbf{E}\bigg(\int_0^t \bigg(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))} dW_r^1\bigg)^2 ds\bigg)$$

$$= \frac{1}{\mu\Lambda^2}\bigg(\int_0^t \bigg(\int_0^s \mathbf{E}e^{-\frac{2}{\mu}(A(\mu,s)-A(\mu,r))} dr\bigg) ds\bigg) \geqslant \frac{1}{\mu\Lambda^2}\bigg(\int_0^t \bigg(\int_0^s e^{-\frac{2}{\mu}\Lambda(s-r)} dr\bigg) ds\bigg)$$

$$= \frac{1}{\mu\Lambda^2}\frac{\mu}{2\Lambda}\int_0^t (1 - e^{-\frac{2\Lambda s}{\mu}}) ds = \frac{t}{2\Lambda^3} - \frac{\mu}{4\Lambda^4}(1 - e^{-\frac{2\Lambda t}{\mu}}),$$

which does not tend to 0 as $\mu \downarrow 0$. Since

$$\mathbf{E}|(I\vec{I}I_3)|^2_{\mathbb{R}^d} \geqslant (\mathbf{E}|(I\vec{I}I_3)|_{\mathbb{R}^d})^2,$$

we see that $\mathbf{E}|(I\vec{I}I_3)|^2_{\mathbb{R}^d}$ does not go to 0 as $\mu \downarrow 0$. Now, we have

$$\mathbf{E}\bigg|\boldsymbol{\gamma}(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s\bigg|^2_{\mathbb{R}^d} \geqslant \frac{1}{4}\mathbf{E}|(I\vec{I}I_3)|^2_{\mathbb{R}^d} - \mathbf{E}|\boldsymbol{R}_{\boldsymbol{\gamma}}(\mu)|^2_{\mathbb{R}^d} - \mathbf{E}|(I\vec{I}I_1)|^2_{\mathbb{R}^d} - \mathbf{E}|(I\vec{I}I_2)|^2_{\mathbb{R}^d}.$$

Therefore,

$$\mathbf{E}\bigg|\boldsymbol{\gamma}(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s\bigg|^2_{\mathbb{R}^d}$$

is uniformly bounded from below by a positive constant as $\mu \downarrow 0$.

We can check that the process $\boldsymbol{q}_t^\mu$, $0 \leqslant t \leqslant T$, does not converge as $\mu \downarrow 0$ to the process $\boldsymbol{q}_t$, $\boldsymbol{q}_0 = \boldsymbol{q}$. We have

$$\boldsymbol{q}_t^\mu = \boldsymbol{q} + \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)} ds + \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s$$

$$+ \boldsymbol{\alpha}(\mu) + \bigg(\boldsymbol{\beta}(\mu) - \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)} ds\bigg) + \bigg(\boldsymbol{\gamma}(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s\bigg),$$

$$\boldsymbol{q}_t = \boldsymbol{q} + \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s)}{\lambda(\boldsymbol{q}_s)} ds + \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s)} d\boldsymbol{W}_s.$$

Suppose that for any κ, $T > 0$, and any $\boldsymbol{p}_0^\mu = \boldsymbol{p} \in \mathbb{R}^d$ fixed

$$\lim_{\mu \downarrow 0} \mathbf{P}\Big(\max_{0 \leqslant t \leqslant T} |\boldsymbol{q}_t^\mu - \boldsymbol{q}_t|^2_{\mathbb{R}^d} \geqslant \kappa \Big) = 0.$$

For some $A > 0$ independent of μ and κ we have

$$\mathbf{E}\left|(\boldsymbol{q}_t^\mu - \boldsymbol{q}_t) - \int_0^t \left(\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)} - \frac{\boldsymbol{b}(\boldsymbol{q}_s)}{\lambda(\boldsymbol{q}_s)} \right) ds - \int_0^t \left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)} - \frac{1}{\lambda(\boldsymbol{q}_s)} \right) d\boldsymbol{W}_s \right|^2_{\mathbb{R}^d} \leqslant A \mathbf{E} \max_{0 \leqslant s \leqslant t} |\boldsymbol{q}_s^\mu - \boldsymbol{q}_s|^2_{\mathbb{R}^d}$$

$$\leqslant A\Big[\mathbf{P}\big(\max_{0 \leqslant s \leqslant t} |\boldsymbol{q}_s^\mu - \boldsymbol{q}_s|^2_{\mathbb{R}^d} \geqslant \kappa\big) \cdot \mathbf{E} \max_{0 \leqslant s \leqslant t} |\boldsymbol{q}_s^\mu - \boldsymbol{q}_s|^2_{\mathbb{R}^d} + \mathbf{P}\big(\max_{0 \leqslant s \leqslant t} |\boldsymbol{q}_s^\mu - \boldsymbol{q}_s|^2_{\mathbb{R}^d} < \kappa\big) \cdot \kappa \Big]$$

$$\leqslant A[\kappa + o(\mu, \kappa)]$$

since $\mathbf{E} \max\limits_{0 \leqslant s \leqslant t} |\boldsymbol{q}_s^\mu - \boldsymbol{q}_s|^2_{\mathbb{R}^d} < \infty$. Here, the term $o(\mu, \kappa)$ converges to 0 as $\mu \downarrow 0$ for every fixed $\kappa > 0$. Fix $\kappa > 0$. Letting $\mu \downarrow 0$, we see that

$$\lim_{\mu \downarrow 0} \mathbf{E}\left|(\boldsymbol{q}_t^\mu - \boldsymbol{q}_t) - \int_0^t \left(\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)} - \frac{\boldsymbol{b}(\boldsymbol{q}_s)}{\lambda(\boldsymbol{q}_s)} \right) ds - \int_0^t \left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)} - \frac{1}{\lambda(\boldsymbol{q}_s)} \right) d\boldsymbol{W}_s \right|^2_{\mathbb{R}^d} \leqslant A\kappa.$$

Since $\kappa > 0$ is arbitrary, we see that

$$\lim_{\mu \downarrow 0} \mathbf{E}\left|(\boldsymbol{q}_t^\mu - \boldsymbol{q}_t) - \int_0^t \left(\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)} - \frac{\boldsymbol{b}(\boldsymbol{q}_s)}{\lambda(\boldsymbol{q}_s)} \right) ds - \int_0^t \left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)} - \frac{1}{\lambda(\boldsymbol{q}_s)} \right) d\boldsymbol{W}_s \right|^2_{\mathbb{R}^d} = 0.$$

On the other hand, let us suppose that $\nabla \lambda(\boldsymbol{q}_t^\mu) = \mathbf{e}_1$ for $0 \leqslant t \leqslant T < \infty$. Here, $\mathbf{e}_1$ is the unit basis vector $\mathbf{e}_1 = (1, 0, \ldots, 0)$ in $\mathbb{R}^d$. We have

$$\mathbf{E}\left| \boldsymbol{\alpha}(\mu) + \left(\boldsymbol{\beta}(\mu) - \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)} ds \right) + \left(\boldsymbol{\gamma}(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s \right) \right|^2_{\mathbb{R}^d}$$

$$\geqslant \frac{1}{3} \mathbf{E}\left| \boldsymbol{\gamma}(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)} d\boldsymbol{W}_s \right|^2_{\mathbb{R}^d} - \mathbf{E}|\boldsymbol{\alpha}(\mu)|^2_{\mathbb{R}^d} - \mathbf{E}\left| \boldsymbol{\beta}(\mu) - \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)} ds \right|^2_{\mathbb{R}^d}.$$

By our estimates, this leads to a contradiction.

3 Regularization via Approximation of the Wiener Process

We can regularize the problem via *approximation of the Wiener process*. FOr this purpose, we introduce the process

$$\boldsymbol{W}_t^\delta = \frac{1}{\delta} \int_0^\infty \boldsymbol{W}_s \rho\left(\frac{s - t}{\delta} \right) ds = \frac{1}{\delta} \int_0^\delta \boldsymbol{W}_{s+t} \rho\left(\frac{s}{\delta} \right) ds,$$

where $\rho(\bullet)$ is a smooth C^∞ function whose support is contained in the interval $[0,1]$ such that

$$\int_0^1 \rho(s)ds = 1.$$

One can prove that (cf. [2] and the references therein)

$$\lim_{\delta\downarrow 0} \mathbf{E} \max_{t\in[0,T]} |\boldsymbol{W}_t^\delta - \boldsymbol{W}_t|^2_{\mathbb{R}^d} = 0.$$

We have

$$\dot{\boldsymbol{W}}_t^\delta = -\frac{1}{\delta}\int_0^1 \boldsymbol{W}_{t+\delta r}\dot{\rho}(r)dr.$$

We can then introduce the following regularization of our problem: first we consider the system

$$\mu\ddot{\boldsymbol{q}}_t^{\mu,\delta} = \boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta}) - \lambda(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{q}}_t^{\mu,\delta} + \dot{\boldsymbol{W}}_t^\delta, \quad \boldsymbol{q}_0^{\mu,\delta} = \boldsymbol{q} \in \mathbb{R}^d, \quad \dot{\boldsymbol{q}}_0^{\mu,\delta} = \boldsymbol{p} \in \mathbb{R}^d. \tag{3.1}$$

Equivalently, it is the first order system

$$\begin{cases} \dot{\boldsymbol{q}}_t^{\mu,\delta} = \boldsymbol{p}_t^{\mu,\delta}, \\ \dot{\boldsymbol{p}}_t^{\mu,\delta} = \dfrac{1}{\mu}\boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta}) - \dfrac{\lambda(\boldsymbol{q}_t^{\mu,\delta})}{\mu}\boldsymbol{p}_t^{\mu,\delta} + \dfrac{1}{\mu}\dot{\boldsymbol{W}}_t^\delta. \end{cases} \tag{3.2}$$

We can proceed with estimates similar to those in the previous sections. Since for fixed $\delta > 0$

$$|\dot{\boldsymbol{W}}_t^\delta|_{\mathbb{R}^d} \leqslant \frac{1}{\delta}\big(\max_{0\leqslant r\leqslant 1}|\dot{\rho}(r)|\big)\big(\max_{t\leqslant s\leqslant t+\delta}|\boldsymbol{W}_s|_{\mathbb{R}^d}\big) < \infty \quad \text{a.s.}, \tag{3.3}$$

we can prove that all the terms

$$\mathbf{E}|\boldsymbol{\alpha}(\mu)|_{\mathbb{R}^d}, \quad \mathbf{E}\bigg|\boldsymbol{\beta}(\mu) - \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds\bigg|_{\mathbb{R}^d}, \quad \mathbf{E}\bigg|\boldsymbol{\gamma}(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s^\delta\bigg|_{\mathbb{R}^d}$$

goes to zero as $\mu \downarrow 0$. (To be precise, we should write $\boldsymbol{\alpha}(\mu,\delta)$, $\boldsymbol{\beta}(\mu,\delta)$, and $\boldsymbol{\gamma}(\mu,\delta)$ to indicate the dependence on δ, but for the sake of brevity we neglect that.) In particular, with $\delta > 0$ fixed, we can estimate the term $(I\vec{I}I_3)$ up to a term which tends to 0 as $\mu \downarrow 0$. We have

$$\begin{aligned}
\mathbf{E}|(I\vec{I}I_3)|_{\mathbb{R}^d} &\leqslant \frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\int_0^t \mathbf{E}\bigg|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}\dot{\boldsymbol{W}}_r^\delta dr\bigg|^2_{\mathbb{R}^d} ds \\
&= \frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\int_0^t \frac{1}{\delta^2}\mathbf{E}\bigg|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}\bigg(\int_0^1 \boldsymbol{W}_{r+\delta m}\dot{\rho}(m)dm\bigg)dr\bigg|^2_{\mathbb{R}^d} ds \\
&= \frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\int_0^t \frac{1}{\delta^2}\mathbf{E}\bigg|\int_0^1 \dot{\rho}(m)\boldsymbol{W}_{r+\delta m}dm\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dr\bigg|^2_{\mathbb{R}^d} ds \\
&\leqslant \frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\int_0^t \frac{1}{\delta^2}\big(\max_{0\leqslant m\leqslant 1}|\dot{\rho}(m)|\big)^2\mathbf{E}\big(\max_{0\leqslant l\leqslant s+\delta}|\boldsymbol{W}_l|_{\mathbb{R}^d}\big)^2\bigg(\int_0^s e^{-\frac{\lambda_0(s-r)}{\mu}}dr\bigg)^2 ds
\end{aligned}$$

$$\leqslant \mu \frac{\|\nabla\lambda\|_\infty}{\lambda_0^4}\frac{t}{\delta^2}\big(\max_{0\leqslant m\leqslant 1}|\dot\rho(m)|\big)^2\mathbf{E}\big(\max_{0\leqslant l\leqslant s+\delta}|\boldsymbol{W}_l|_{\mathbb{R}^d}\big)^2.$$

Therefore, for fixed $\delta > 0$ we have $\mathbf{E}|(I\vec{I}I_3)|_{\mathbb{R}^d} \to 0$ as $\mu \downarrow 0$. By (2.4), we get

$$\begin{aligned}\boldsymbol{q}_t^{\mu,\delta} &= \boldsymbol{q} + \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^{\mu,\delta})}{\lambda(\boldsymbol{q}_s^{\mu,\delta})}ds + \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^{\mu,\delta})}d\boldsymbol{W}_s^\delta \\ &+ \boldsymbol{\alpha}(\mu) + \left(\boldsymbol{\beta}(\mu) - \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^{\mu,\delta})}{\lambda(\boldsymbol{q}_s^{\mu,\delta})}ds\right) + \left(\boldsymbol{\gamma}(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^{\mu,\delta})}d\boldsymbol{W}_s^\delta\right).\end{aligned} \tag{3.4}$$

Let the process $\widetilde{\boldsymbol{q}}_t^\delta$ be governed by the equation

$$\dot{\widetilde{\boldsymbol{q}}}_t^\delta = \frac{\boldsymbol{b}(\widetilde{\boldsymbol{q}}_t^\delta)}{\lambda(\widetilde{\boldsymbol{q}}_t^\delta)} + \frac{1}{\lambda(\widetilde{\boldsymbol{q}}_t^\delta)}\dot{\boldsymbol{W}}_t^\delta, \quad \widetilde{\boldsymbol{q}}_0^\delta = \boldsymbol{q} \in \mathbb{R}^d. \tag{3.5}$$

Then

$$\widetilde{\boldsymbol{q}}_t^\delta = \boldsymbol{q} + \int_0^t \frac{\boldsymbol{b}(\widetilde{\boldsymbol{q}}_s^\delta)}{\lambda(\widetilde{\boldsymbol{q}}_s^\delta)}ds + \int_0^t \frac{1}{\lambda(\widetilde{\boldsymbol{q}}_s^\delta)}d\boldsymbol{W}_s^\delta. \tag{3.6}$$

Let

$$M(t,\delta,\mu) = \mathbf{E}\max_{0\leqslant s\leqslant t}|\boldsymbol{q}_s^{\mu,\delta} - \widetilde{\boldsymbol{q}}_s^\delta|_{\mathbb{R}^d}.$$

By (3.4) and (3.6), using the estimate (3.3), we get

$$M(t,\delta,\mu) \leqslant K_1\int_0^t M(s,\delta,\mu)ds + K_2(t,\delta)\int_0^t M(s,\delta,\mu)ds + o_\mu(1).$$

Here, $o_\mu(1)$ is a term which goes to 0 as $\mu \downarrow 0$. The positive constant K_1 is independent of μ, δ, and t. The positive constant $K_2 = K_2(t,\delta)$ may depend on t and δ, but is independent of μ. Now, we use the Bellman–Gronwall inequality:

$$M(t,\delta,\mu) \leqslant o_\mu(1)\exp((K_1 + K_2(t,\delta))t).$$

We conclude that for any $\delta, \kappa, T > 0$ fixed and any $\boldsymbol{p}_0^{\mu,\delta} = \boldsymbol{p}$ fixed,

$$\lim_{\mu\downarrow 0}\mathbf{P}\big(\max_{0\leqslant t\leqslant T}|\boldsymbol{q}_t^{\mu,\delta} - \widetilde{\boldsymbol{q}}_t^\delta|_{\mathbb{R}^d} > \kappa\big) = 0.$$

Now we can take $\delta \downarrow 0$. Using Theorem 6.7.2 from [3] we obtain the following result.

Theorem 3.1. *We have, as $\delta \downarrow 0$,*

$$\lim_{\delta\to 0}\mathbf{E}\max_{t\in[0,T]}|\widetilde{\boldsymbol{q}}_t^\delta - \widehat{\boldsymbol{q}}_t|_{\mathbb{R}^d} = 0,$$

where $\widehat{\boldsymbol{q}}_t$ is the solution of the problem

$$\dot{\widehat{\boldsymbol{q}}}_t = \frac{\boldsymbol{b}(\widehat{\boldsymbol{q}}_t)}{\lambda(\widehat{\boldsymbol{q}}_t)} + \frac{1}{\lambda(\widehat{\boldsymbol{q}}_t)} \circ \dot{\boldsymbol{W}}_t, \quad \widehat{\boldsymbol{q}}_0 = \boldsymbol{q} \in \mathbb{R}^d. \tag{3.6}$$

Here, the stochastic term is understood in the Stratonovich sense.

In the general case,

$$\mu\ddot{\boldsymbol{q}}_t^{\mu,\delta} = \boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta}) - \lambda(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{q}}_t^{\mu,\delta} + \sigma(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{W}}_t^{\delta}, \quad \boldsymbol{q}_0^{\mu,\delta} = \boldsymbol{q}, \quad \dot{\boldsymbol{q}}_0^{\mu,\delta} = \boldsymbol{p}, \tag{3.7}$$

where the matrix $\sigma(\bullet)$ satisfies the assumptions made in Section 1, for any $\delta, \kappa, T > 0$ fixed and any $\boldsymbol{p}_0^{\mu,\delta} = \boldsymbol{p}$ fixed we have

$$\lim_{\mu\downarrow 0} \mathbf{P}\big(\max_{0\leqslant t\leqslant T} |\boldsymbol{q}_t^{\mu,\delta} - \widetilde{\boldsymbol{q}}_t^{\delta}|_{\mathbb{R}^d} > \kappa\big) = 0.$$

The process $\widetilde{\boldsymbol{q}}_t^{\delta}$ is governed by the equation

$$\dot{\widetilde{\boldsymbol{q}}}_t^{\delta} = \frac{\boldsymbol{b}(\widetilde{\boldsymbol{q}}_t^{\delta})}{\lambda(\widetilde{\boldsymbol{q}}_t^{\delta})} + \frac{\sigma(\widetilde{\boldsymbol{q}}_t^{\delta})}{\lambda(\widetilde{\boldsymbol{q}}_t^{\delta})}\dot{\boldsymbol{W}}_t^{\delta}, \quad \widetilde{\boldsymbol{q}}_0^{\delta} = \boldsymbol{q} \in \mathbb{R}^d. \tag{3.8}$$

We conclude the section with the following assertion.

Theorem 3.2. Under the above assumptions,

$$\lim_{\delta\to 0} \mathbf{E} \max_{t\in[0,T]} |\widetilde{\boldsymbol{q}}_t^{\delta} - \widehat{\boldsymbol{q}}_t|_{\mathbb{R}^d} = 0,$$

where $\widehat{\boldsymbol{q}}_t$ is the solution of the problem

$$\dot{\widehat{\boldsymbol{q}}}_t = \frac{\boldsymbol{b}(\widehat{\boldsymbol{q}}_t)}{\lambda(\widehat{\boldsymbol{q}}_t)} + \frac{\sigma(\widehat{\boldsymbol{q}}_t)}{\lambda(\widehat{\boldsymbol{q}}_t)} \circ \dot{\boldsymbol{W}}_t, \quad \widehat{\boldsymbol{q}}_0 = \boldsymbol{q} \in \mathbb{R}^d. \tag{3.9}$$

4 One-Dimensional Case

In the case of one spatial variable, the Smoluchowski–Kramers approximation leads to a one-dimensional diffusion process q_t which is defined by the following stochastic differential equation written in the Itô form:

$$\dot{q}_t = \frac{b(q_t)}{\lambda(q_t)} - \frac{\lambda'(q_t)}{2\lambda^3(q_t)} + \frac{1}{\lambda(q_t)}\dot{W}_t, \quad q_0 = q \in \mathbb{R}^1. \tag{4.1}$$

We set

$$\begin{aligned} u(q) &= \int_0^q \lambda(x)\exp\Big(-2\int_0^x b(y)\lambda(y)dy\Big)dx, \\ v(q) &= 2\int_0^q \lambda(x)\exp\Big(2\int_0^x b(y)\lambda(y)dy\Big)dx. \end{aligned} \tag{4.2}$$

Since $\lambda(x) > 0$, the functions $u(q)$ and $v(q)$ are strictly increasing. Following [4], we introduce an operator D_vD_u, where D_u means the differentiation with respect to the monotone function $u(q)$: $D_uf(q) = \lim\limits_{h\to 0} \dfrac{f(x+h)-f(x)}{u(x+h)-u(x)}$; the operator D_v is defined in a similar way. One can check that D_vD_u is the generator of the diffusion process q_t defined by (4.1).

Suppose that the friction coefficient $\lambda(q) = \lambda_\varepsilon(q)$ depends on a parameter $\varepsilon > 0$. We assume that for each $\varepsilon \in (0,1]$ the function $\lambda_\varepsilon(q)$ has a bounded continuous derivative $\lambda'_\varepsilon(q)$ and $0 < \underline{\lambda} \leqslant \lambda_\varepsilon(q) \leqslant \overline{\lambda} < \infty$. Let $u_\varepsilon(q)$ and $v_\varepsilon(q)$ be the functions defined by (4.2) with $\lambda(q)$ replaced by $\lambda_\varepsilon(q)$. Consider the stochastic process $q_t^{\mu,\delta,\varepsilon}$ in $\mathbb{R}^1$ defined by the equation

$$\mu \ddot{q}_t^{\mu,\delta,\varepsilon} = b(q_t^{\mu,\delta,\varepsilon}) - \lambda^\varepsilon(q_t^{\mu,\delta,\varepsilon})q_t^{\mu,\delta,\varepsilon} + \dot{W}_t^\delta, \quad q_0^{\mu,\delta,\varepsilon} = q, \quad \dot{q}_0^{\mu,\delta,\varepsilon} = p, \tag{4.3}$$

where $\dot{W}_t^\delta$ is, as above, a "smoothed" white noise converging to $\dot{W}_t$ as $\delta \downarrow 0$.

Theorem 4.1. *Assume that $\lambda_\varepsilon(q)$ converge weakly as $\varepsilon \downarrow 0$ on each finite interval $[\alpha,\beta] \subset \mathbb{R}^1$ to a function $\overline{\lambda}(q)$ (possibly, discontinuous). Then the processes $q_t^{\mu,\delta,\varepsilon}$ converge weakly on each finite time interval to the diffusion process $\overline{q}_t$ governed by the generator $D_{\overline{v}}D_{\overline{u}}$ (where $\overline{u}(q)$ and $\overline{v}(q)$ defined by (4.2) with $\lambda = \overline{\lambda}(q)$) as, first $\mu \downarrow 0$, then $\delta \downarrow 0$, and then $\varepsilon \downarrow 0$.*

Proof. According to Section 3, the processes $q_t^{\mu,\delta,\varepsilon}$ converge weakly as first $\mu \downarrow 0$ and then $\delta \downarrow 0$ to the process $\widehat{q}_t^\delta$ which solves Equation (4.1) with $\lambda(q) = \lambda^\varepsilon(q)$. By assumptions, the functions $u_\varepsilon(q)$ and $v_\varepsilon(q)$ converge as $\varepsilon \downarrow 0$ to functions $\overline{u}(q)$ and $\overline{v}(q)$ respectively for each $q \in \mathbb{R}^1$. The functions $\overline{u}(q)$ and $\overline{v}(q)$ are continuous and strictly increasing. Therefore (cf. [4]), there exists a diffusion process $\overline{q}_t$ governed by $D_{\overline{v}}D_{\overline{u}}$. As was shown in [5], the convergence of $u_\varepsilon(q)$ and $v_\varepsilon(q)$ as $\varepsilon \downarrow 0$ to $\overline{u}(q)$ and $\overline{v}(q)$ respectively implies the weak convergence of processes q_t^ε to the process corresponding to $D_{\overline{v}}D_{\overline{u}}$ as $\varepsilon \downarrow 0$. □

Theorem 4.2. *Let $\lambda_\varepsilon(q) = \widetilde{\lambda}(q/\varepsilon)$, and let one of the following conditions be satisfied:*

(1) *$\widetilde{\lambda}(q)$ is a continuously differentiable positive 1-periodic function,*
(2) *$\widetilde{\lambda}(q)$ is an ergodic stationary process (independent of the process W_t in (4.3)) with continuously differentiable trajectories and $0 < \lambda_- \leqslant \widetilde{\lambda}(q) \leqslant \lambda_+ < \infty$ for some constants λ_-, λ_+,*

Put $\overline{\lambda} = \int_0^1 \widetilde{\lambda}(q)dq$ if condition (1) is satisfied and $\overline{\lambda} = \mathbf{E}\widetilde{\lambda}(q)$ is condition (2) is satisfied.

Then the process $q_t^{\mu,\delta,\varepsilon}$ defined by (5.3) converges weakly as first $\mu \downarrow 0$ and then $\varepsilon \downarrow 0$ to the process $\overline{q}_t$ defined by the equation

$$\overline{q}_t = \frac{1}{\overline{\lambda}} b(\overline{q}_t) + \frac{1}{\overline{\lambda}} \dot{W}_t, \quad \overline{q}_0 = q.$$

The proof of this theorem follows from Theorem 4.1 since each of conditions (1) and (2) implies the validity of the assumptions of Theorem 4.1 and $\overline{\lambda}(q) = \overline{\lambda}$.

Assume that $\lambda_\varepsilon(q)$ is a bounded and separated from zero uniformly in $\varepsilon \in (0,1]$ positive function such that $\lim\limits_{\varepsilon\downarrow 0} \lambda_\varepsilon(q) = \lambda_1$ for $q < 0$ and $\lim\limits_{\varepsilon\downarrow 0} \lambda_\varepsilon(q) = \lambda_2$ for $q > 0$. Assume that $\lambda_\varepsilon(q)$ is continuously differentiable for each $\varepsilon > 0$. Let $\widehat{\lambda}(q)$ be the step function equal to λ_1 for $q \leqslant 0$ and λ_2 for $q > 0$. Let functions $\widehat{u}(q)$ and $\widehat{v}(q)$ be defined by formula (4.2) with $\lambda(q) = \widehat{\lambda}(q)$; $\widehat{u}(q)$ and $\widehat{v}(q)$ are continuous strictly increasing functions. Denote by $\widehat{q}_t$ the diffusion process in $\mathbb{R}^1$ governed by the generator $A = D_{\widehat{v}} D_{\widehat{u}}$. The process $\widehat{q}_t$ behaves as $\dfrac{1}{\lambda_1} W_t$ on the negative part of the q-axis and as $\dfrac{1}{\lambda_2} W_t$ on the positive part. Its behavior at $q = 0$ is defined by the domain of definition $\mathfrak{D}_A$ of the generator A: a continuous bounded function $f(q)$, $q \in \mathbb{R}^1$, twice continuously differentiable at $q \in \{\mathbb{R}^1 \setminus \{q = 0\}\}$ belongs to $\mathfrak{D}_A$ if and only if its left and right derivatives at $q = 0$, $f'_-(0)$ and $f'_+(0)$ respectively, satisfy the equality

$$\frac{1}{\lambda_1} f'_-(0) = \frac{1}{\lambda_2} f'_+(0)$$

and $Af(q)$ is continuous.

It is easy to see that the functions $u_\varepsilon(q)$ and $v_\varepsilon(q)$ defined by (4.2) with $\lambda(q) = \lambda_\varepsilon(q)$ converge as $\varepsilon \downarrow 0$ to $\widehat{u}(q)$ and $\widehat{v}(q)$ respectively for each $q \in \mathbb{R}^1$. This implies the following result.

Theorem 4.3. *Let the friction coefficient $\lambda_\varepsilon(q)$ satisfy the above conditions. Then the stochastic process $q_t^{\mu,\delta,\varepsilon}$ defined by (4.3) converges weakly to the diffusion process $\widehat{q}_t$ in $\mathbb{R}^1$ governed by $A = D_{\widehat{v}} D_{\widehat{u}}$ as first $\mu \downarrow 0$, then $\delta \downarrow 0$, and then $\varepsilon \downarrow 0$.*

This means, roughly speaking, that, if the friction coefficient is close to the step-function $\widehat{\lambda}(q)$, then the process q_t^μ, for $0 < \mu \ll 1$, can be approximated by the diffusion process $\widehat{q}_t$.

5 Multidimensional Case

In this section, we consider the problem of fast oscillating periodic environment in multidimensional case. We consider the system

$$\mu \ddot{\boldsymbol{q}}_t^{\mu,\delta,\varepsilon} = \boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta,\varepsilon}/\varepsilon) - \lambda(\boldsymbol{q}_t^{\mu,\delta,\varepsilon}/\varepsilon) \dot{\boldsymbol{q}}_t^{\mu,\delta,\varepsilon} + \dot{\boldsymbol{W}}_t^\delta, \quad \boldsymbol{q}_0^{\mu,\delta,\varepsilon} = \boldsymbol{q} \in \mathbb{R}^d, \quad \dot{\boldsymbol{q}}_0^{\mu,\delta,\varepsilon} = \boldsymbol{p} \in \mathbb{R}^d. \tag{5.1}$$

Here, as in Section 3, the process $\boldsymbol{W}_t^\delta$ is the approximation of the Wiener process in $\mathbb{R}^d$. We make the same assumptions on the functions $\lambda(\bullet)$ and $\boldsymbol{b}(\bullet)$ as in Section 2. In addition, we assume that the functions $\lambda(\bullet)$ and $\boldsymbol{b}(\bullet)$ are 1-periodic, i.e., $\lambda(\boldsymbol{x} + \mathbf{e}_k) = \lambda(\boldsymbol{x})$ and $\boldsymbol{b}(\boldsymbol{x} + \mathbf{e}_k) = \boldsymbol{b}(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{R}^d$ and $\mathbf{e}_k = (0, 0, \ldots, 1(k\text{th coordinate}), \ldots, 0)$, $1 \leqslant k \leqslant d$. Under this assumption, the system (5.1) can be regarded as a system on the d-torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Fix $\varepsilon > 0$. We can proceed as in Section 3 to see that first as $\mu \downarrow 0$, then as $\delta \downarrow 0$ the process $\boldsymbol{q}_t^{\mu,\delta,\varepsilon}$ converges in probability to the process $\boldsymbol{q}_t^\varepsilon$ such that

$$\dot{\boldsymbol{q}}_t^\varepsilon = \frac{\boldsymbol{b}(\boldsymbol{q}_t^\varepsilon/\varepsilon)}{\lambda(\boldsymbol{q}_t^\varepsilon/\varepsilon)} + \frac{1}{\lambda(\boldsymbol{q}_t^\varepsilon/\varepsilon)} \circ \dot{\boldsymbol{W}}_t, \quad \boldsymbol{q}_0^\varepsilon = \boldsymbol{q} \in \mathbb{R}^d.$$

The above equation, written in the form of the Itô integral, takes the form

$$\dot{\boldsymbol{q}}_t^\varepsilon = \frac{\boldsymbol{b}(\boldsymbol{q}_t^\varepsilon/\varepsilon)}{\lambda(\boldsymbol{q}_t^\varepsilon/\varepsilon)} - \frac{1}{2\varepsilon} \frac{\nabla\lambda(\boldsymbol{q}_t^\varepsilon/\varepsilon)}{\lambda^3(\boldsymbol{q}_t^\varepsilon/\varepsilon)} + \frac{1}{\lambda(\boldsymbol{q}_t^\varepsilon/\varepsilon)} \dot{\boldsymbol{W}}_t, \quad \boldsymbol{q}_0^\varepsilon = \boldsymbol{q} \in \mathbb{R}^d. \tag{5.2}$$

The generator corresponding to (5.2) is the second order differential operator

$$L^\varepsilon u(\boldsymbol{x}) = \left(\frac{\boldsymbol{b}(\boldsymbol{x}/\varepsilon)}{\lambda(\boldsymbol{x}/\varepsilon)} - \frac{1}{2\varepsilon} \frac{\nabla\lambda(\boldsymbol{x}/\varepsilon)}{\lambda^3(\boldsymbol{x}/\varepsilon)} \right) \cdot \nabla u(\boldsymbol{x}) + \frac{1}{2} \frac{1}{\lambda^2(\boldsymbol{x}/\varepsilon)} \Delta u(\boldsymbol{x}). \tag{5.3}$$

Our goal is to study the homogenization properties of (5.3) in the general multidimensional case. Homogenization problems were considered by many authors (cf., for example, [6]–[10].) However, we provide here an elementary probabilistic way of doing this. Our method follows [6] and [11, pp. 104–106].

We first make the change of variable $\dfrac{\boldsymbol{q}}{\varepsilon} = \boldsymbol{y}$ and $\dfrac{t}{\varepsilon^2} = s$. The process $\boldsymbol{y}_s^\varepsilon = \dfrac{1}{\varepsilon}\boldsymbol{q}_t^\varepsilon$ corresponds to the generator

$$A^\varepsilon = \frac{1}{2\lambda^2(\boldsymbol{y})}\Delta_{\boldsymbol{y}} - \frac{\nabla\lambda(\boldsymbol{y})}{2\lambda^3(\boldsymbol{y})}\cdot\nabla_{\boldsymbol{y}} + \varepsilon\frac{\boldsymbol{b}(\boldsymbol{y})}{\lambda(\boldsymbol{y})}\cdot\nabla_{\boldsymbol{y}}.$$

We regard $\boldsymbol{y}_s^\varepsilon$ as a process on $\mathbb{T}^d$. Then we have the bound

$$\Bigg|\mathbf{E}_{\boldsymbol{q}/\varepsilon} f(\boldsymbol{y}_s^\varepsilon) - \int\limits_{\mathbb{T}^d} f(\boldsymbol{x})\mu^\varepsilon(\boldsymbol{x})d\boldsymbol{x}\Bigg| < Ke^{-as}.$$

Here, $K > 0$ and $a > 0$ are independent of ε for small ε. The function f is bounded and measurable. The function $\mu^\varepsilon(\boldsymbol{x})$ is the density of the unique invariant measure of $\boldsymbol{y}_s^\varepsilon$ on $\mathbb{T}^d$ and

$$\int\limits_{\mathbb{T}^d} \mu^\varepsilon(\boldsymbol{x})d\boldsymbol{x} = 1.$$

We have

$$\lim_{\varepsilon\downarrow 0}\mu^\varepsilon(\boldsymbol{x}) = \mu(\boldsymbol{x}), \quad \lim_{\varepsilon\downarrow 0}\int\limits_{\mathbb{T}^d} f(\boldsymbol{x})\mu^\varepsilon(\boldsymbol{x})d\boldsymbol{x} = \int\limits_{\mathbb{T}^d} f(\boldsymbol{x})\mu(\boldsymbol{x})d\boldsymbol{x}$$

for $f \in C(\mathbb{T}^d)$, where $\mu(\boldsymbol{x})$ is a unique invariant measure for the process with generator A^0 on $\mathbb{T}^d$ and

$$\int\limits_{\mathbb{T}^d} \mu(\boldsymbol{x})d\boldsymbol{x} = 1.$$

Combining these estimates, we conclude that for any n and $t \geqslant \delta > 0$ there exists $\varepsilon_0(n,\delta) > 0$ such that for any $0 < \varepsilon < \varepsilon_0(n,\delta)$

$$\Bigg|\mathbf{E}_{\boldsymbol{q}} f\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \int\limits_{\mathbb{T}^d} f(\boldsymbol{x})\mu(\boldsymbol{x})d\boldsymbol{x}\Bigg| < \frac{1}{n}.$$

This implies that for any $f \in C(\mathbb{T}^d)$

$$\lim_{\varepsilon\downarrow 0}\sup_{t\geqslant\delta}\Bigg|\mathbf{E}_{\boldsymbol{q}} f\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \int\limits_{\mathbb{T}^d} f(\boldsymbol{x})\mu(\boldsymbol{x})d\boldsymbol{x}\Bigg| = 0.$$

Finally, we calculate the density $\mu(\boldsymbol{x})$. Since

$$A^0 = \frac{1}{2\lambda^2(\boldsymbol{y})}\Delta_{\boldsymbol{y}} - \frac{\nabla\lambda(\boldsymbol{y})}{2\lambda^3(\boldsymbol{y})}\cdot\nabla_{\boldsymbol{y}} = \frac{1}{2\lambda^2(\boldsymbol{y})}(\Delta_{\boldsymbol{y}} - \nabla(\ln\lambda(\boldsymbol{y}))\cdot\nabla_{\boldsymbol{y}}),$$

we see that

$$\mu(\boldsymbol{x}) = C\lambda(\boldsymbol{x}), \quad C = \Bigg(\int\limits_{\mathbb{T}^d}\lambda(\boldsymbol{x})d\boldsymbol{x}\Bigg)^{-1},$$

and we have the following result.

Lemma 5.1. *For any* $f \in C(\mathbb{T}^d)$

$$\lim_{\varepsilon\downarrow 0}\sup_{t\geqslant\delta}\left|\mathbf{E}_{\boldsymbol{q}} f\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \frac{\int_{\mathbb{T}^d} f(\boldsymbol{x})\lambda(\boldsymbol{x})d\boldsymbol{x}}{\int_{\mathbb{T}^d}\lambda(\boldsymbol{x})d\boldsymbol{x}}\right| = 0. \tag{5.4}$$

Corollary. *For any bounded continuous function* $f(\boldsymbol{x})$ *on* $\mathbb{T}^d$, $\boldsymbol{q}\in\mathbb{T}^d$,

$$\mathbf{E}_{\boldsymbol{q}}\left[\int_0^t f\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)ds - \frac{t\int_{\mathbb{T}^d} f(\boldsymbol{x})\lambda(\boldsymbol{x})d\boldsymbol{x}}{\int_{\mathbb{T}^d}\lambda(\boldsymbol{x})d\boldsymbol{x}}\right]^2 \to 0$$

as $\varepsilon\downarrow 0$, *for* $0<t<\infty$.

The proof of this corollary is the same as that of the corollary after Lemma 1 in [6].

Now, let us consider auxiliary functions $N_k(\boldsymbol{y})$, $k=1,\dots,d$, which are periodic bounded solutions (i.e., on $\mathbb{T}^d$) of the equation

$$\frac{1}{2\lambda^2(\boldsymbol{y})}\Delta_{\boldsymbol{y}}N_k(\boldsymbol{y}) - \frac{\nabla_{\boldsymbol{y}}\lambda(\boldsymbol{y})}{2\lambda^3(\boldsymbol{y})}\cdot\nabla_{\boldsymbol{y}}N_k(\boldsymbol{y}) = A^0(N_k(\boldsymbol{y})) = \frac{1}{2\lambda^3(\boldsymbol{y})}\frac{\partial\lambda}{\partial y_k}(\boldsymbol{y}), \quad \boldsymbol{y}\in\mathbb{T}^d. \tag{5.5}$$

The solvability of this equation comes from the fact that $(A^0)^*\lambda(\boldsymbol{y})=0$ and

$$\int_{\mathbb{T}^d}\frac{1}{2\lambda^3(\boldsymbol{y})}\frac{\partial\lambda}{\partial y_k}(\boldsymbol{y})\lambda(\boldsymbol{y})d\boldsymbol{y} = 0.$$

The boundedness of a solution comes from our assumptions on the function $\lambda(\bullet)$. Now, we apply the Itô formula:

$$\begin{aligned}\varepsilon N_k\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \varepsilon N_k\left(\frac{\boldsymbol{q}}{\varepsilon}\right) &= \varepsilon\Bigg[\int_0^t \nabla N_k\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)\cdot\frac{1}{\varepsilon}\left(\frac{\boldsymbol{b}}{\lambda}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) - \frac{1}{2\varepsilon}\frac{\nabla\lambda}{\lambda^3}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) + \frac{\dot{\boldsymbol{W}}_s}{\lambda(\boldsymbol{q}_s^\varepsilon/\varepsilon)}\right)ds \\ &\quad + \frac{1}{2}\int_0^t \Delta N_k\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)\frac{1}{\varepsilon^2}\frac{1}{\lambda^2(\boldsymbol{q}_s^\varepsilon/\varepsilon)}ds\Bigg] \\ &= \int_0^t \nabla N_k\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)\cdot\left(\frac{\boldsymbol{b}}{\lambda}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) + \frac{\dot{\boldsymbol{W}}_s}{\lambda(\boldsymbol{q}_s^\varepsilon/\varepsilon)}\right)ds + \frac{1}{2\varepsilon}\int_0^t \frac{\frac{\partial\lambda}{\partial y_k}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)}{\lambda^3(\boldsymbol{q}_s^\varepsilon/\varepsilon)}ds.\end{aligned} \tag{5.6}$$

Let $\boldsymbol{N}(\boldsymbol{y}) = (N_1(\boldsymbol{y}),\dots,N_d(\boldsymbol{y}))$. Using (5.5), we have

$$\begin{aligned}\boldsymbol{q}_t^\varepsilon - \boldsymbol{q} &= \int_0^t\left(\frac{\boldsymbol{b}}{\lambda(\boldsymbol{q}_s^\varepsilon/\varepsilon)} + \frac{\dot{\boldsymbol{W}}_s}{\lambda(\boldsymbol{q}_s^\varepsilon/\varepsilon)}\right)ds \\ &\quad + \int_0^t (D\boldsymbol{N})\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)\left(\frac{\boldsymbol{b}}{\lambda}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) + \frac{\dot{\boldsymbol{W}}_s}{\lambda(\boldsymbol{q}_s^\varepsilon/\varepsilon)}\right)ds - \varepsilon\left(\boldsymbol{N}\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \boldsymbol{N}\left(\frac{\boldsymbol{q}}{\varepsilon}\right)\right).\end{aligned}$$

Here,

$$(D\boldsymbol{N})(\boldsymbol{y}) = \left(\frac{\partial N_i}{\partial y_j}\right)_{1\leqslant i,j\leqslant d}, \quad \boldsymbol{y} = (y_1, \ldots, y_d) \in \mathbb{T}^d.$$

Therefore, using the corollary after Lemma 5.1, we see that $\boldsymbol{q}_t^\varepsilon$ converges weakly to a process $\boldsymbol{q}_t$, $\boldsymbol{q}_0 = \boldsymbol{q} \in \mathbb{R}^d$ governed by the operator

$$\overline{L} = \frac{1}{2}\sum_{i,j=1}^{d} \overline{a}_{ij}\frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^{d} \overline{b}_i \frac{\partial}{\partial y_i} \tag{5.7}$$

with the coefficients

$$\overline{a}_{ij} = \int\limits_{\mathbb{T}^d} \left(\frac{\nabla N_i(\boldsymbol{y})\cdot\nabla N_j(\boldsymbol{y})}{\lambda(\boldsymbol{y})} + \frac{1}{\lambda(\boldsymbol{y})}\left(\frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) + \frac{\partial N_i}{\partial y_j}(\boldsymbol{y})\right) + \delta_{ij}\frac{1}{\lambda(\boldsymbol{y})}\right) d\boldsymbol{y} \Bigg/ \left(\int\limits_{\mathbb{T}^d} \lambda(\boldsymbol{y})d\boldsymbol{y}\right),$$

$$\overline{b}_i = \frac{\displaystyle\int_{\mathbb{T}^d} b_i(\boldsymbol{y})d\boldsymbol{y}}{\displaystyle\int_{\mathbb{T}^d}\lambda(\boldsymbol{y})d\boldsymbol{y}} + \sum_{k=1}^{d}\frac{\displaystyle\int_{\mathbb{T}^d} b_k(\boldsymbol{y})\frac{\partial N_i}{\partial y_k}(\boldsymbol{y})d\boldsymbol{y}}{\displaystyle\int_{\mathbb{T}^d}\lambda(\boldsymbol{y})d\boldsymbol{y}}. \tag{5.8}$$

Here, $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise.

We can simplify the expression for $\overline{a}_{ij}$: using (5.5), we get

$$\begin{aligned}
\overline{a}_{ij} &= \int\limits_{\mathbb{T}^d} \left(\frac{\nabla N_i(\boldsymbol{y})\cdot\nabla N_j(\boldsymbol{y})}{\lambda(\boldsymbol{y})} + \frac{1}{\lambda(\boldsymbol{y})}\left(\frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) + \frac{\partial N_i}{\partial y_j}(\boldsymbol{y})\right) + \delta_{ij}\frac{1}{\lambda(\boldsymbol{y})}\right) d\boldsymbol{y} \Bigg/ \left(\int\limits_{\mathbb{T}^d} \lambda(\boldsymbol{y})d\boldsymbol{y}\right) \\
&= \int\limits_{\mathbb{T}^d} \Bigg(\operatorname{div}\left(\frac{N_i(\boldsymbol{y})}{\lambda(\boldsymbol{y})}\nabla N_j(\boldsymbol{y})\right) - \frac{N_i(\boldsymbol{y})}{\lambda(\boldsymbol{y})}\Delta N_j(\boldsymbol{y}) - N_i(\boldsymbol{y})\nabla N_j(\boldsymbol{y})\cdot\nabla\left(\frac{1}{\lambda(\boldsymbol{y})}\right) \\
&\quad + \frac{1}{\lambda(\boldsymbol{y})}\left(\frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) + \frac{\partial N_i}{\partial y_j}(\boldsymbol{y})\right) + \delta_{ij}\frac{1}{\lambda(\boldsymbol{y})}\Bigg) d\boldsymbol{y} \Bigg/ \left(\int\limits_{\mathbb{T}^d} \lambda(\boldsymbol{y})d\boldsymbol{y}\right) \\
&= \int\limits_{\mathbb{T}^d} \Bigg(\frac{\partial}{\partial y_j}\left(N_i(\boldsymbol{y})\frac{1}{\lambda(\boldsymbol{y})}\right) - \frac{1}{\lambda(\boldsymbol{y})}\frac{\partial N_i}{\partial y_j}(\boldsymbol{y}) \\
&\quad + \frac{1}{\lambda(\boldsymbol{y})}\left(\frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) + \frac{\partial N_i}{\partial y_j}(\boldsymbol{y})\right) + \delta_{ij}\frac{1}{\lambda(\boldsymbol{y})}\Bigg) d\boldsymbol{y} \Bigg/ \left(\int\limits_{\mathbb{T}^d} \lambda(\boldsymbol{y})d\boldsymbol{y}\right) \\
&= \frac{\displaystyle\int_{\mathbb{T}^d}\frac{\partial N_j}{\partial y_i}(\boldsymbol{y})\frac{1}{\lambda(\boldsymbol{y})}d\boldsymbol{y}}{\displaystyle\int_{\mathbb{T}^d}\lambda(\boldsymbol{y})d\boldsymbol{y}} + \delta_{ij}\frac{\displaystyle\int_{\mathbb{T}^d}\frac{1}{\lambda(\boldsymbol{y})}d\boldsymbol{y}}{\displaystyle\int_{\mathbb{T}^d}\lambda(\boldsymbol{y})d\boldsymbol{y}}.
\end{aligned} \tag{5.9}$$

So, we have the following assertion.

Theorem 5.1. *As $\varepsilon \downarrow 0$, the process $\boldsymbol{q}_t^\varepsilon$ converges weakly to the process $\boldsymbol{q}_t$, $\boldsymbol{q}_0 = \boldsymbol{q} \in \mathbb{R}^d$, governed by the operator* (5.7) *with coefficients given by* (5.8) *and* (5.9).

This Theorem implies a homogenization result for the process $\boldsymbol{q}_t^{\mu,\delta,\varepsilon}$ defined by (5.1).

6 Remarks and Generalizations

6.1 Small mass – large friction asymptotics

Let the friction coefficient in (1.1) be as follows: $\lambda^\varepsilon(\boldsymbol{q}) = \varepsilon^{-1}\lambda(\boldsymbol{q})$, $\boldsymbol{q} \in \mathbb{R}^n$, $0 < \varepsilon \ll 1$. As it follows from Theorem 3.1, the Smoluchowski–Kramers approximation in this case has the form

$$\dot{\overline{\boldsymbol{q}}}_t^\varepsilon = \frac{\varepsilon \boldsymbol{b}(\overline{\boldsymbol{q}}_t^\varepsilon)}{\lambda(\overline{\boldsymbol{q}}_t^\varepsilon)} - \frac{\varepsilon^2 \nabla\lambda(\overline{\boldsymbol{q}}_t^\varepsilon)}{2\lambda^3(\overline{\boldsymbol{q}}_t^\varepsilon)} + \frac{\varepsilon}{\lambda(\overline{\boldsymbol{q}}_t^\varepsilon)}\dot{\boldsymbol{W}}_t, \quad \overline{\boldsymbol{q}}_0^\varepsilon = \boldsymbol{q}.$$

Put $\widetilde{\boldsymbol{q}}_t^\varepsilon = \overline{\boldsymbol{q}}_{t/\varepsilon}^\varepsilon$. Then $\widetilde{\boldsymbol{q}}_t^\varepsilon$ satisfies the equation

$$\dot{\widetilde{\boldsymbol{q}}}_t^\varepsilon = \frac{\boldsymbol{b}(\widetilde{\boldsymbol{q}}_t^\varepsilon)}{\lambda(\widetilde{\boldsymbol{q}}_t^\varepsilon)} - \frac{\varepsilon \nabla\lambda(\widetilde{\boldsymbol{q}}_t^\varepsilon)}{2\lambda^3(\widetilde{\boldsymbol{q}}_t^\varepsilon)} + \frac{\sqrt{\varepsilon}}{\lambda(\widetilde{\boldsymbol{q}}_t^\varepsilon)}\dot{\widetilde{\boldsymbol{W}}}_t, \quad \widetilde{\boldsymbol{q}}_0^\varepsilon = \boldsymbol{q}, \tag{6.1}$$

where $\widetilde{\boldsymbol{W}}_t$ is a Wiener process.

Assume that the vector field $\boldsymbol{b}(\boldsymbol{q})$, $\boldsymbol{q} \in \mathbb{R}^n$, has a finite number of compact attractors $K_1, \dots, K_l$. Let, for the sake of brevity, each K_i be an asymptotically stable equilibrium, and let each point of $\mathbb{R}^n$, except for a separatrix set $\mathscr{E} \subset \mathbb{R}^n$, be attracted to one of these equilibriums. The separatrix set $\mathscr{E}$ is assumed to have dimension less than n. Then, if $\widetilde{\boldsymbol{q}}_0^\varepsilon = \boldsymbol{q} \notin \mathscr{E}$, $\widetilde{\boldsymbol{q}}_t^\varepsilon$ first comes to a small neighborhood of a stable equilibrium K_i, $i = i(\boldsymbol{q})$, with the probability close to 1 as $\varepsilon \downarrow 0$ and spends in this neighborhood a long time. Because of the large deviations, the trajectory will switch to the neighborhood of another attractor, then to another, and so on. We see from (6.1) that the long-time behavior of the system with a large friction is similar to the behavior of a system with small noise. Applying the results of [12, Chapters 4 and 6], we see that for $0 < \varepsilon \ll 1$ the sequence of transitions between the attractors, the main term of transition time logarithmic asymptotics, and the most probable transition paths are not random for generic systems. These characteristics of the long-time behavior are defined by the function

$$V(\boldsymbol{x}, \boldsymbol{y}) = \inf\left\{\frac{1}{2}\int_0^T |\lambda(\boldsymbol{\varphi}_s)\dot{\boldsymbol{\varphi}}_s - \boldsymbol{b}(\boldsymbol{\varphi}_s)|^2 ds : \boldsymbol{\varphi}_0 = \boldsymbol{x}, \quad \boldsymbol{\varphi}_T = \boldsymbol{y}, T \geqslant 0\right\}, \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$$

and the extremals of this variational problem.

Assume that the dynamical system $\dot{\boldsymbol{q}}_t = \boldsymbol{b}(\boldsymbol{q}_t)$ has a first integral. Assume that, say, $n = 2$ and $\boldsymbol{b}(\boldsymbol{q}) = \overline{\nabla} H(\boldsymbol{q})$ for some smooth generic function $H(\boldsymbol{q})$, $\boldsymbol{q} \in \mathbb{R}^2$, such that $\lim\limits_{|\boldsymbol{q}| \to \infty} H(\boldsymbol{q}) = \infty$. In this case, $H(\boldsymbol{q}_t) \equiv H(\boldsymbol{q}_0)$.

Assume that the friction is strong: $\lambda^\varepsilon(\boldsymbol{q}) = \varepsilon^{-1}\lambda(\boldsymbol{q})$. Making the time change $\widehat{\boldsymbol{q}}_t^\varepsilon = \boldsymbol{q}_{t/\varepsilon^2}$, we have

$$\dot{\widehat{\boldsymbol{q}}}_t^\varepsilon = \frac{1}{\varepsilon^2}\nabla H(\widehat{\boldsymbol{q}}_t^\varepsilon) - \frac{\nabla\lambda(\widehat{\boldsymbol{q}}_t^\varepsilon)}{2\lambda^3(\widehat{\boldsymbol{q}}_t^\varepsilon)} + \frac{1}{\lambda(\widehat{\boldsymbol{q}}_t^\varepsilon)}\dot{\boldsymbol{W}}_t, \quad \widehat{\boldsymbol{q}}_0^\varepsilon = \boldsymbol{q} \in \mathbb{R}^2.$$

We identify points of each connected component of every level set of $H(\boldsymbol{q})$. The set obtained after such an identification is homeomorphic in the natural topology to the graph Γ. Let $Y : \mathbb{R}^2 \to \Gamma$ be the identification mapping. Then the long-time evolution of the system can be characterized by the stochastic process $\mathscr{Y}_t^\varepsilon = Y(\widehat{\boldsymbol{q}}_t^\varepsilon)$ on Γ. The process $\mathscr{Y}_t^\varepsilon$, in general, is not Markovian. But

$\mathscr{Y}_t^\varepsilon$ converges weakly in the space of continuous functions $\varphi : [0,T] \to \Gamma$ as $\varepsilon \downarrow 0$ to a diffusion process on the graph Γ (cf. [12, Chapter 8]). This limiting process is defined by a family of second order differential operators, one on each edge of Γ, and by the gluing conditions at the vertices. Following [12], one can evaluate these operators and the gluing conditions.

6.2 Fast oscillating random friction in multidimensional case

Let us consider the case of fast oscillating in the space variable, random friction, in dimension $d \geqslant 2$. Let $(\Omega, \mathscr{F}, \mathbf{P})$ be a probability space. Let $\lambda(\boldsymbol{x}, \omega)$, $\omega \in \Omega$ be a random field in $\mathbb{R}^d$ with the following properties:

(i) For any fixed $\omega \in \Omega$ and $\boldsymbol{x} \in \mathbb{R}^d$, the function $\infty > \Lambda \geqslant \lambda(\boldsymbol{x}, \omega) \geqslant \lambda_0 > 0$.

(ii) For every $\boldsymbol{x} \in \mathbb{R}^d$ the random variable $\lambda(\boldsymbol{x}, \omega)$ is independent of the Wiener process $\boldsymbol{W}_t$.

(iii) The random field $\lambda(\boldsymbol{x}, \omega)$ has the form $\lambda(\boldsymbol{x}, \omega) = \lambda(T(\boldsymbol{x})\omega)$ where $T(\boldsymbol{x}) : \Omega \to \Omega$ is a d-dimensional dynamical system which preserves the measure $\mathbf{P}$ and is ergodic with respect to $\mathbf{P}$.

Let us now consider an analogue of (5.1):

$$\mu \ddot{\boldsymbol{q}}_t^{\mu,\delta,\varepsilon} = -\lambda\left(\frac{\boldsymbol{q}_t^{\mu,\delta,\varepsilon}}{\varepsilon}, \omega\right) \dot{\boldsymbol{q}}_t^{\mu,\delta,\varepsilon} + \dot{\boldsymbol{W}}_t^\delta \ , \ \boldsymbol{q}_0^{\mu,\delta,\varepsilon} = \boldsymbol{q} \in \mathbb{R}^d \ , \ \dot{\boldsymbol{q}}_0^{\mu,\varepsilon,\delta} = \boldsymbol{p} \in \mathbb{R}^d \ . \tag{6.2}$$

For each fixed $\omega \in \Omega$, as we proved in Section 3, we have that $\boldsymbol{q}_t^{\mu,\delta,\varepsilon}(\omega)$ converges weakly to a process $\boldsymbol{q}_t^\varepsilon(\omega)$ as first $\mu \downarrow 0$ and then $\delta \downarrow 0$. The process $\boldsymbol{q}_t^\varepsilon$ is subject to

$$\dot{\boldsymbol{q}}_t^\varepsilon = -\frac{1}{2\varepsilon} \frac{\nabla\lambda\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}, \omega\right)}{\lambda^3\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}, \omega\right)} + \frac{1}{\lambda\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}, \omega\right)} \dot{\boldsymbol{W}}_t \ , \ \boldsymbol{q}_0^\varepsilon = \boldsymbol{q} \in \mathbb{R}^d \ . \tag{6.3}$$

We conjecture that as $\varepsilon \downarrow 0$, $\boldsymbol{q}_t^\varepsilon$ converges weakly to a process $\boldsymbol{q}_t$, $\boldsymbol{q}_0 = \boldsymbol{q} \in \mathbb{R}^d$ subject to the operator $\overline{\overline{L}} = \frac{1}{2}\sum\limits_{i,j=1}^d \overline{\overline{a}}_{ij} \frac{\partial^2}{\partial x_i \partial x_j}$ with effective diffusivity

$$\overline{\overline{a}}_{ij} = \mathbf{E}\left[\frac{\displaystyle\int_{\mathbb{T}^d} \frac{\partial N_j}{\partial y_i}(\boldsymbol{x}, \omega)\frac{1}{\lambda(\boldsymbol{x}, \omega)} d\boldsymbol{x}}{\displaystyle\int_{\mathbb{T}^d} \lambda(\boldsymbol{x}, \omega) d\boldsymbol{x}} + \delta_{ij} \frac{\displaystyle\int_{\mathbb{T}^d} \frac{1}{\lambda(\boldsymbol{x}, \omega)} d\boldsymbol{x}}{\displaystyle\int_{\mathbb{T}^d} \lambda(\boldsymbol{x}, \omega) d\boldsymbol{x}}\right] \ . \tag{6.4}$$

Here the functions $N_k(\boldsymbol{x}, \omega)$ $(1 \leqslant k \leqslant d)$ shall satisfy certain auxiliary problem. (We actually have a formulation of this problem but we are not sure about its validity: we let $N_k(\boldsymbol{x}, \omega)$ be the solution of the equation

$$\mathbf{E}[(\nabla_{\boldsymbol{x}} N_k(\boldsymbol{x}, \omega) - \boldsymbol{e}_k) \cdot \nabla_{\boldsymbol{x}} \varphi(\boldsymbol{x}, \omega)] = 0 \ ,$$

for all $\varphi(\boldsymbol{x}, \omega)$ smooth and compactly supported in $\boldsymbol{x} \in \mathbb{R}^d$ and measurable with respect to $\omega \in \Omega$. The existence of solutions to this problem is guaranteed by the Lax-Milgram lemma.)

However, we are not aware of the validity of this conjecture nor a proof of it. We are also not sure about the correct reference of such a problem. (We thank E.Kosygina for pointing out to us two relevant papers [13] and [14].)

6.3 Motion of charged particles in a magnetic field

One can expect that, using the regularization by smoothed white noise, it is possible to get the Smoluchowski–Kramers approximation for the equation

$$\mu \ddot{\boldsymbol{q}}_t^\mu = \boldsymbol{b}(\boldsymbol{q}_t^\mu) - A(\boldsymbol{q}_t^\mu)\dot{\boldsymbol{q}}_t^\mu + \sigma \dot{\boldsymbol{W}}_t, \quad \boldsymbol{q}_0^\mu = \boldsymbol{q}, \dot{\boldsymbol{q}}_0^\mu = \boldsymbol{p}, \quad \boldsymbol{q}, \boldsymbol{p} \in \mathbb{R}^n, \tag{6.5}$$

Here, $\sigma > 0$ and $A(\boldsymbol{q})$ are matrix-valued functions having strictly positive eigenvalues for each $\boldsymbol{q} \in \mathbb{R}^n$. In particular, if $A(\boldsymbol{q}) = \lambda(\boldsymbol{q})A$, where $\lambda(\boldsymbol{q}) > 0$ and A is a constant positive definite symmetric matrix, the problem can be reduced to the case considered in Section 3 by an appropriate linear change of variables.

If A has a negative eigenvalue, the Smoluchowski–Kramers approximation is not applicable. The case of A with pure imaginary eigenvalues is of interest since such equations describe the motion of charged particles in a magnetic field. In the case $\lambda = \text{const}$, $n = 2$, and $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the problem was considered in [2]. In this case, the Smoluchowski–Kramers approximation holds after a regularization. If $\boldsymbol{b}(\boldsymbol{q}) = -\nabla F(\boldsymbol{q})$, $\boldsymbol{q} \in \mathbb{R}^2$, and $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, one can show that the regularization by the smoothed white noise leads to the equation

$$\dot{\boldsymbol{q}}_t = \frac{1}{\lambda(\boldsymbol{q}_t)} \overline{\nabla} F(\boldsymbol{q}_t) - \sigma^2 \frac{\overline{\nabla}\lambda(\boldsymbol{q}_t)}{2\lambda^3(\boldsymbol{q}_t)} + \frac{\sigma}{\lambda(\boldsymbol{q}_t)} \dot{\widetilde{\boldsymbol{W}}}_t \quad \boldsymbol{q}_0 = \boldsymbol{q} \in \mathbb{R}^2. \tag{6.6}$$

If the noise in (6.5) is small ($0 < \sigma \ll 1$), the motion described by (6.6) has a fast component and a slow component. Applying the results of [12, 15], one can describe the limiting (as $\sigma \downarrow 0$) slow component of $\widehat{\boldsymbol{q}}_t^\sigma = \boldsymbol{q}_{t/\sigma^2}$ as a diffusion process on the graph corresponding to the potential $F(\boldsymbol{q})$ (the graph is homeomorphic to the set of connected components of the level sets of $F(\boldsymbol{q})$ provided with the natural topology).

Acknowledgments

This work is supported in part by NSF Grants DMS-0803287 and DMS-0854982.

References

1. M. Freidlin, "Some remarks on the smoluchowski–kramers approximation," *J. Stat. Phys.* **117**, No.3-4, 617-634, (2004).

2. S. Cerrai and M. Freidlin, "Small mass asymptotics of a charged particle in a magnetic field and long-time influence of small perturbations," *J. Stat. Phys.* **144**, 101–123 (2011).

3. N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam (1989).

4. W. Feller, "Generalized second order differential operators and their lateral conditions," *Ill. J. Math.* **1**, 459–504 (1957).

5. M. Freidlin and A. Wentzell, "Necessary and sufficient conditions for weak convergence of one-dimensional Markov processes," In: *The Dynkin Festschrift: Markov Processes and their Applications In celebration of Eugene B. Dynkin's 70th birthday*, pp. 95–109, Birkhäuser (1994).

6. M. Freidlin, “Dirichlet’s problem for an equation with periodic coefficients depending on a small parameter” [in Russian], *Teor. Veroyatn. Primen.* **9**, 133-139 (1964).

7. V. V. Zhikov, S. M. Kozlov, O. A. Oleinik, and Ha Tien Ngoan, “Averaging and G-convergence of differential operators,” *Russ. Math. Surv.* **34**, No. 5, 65-147 (1979).

8. V. V. Zhikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer, Berlin (1994).

9. G. Papanicolaou, and S. R. S. Varadhan, “Boundary value problems with rapidly oscillating random coefficients,” In: *Proceedings of Conference on Random Fields, Esztergom, Hungary, 1979*, pp. 835–873, North-Holland, Amsterdam (1981).

10. N. Krylov, “G-convergence of elliptic operators in nondivergence form” [in Russian], *Mat. Zametki* **37**, No. 4, 522–527 (1985); English transl.: *Math Notes* **37**, 290–292 (1985).

11. M. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton Univ. Press, Princeton (1985).

12. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer, New York etc. (1998).

13. G. Allaire and R. Orive, “Homogenization of periodic non self-adjoint problems with large drift and potential,” *ESAIM, Control Optim. Calc. Var.* **13**, No. 4, 735–749 (2007).

14. M. Klepsyna and A. Piatnitski, *Averaging of Nonself Adjoint Parabolic Equations with Random Evolution*, INRIA, Report 3951 (2000).

15. M. Brin and M. Freidlin, “On stochastic behavior of perturbed Hamiltonian systems,” *Ergodic Theory Dyn. Syst.* **20**, No. 1. 55–76 (2000).

Submitted on July 6, 2011