Smoluchowski–Kramers Approximation

Minh Hoang

November 26, 2025

1 Introduction

Newton's second law of motion is one of the foundational principles in classical mechanics. It states that the acceleration of an object is directly proportional to the net external force acting upon it and inversely proportional to its mass. Mathematically, this relationship is expressed as:

$$\mathbf{F}_{\text{net}} = m\mathbf{a} = m\frac{d^2\mathbf{s}(t)}{dt^2} \tag{1}$$

where $\mathbf{F}_{\text{net}}$ is the total net force acting on the object, m is its mass, $\mathbf{a}$ is the resulting acceleration, and $\mathbf{s}(t)$ describes the position of the object as a function of time. In this work, we explore how the classical second-order dynamics described by Newton's law reduce to a first-order stochastic differential equation as the object's mass goes to zero. This limiting procedure, known as the *Smoluchowski-Kramers approximation.*

We consider the motion of a particle of mass μ in a deterministic force field $b(q)$, where $q \in \mathbb{R}^n$, and subject to both random fluctuations modeled by Gaussian white noise and a friction force proportional to the velocity. Denote the position of the particle q_t^μ. According to Newton's second law, the dynamics of the particle are governed by the second-order stochastic differential equation:

$$\mu \ddot{q}_t^\mu = b(q_t^\mu) + \sigma(q_t^\mu)\dot{W}_t - \alpha \dot{q}_t^\mu, \quad q_0^\mu = q \in \mathbb{R}^n, \quad \dot{q}_0^\mu = p \in \mathbb{R}^n \tag{2}$$

In Eq. (2), $b(q)$ is the deterministic vector field, $\dot{W}_t$ is a Gaussian White Noise process in $\mathbb{R}^n$, and $\sigma(q)$ is an $n \times n$ matrix. The term $\alpha \dot{q}_t^\mu$ describes the friction to the motion that is proportional to the velocity. The functions $b(q)$ and $\sigma(q)$ are supposed to be regular satisfying Assumption 1, so that the solution of Eq. (2) exists and is unique.

It is usually assumed in Eq. (2) the friction coefficient α is a positive constant. Under this assumption, one can prove that q_t^μ converges in probability as $\mu \downarrow 0$ uniformly on each finite time interval $[0, T]$ to an n-dimensional Itô diffusion process q_t.

$$\dot{q}_t = \frac{1}{\alpha}b(q_t) + \frac{1}{\alpha}\sigma(q_t)\dot{W}_t, \quad q_0 = q_0^\mu = q \in \mathbb{R}^n \tag{3}$$

In particular, consider the second-order Newton equation (2) and its first-order limiting counterpart (3). In Freidlin (2004), it was established that on each finite time interval $[0, T]$, for any $\epsilon > 0$, and any initial conditions $p, q \in \mathbb{R}^n$, one has:

$$\lim_{\mu \downarrow 0} \mathbb{P}\left(\max_{0 \leq t \leq T} |q_t^\mu - q_t| > \epsilon \right) = 0 \tag{4}$$

The approximation of q_t^μ by q_t for $0 < \mu \ll 1$ in Eq. (4) is called the *Smoluchowski–Kramers* approximation which justifies replacing the second-order (acceleration) equation with a first-order (velocity) equation in physical experiments.

The key idea behind this approximation lies in the physical interpretation of the inertial term $\mu \ddot{q}_t^\mu$, which represents the effect of mass times acceleration. When the mass μ is very small, this term becomes negligible compared to the dominant forces—friction and noise. Physically, when mass is negligible, friction instantly adjusts the velocity to match the applied forces and noise, thus acceleration no longer contributes significantly to the dynamics. This eliminates the need to track acceleration separately, allowing one to describe the motion entirely in terms of position and velocity.

2 Result

We begin with the following conjecture to develop intuition about the relationship between the second-order differential equation derived from Newton's second law and their corresponding first-order stochastic limit. Without loss of generality, we assume the constant friction coefficient $\alpha = 1$.

Conjecture 2.1. *Consider the second-order differential equation arising from Newton's second law:*

$$\mu \ddot{q}_t^\mu = b(t, q_t^\mu) + \sigma(t, q_t^\mu)\dot{W}_t - \dot{q}_t^\mu, \quad q_0^\mu = q, \quad \dot{q}_0^\mu = p, \tag{5}$$

and the associated Itô stochastic differential equation:

$$\dot{q}_t = b(t, q_t) + \sigma(t, q_t)\dot{W}_t, \quad q_0 = q. \tag{6}$$

Then, both equations can be respectively expressed that:

$$q_t^\mu = q + \int_0^t b(s, q_s^\mu)\, ds + \int_0^t \sigma(s, q_s^\mu)\, dW_s + \underbrace{\cdots\cdots\cdots\cdots\cdots\cdots\cdots}_{\leq C\mu},$$

$$q_t = q + \int_0^t b(s, q_s)\, ds + \int_0^t \sigma(s, q_s)\, dW_s.$$

It is conjectured that the residual terms, $|q_t^\mu - q_t|$, are bounded in magnitude by $C\mu$ for some constant $C > 0$. Therefore, as the mass $\mu \to 0$, the difference between the trajectories q_t^μ and q_t becomes arbitrarily small.

To rigorously analyze this convergence, we make the following assumptions on the drift and diffusion coefficients.

Assumption 1. *It is assumed that both $b(t, q)$ and $\sigma(t, q)$ Lipschitz continuous in q with constant K, and bounded uniformly in t and q.*

- *Lipschitz continuity in q with constant $K < \infty$:*

  $$|b(t, q_1) - b(t, q_2)| \leq K|q_1 - q_2| \quad \text{and} \quad |\sigma(t, q_1) - \sigma(t, q_2)| \leq K|q_1 - q_2|$$

  for all $q_1, q_2 \in \mathbb{R}^n$, where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$ for b and the operator norm on $\mathbb{R}^{n\times n}$ for σ.

- *Bounded uniformly in t and q:*

  $$|b(t, q)| \leq M \quad \text{and} \quad |\sigma(t, q)| \leq M$$

  for all $q \in \mathbb{R}^n$, where $M < \infty$ is a constant.

These regularity conditions ensure the existence and uniqueness of strong solutions to both the second-order system (5) and the limiting first-order Itô SDE (6). Moreover, they provide stability and control over the nonlinearities in the dynamics, which are essential for proving convergence results as $\mu \to 0$.

Lemma 2.2 (Freidlin (2004)). *Let q_t^μ and q_t be defined by Eqs. (5) and (6) respectively.*

(i) Assuming $\sigma(t,q)=0$ in both equations, then for any $T>0$

$$\max_{0\le t\le T}|q_t^\mu - q_t| \le e^{KT}\mu\left[\sup_{0\le t\le T, q\in\mathbb{R}^n}|b(t,q)|+|p|\right]$$

(ii) There exists $c_1>0$ defined by T,n and the Lipschitz constant K of the coefficients $b(t,q)$ and $\sigma(t,q)$, such that

$$\max_{0\le t\le T}\mathbb{E}\left[|q_t^\mu - q_t|^2\right] \le c_1\mu\left[|p|^2+\|b\|^2+\|\sigma\sigma^\top\|\right]$$

$\|b\| = \sup_{0\le t\le\infty, q\in\mathbb{R}^n, 1\le i\le n}|b_i(t,q)|$ *and* $\|\sigma\sigma^\top\| = \sup_{0\le t\le\infty, q\in\mathbb{R}^n, 1\le i,j\le n}\sum_{k=1}^n|\sigma_{ik}(t,q)\sigma_{kj}(t,q)|$.

(iii) For any $\epsilon>0$ and $T>0$, there exists a constant $c_2=(c_1,n)$ such that

$$\mathbb{P}\left(\max_{0\le t\le T}|q_t^\mu - q_t|>\epsilon\right) \le \frac{\mu}{\epsilon^2}c_2T\left[|p|^2+\|b\|^2+\|\sigma\sigma^\top\|\right]$$

Proof. Denoting $\dot{q}_t^\mu = p_t^\mu$, we transform Eq. (5) into a linear first order ODE

$$\dot{p}_t^\mu + \frac{1}{\mu}p_t^\mu = \frac{1}{\mu}b(t,q_t^\mu)+\frac{1}{\mu}\sigma(t,q_t^\mu)\dot{W}_t, \quad p_0^\mu = p$$

Notice that this is a first-order linear equation that can be solved using Integrating factor method

$$\dot{y}+p(t)y=f(t), \quad y(x_0)=y_0$$

with the corresponding solution as

$$y_t = e^{-\int_{x_0}^t p(s)ds}\left(\int_0^t e^{\int_{x_0}^s p(u)du}f(s)ds + y_0\right)$$

It follows that the solution is

$$p_t^\mu = pe^{\frac{-t}{\mu}} + \frac{1}{\mu}e^{\frac{-t}{\mu}}\int_0^t e^{\frac{s}{\mu}}b(s,q_s^\mu)ds + \frac{1}{\mu}e^{\frac{-t}{\mu}}\int_0^t e^{\frac{s}{\mu}}\sigma(s,q_s^\mu)dW_s$$

Since $p_t^\mu = \dot{q}_t^\mu$. Integrating this equation, we get

$$q_t^\mu = q_0^\mu + \int_0^t p_s^\mu\, ds = q + \int_0^t p_s^\mu\, ds$$

With Integration by parts, we obtain the final Eq. (5) as

$$q_t^\mu = q + p\mu\left(1-e^{-\frac{t}{\mu}}\right) + \left[-e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}b(s,q_s^\mu)\,ds + \int_0^t b(s,q_s^\mu)\,ds\right] + \left[-e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}\sigma(s,q_s^\mu)\,dW_s + \int_0^t \sigma(s,q_s^\mu)\,dW_s\right]$$

And the corresponding integral version of Itô diffusion stochastic differential equation in Eq. (6) is

$$q_t = q + \int_0^t b(s,q_s)ds + \int_0^t \sigma(s,q_s)dW_s$$

Now, we are well ready to prove each statement in the Lemma 2.2.

1. *Proof.* Assuming the diffusion part $\sigma(t,q)=0$ in both equations, we get:

$$\begin{aligned}
|q_t^\mu - q_t| &= \left| p\mu\left(1-e^{-\frac{t}{\mu}}\right) + \int_0^t \left[b(s,q_s^\mu)-b(s,q_s)\right]\,ds - e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}} b(s,q_s^\mu)\,ds \right| \\
&\leq \left| p\mu\left(1-e^{-\frac{t}{\mu}}\right)\right| + \left|\int_0^t \left[b(s,q_s^\mu)-b(s,q_s)\right]\,ds\right| + \left|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}} b(s,q_s^\mu)\,ds \right| \\
&\leq |p|\,\mu\left(1-e^{-\frac{t}{\mu}}\right) + \int_0^t K|q_s^\mu - q_s|\,ds + \left|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}} \sup_{0\leq t\leq T, q\in\mathbb{R}^n} b(t,q)\,ds\right| \qquad (7) \\
&\leq |p|\mu + K\int_0^t |q_s^\mu - q_s|ds + \mu \sup_{0\leq t\leq T, q\in\mathbb{R}^n} |b(t,q)| \\
&= \mu\left(|p| + \sup_{0\leq t\leq T, q\in\mathbb{R}^n} b(t,q)\right) + K\int_0^t |q_s^\mu - q_s|ds
\end{aligned}$$

where we have used the uniformly boundedness of $b(t,q)$ to obtain:

$$\begin{aligned}
\left|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}} b(s,q_s^\mu)\,ds\right| &\leq \left|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}} \sup_{0\leq t\leq T, q\in\mathbb{R}^n} b(t,q)\,ds\right| = \left|\mu \sup_{0\leq t\leq T, q\in\mathbb{R}^n} b(t,q)\left(1-e^{-\frac{t}{\mu}}\right)\right| \\
&\leq \mu \sup_{0\leq t\leq T, q\in\mathbb{R}^n} |b(t,q)|
\end{aligned} \qquad (8)$$

Applying the Grownwall Inequality (see Karatzas and Shreve (1998c) p. 287) to Eq. (10) to Eq. (7) leads to

$$\begin{aligned}
|q_t^\mu - q_t| &\leq \mu\left(|p| + \sup_{0\leq t\leq T, q\in\mathbb{R}^n} b(t,q)|\right)\exp\left(\int_0^t K ds\right) \\
&\leq e^{KT}\mu\left[|p| + \sup_{0\leq t\leq T, q\in\mathbb{R}^n} |b(t,q)|\right], \quad \forall t\in[0,T]
\end{aligned}$$

Since this inequality is valid for all $t\in[0,T]$, it follows that

$$\max_{0\leq t\leq T} |q_t^\mu - q_t| \leq e^{KT}\mu\left[|p| + \sup_{0\leq t\leq T, q\in\mathbb{R}^n} |b(t,q)|\right] \qquad (9)$$

Thus statement (i) is proved.

□

2. *Proof.* Similar to the proof of statement (i), by Cauchy-Swartz Inequality, Lipschitz Continuity, Uniformly Boundedness and Itô Isometry, we can obtain the following bound:

$$
\begin{aligned}
&\mathbb{E}\left[|q_t^\mu - q_t|^2\right] \\
&\le 5\left|p\mu\left(1-e^{-\frac{t}{\mu}}\right)\right|^2 + 5\mathbb{E}\left[\left|\int_0^t \left[b(s,q_s^\mu)-b(s,q_s)\right]\,ds\right|^2\right] + 5\mathbb{E}\left[\left|\int_0^t \left[\sigma(s,q_s^\mu)-\sigma(s,q_s)\right]\,dW_s\right|^2\right] \\
&\quad + 5\left|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}} b(s,q_s^\mu)\,ds\right|^2 + 5\mathbb{E}\left[\left|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}\sigma(s,q_s^\mu)\,dW_s\right|^2\right] \\
&\le 5|p\mu|^2 + 5t\mathbb{E}\left[\int_0^t \left|b(s,q_s^\mu)-b(s,q_s)\right|^2\,ds\right] + 5\mathbb{E}\left[\int_0^t \left|\sigma(s,q_s^\mu)-\sigma(s,q_s)\right|^2\,ds\right] \\
&\quad 5\mu^2\|b\|^2 + 5\mathbb{E}\left[\left|e^{-\frac{t}{\mu}}\right|^2\int_0^t \left|e^{\frac{s}{\mu}}\sigma(s,q_s^\mu)\right|^2\,ds\right] \\
&\le 5|p\mu|^2 + 5T\mathbb{E}\left[\int_0^t K^2|q_s^\mu - q_s|^2 ds\right] + 5\mathbb{E}\left[\int_0^t K^2|q_s^\mu - q_s|^2 ds\right] + 5\mu^2\|b\|^2 + 5\|\sigma\sigma^\top\|^2\frac{\mu}{2} \\
&= 5\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right) + (5T+1)K^2\int_0^t \mathbb{E}\left[|q_s^\mu - q_s|^2\right]ds
\end{aligned}
\tag{10}
$$

Again we bound the last term by Itô Isometry and the uniformly boundedness of $\sigma(t,q)$:

$$
\begin{aligned}
\mathbb{E}\left[\left|e^{-\frac{t}{\mu}}\int_0^t e^{\frac{s}{\mu}}\sigma(s,q_s^\mu)\,dW_s\right|^2\right] &= \mathbb{E}\left[\left|e^{-\frac{t}{\mu}}\right|^2\int_0^t \left|e^{\frac{s}{\mu}}\sigma(s,q_s^\mu)\right|^2\,ds\right] \le \mathbb{E}\left[e^{-\frac{2t}{\mu}}\int_0^t e^{\frac{2s}{\mu}}\|\sigma\sigma^\top\|^2\,ds\right] \\
&= \|\sigma\sigma^\top\|^2\mathbb{E}\left[\frac{\mu}{2}\left(1-e^{-\frac{2t}{\mu}}\right)\right] \\
&\le \|\sigma\sigma^\top\|^2\frac{\mu}{2}
\end{aligned}
\tag{11}
$$

Applying the Grownwall Inequality, we have

$$
\begin{aligned}
\mathbb{E}\left[|q_t^\mu - q_t|^2\right] &\le 5\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right)\exp\left(\int_0^t (5T+1)K^2\right)ds \\
&\le \underbrace{e^{(5T+1)TK^2}5}_{:=c_1}\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right) \\
&= c_1\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right), \quad \forall t\in[0,T]
\end{aligned}
$$

As a result, we obtain the statement (ii):

$$
\max_{0\le t\le T}\mathbb{E}\left[|q_t^\mu - q_t|^2\right] \le c_1\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right) = O(\mu) \tag{12}
$$

□

3. *Proof.* Consider the following difference

$$
\begin{aligned}
q_t^\mu - q_t = p\mu\left(1 - e^{-\frac{t}{\mu}}\right) + \int_0^t \left[b(s, q_s^\mu) - b(s, q_s)\right]\, ds + \int_0^t \left[\sigma(s, q_s^\mu) - \sigma(s, q_s)\right]\, dW_s \\
- e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} b(s, q_s^\mu)\, ds - e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} \sigma(s, q_s^\mu)\, dW_s
\end{aligned}
\tag{13}
$$

We decompose the difference into four principal components:

- Drift term: $X_t = \int_0^t \left[b(s, q_s^\mu) - b(s, q_s)\right] ds$
- Martingale term: $I_t = \int_0^t \left[\sigma(s, q_s^\mu) - \sigma(s, q_s)\right] dW_s$ where I_t is a martingale, so I_t^2 is a submartingale.
- Weighted/Scaled Martingale:

$$
M_t = -e^{-\frac{t}{\mu}} \int_0^t e^{\frac{s}{\mu}} \sigma(s, q_s^\mu)\, dW_s = -e^{-\frac{t}{\mu}} N_t
$$

  where N_t is a martingale, and is scaled by $-e^{-\frac{t}{\mu}}$. Importantly, by Eq. (11) observe that $\mathbb{E}\left[|M_t|^2\right] \le \|\sigma\sigma^\top\|^2 \frac{\mu}{2}$ for all $t \in [0, T]$.[1]

- Remaining terms: $O(\mu)$

This yields the compact decomposition of Eq. (13) :

$$
q_t^\mu - q_t = X_t + I_t + M_t + O(\mu) \tag{14}
$$

It follows that

$$
\mathbb{E}\left[\max_{0\le t\le T} |q_t^\mu - q_t|^2\right] \le 4\mathbb{E}\left[\max_{0\le t\le T} |X_t|^2\right] + 4\mathbb{E}\left[\max_{0\le t\le T} |I_t|^2\right] + 4\mathbb{E}\left[\max_{0\le t\le T} |M_t|^2\right] + O(\mu^2) \tag{15}
$$

Let us consider the terms separately.

$$
\begin{aligned}
\mathbb{E}\left[\max_{0\le t\le T} |X_t|^2\right] &= \mathbb{E}\left[\max_{0\le t\le T} \left|\int_0^t \left[b(s, q_s^\mu) - b(s, q_s)\right]\, ds\right|^2\right] \le \mathbb{E}\left[\max_{0\le t\le T} \left\{t \int_0^t |b(s, q_s^\mu) - b(s, q_s)|^2\, ds\right\}\right] \\
&\le \mathbb{E}\left[\max_{0\le t\le T} \left\{tK^2 \int_0^t |q_s^\mu - q_s|^2\, ds\right\}\right]
\end{aligned}
$$

Since the integral non-decreasing in t, so the bound for the maximum satisfies:

$$
\mathbb{E}\left[\max_{0\le t\le T} |X_t|^2\right] \le \mathbb{E}\left[TK^2 \int_0^T |q_s^\mu - q_s|^2\, ds\right] = TK^2 \int_0^T \mathbb{E}\left[|q_s^\mu - q_s|^2\right]\, ds
$$

[1] M_t^2 is in general neither a submartingale nor a supermartingale.

Recall in Eq. (10) that $\mathbb{E}\left[|q_t^\mu - q_t|^2\right] \leq c_1\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right)$ for all $t \in [0,T]$. Finally, we obtain

$$\begin{aligned}\mathbb{E}\left[\max_{0\leq t\leq T}|X_t|^2\right] &\leq TK^2\int_0^T c_1\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right)ds \\ &= c_1\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right)T^2K^2 = O(\mu)\end{aligned}$$

For the second term, by the Doob-Maximal inequality (see (Karatzas and Shreve, 1998a) p.14)

$$\begin{aligned}\mathbb{E}\left[\max_{0\leq t\leq T}|I_t|^2\right] &\leq 4\mathbb{E}[|I_T|^2] = 4\mathbb{E}\left[\left|\int_0^T\left[\sigma(s,q_s^\mu) - \sigma(s,q_s)\right]\,dW_s\right|^2\right] \leq 4\mathbb{E}\left[\int_0^T K^2|q_s^\mu - q_s|^2 ds\right] \\ &= 4K^2\int_0^T\mathbb{E}\left[|q_s^\mu - q_s|^2\right]ds \leq 4K^2\int_0^T c_1\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right)ds \\ &= c_1\mu\left(|p|^2\mu + \mu\|b\|^2 + \frac{1}{2}\|\sigma\sigma^\top\|^2\right)4K^2T = O(\mu)\end{aligned}$$

To bound the third term, we find the semimartingale decomposition (see (Karatzas and Shreve, 1998d) p.149) of $|M_t|^2 = e^{\frac{-2t}{\mu}}N_t^2$. Since the differential terms of N_t are

$$dN_t = e^{\frac{t}{\mu}}\sigma(t,q_t^\mu)dW_t, \quad d\langle N_t\rangle = e^{\frac{2t}{\mu}}|\sigma(t,q_t^\mu)|^2dt$$

And by Itô formula on $|M_t|^2 = f(t,N_t)$, where $f(t,x) = e^{\frac{-2t}{\mu}}x^2$, we get:

$$\begin{aligned}d(M_t^2) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dN_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}d\langle N_t\rangle \\ &= \left(-\tfrac{2}{\mu}e^{-\frac{2t}{\mu}}N_t^2 + |\sigma(t,q_t^\mu)|^2\right)dt + 2e^{-\frac{t}{\mu}}N_t\sigma(t,q_t^\mu)\,dW_t \\ &= \left(-\tfrac{2}{\mu}M_t^2 + |\sigma(t,q_t^\mu)|^2\right)dt - 2M_t\sigma(t,q_t^\mu)\,dW_t\end{aligned}$$

The integral version of this gives the decomposition:

$$M_t^2 = \underbrace{\int_0^t\left(-\frac{2}{\mu}M_s^2 + |\sigma(s,q_s^\mu)|^2\right)ds}_{\text{Finite Variation Part } A_t} + \underbrace{\int_0^t -2\sigma(s,q_s^\mu)M_s\,dW_s}_{\text{Local Martingale Part } Y_t} \tag{16}$$

We can now bound each term

$$\begin{aligned}\mathbb{E}\left[\max_{0\leq t\leq T}|A_t|\right] &\leq \mathbb{E}\left[\max_{0\leq t\leq T}\left\{\int_0^t\left|-\frac{2}{\mu}M_s^2 + |\sigma(s,q_s^\mu)|^2\right|ds\right\}\right] \leq \mathbb{E}\left[\int_0^T\frac{2}{\mu}M_s^2ds + \int_0^T\|\sigma\sigma^\top\|^2ds\right] \\ &= \int_0^T\frac{2}{\mu}\mathbb{E}\left[|M_s|^2\right]ds + \|\sigma\sigma^\top\|^2T \\ &\leq \int_0^T\frac{2}{\mu}\|\sigma\sigma^\top\|^2\frac{\mu}{2}ds + \|\sigma\sigma^\top\|^2T \\ &= 2\|\sigma\sigma^\top\|^2T\end{aligned}$$

Since the entire bound is independent of μ, we have as $\mu \to 0$:

$$\mathbb{E}\left[\max_{0\leq t\leq T}|A_t|\right] = O(1)$$

By the Burkholder-Davis-Gundy Inequality for continuous local martingale (see (Karatzas and Shreve, 1998d) p.166) and Jensen Inequality for concave function, we have:

$$
\begin{aligned}
\mathbb{E}\left[\max_{0\le t\le T}|Y_t|\right] &\le C_1\mathbb{E}\left[\langle Y_T\rangle^{1/2}\right] = C_1\mathbb{E}\left[\left(\int_0^T 4|\sigma(s,q_s^\mu)|^2|M_s|^2 ds\right)^{1/2}\right] \\
&\le C_1 2\|\sigma\sigma^\top\|\mathbb{E}\left[\left(\int_0^T |M_s|^2 ds\right)^{1/2}\right] \\
&\le C_1 2\|\sigma\sigma^\top\|\left(\mathbb{E}\left[\int_0^T |M_s|^2 ds\right]\right)^{1/2} \\
&\le C_1 2\|\sigma\sigma^\top\|\left(\|\sigma\sigma^\top\|^2\frac{\mu}{2}T\right)^{1/2} \\
&= C_1\|\sigma\sigma^\top\|^2\sqrt{2\mu T}
\end{aligned}
$$

This means the expected maximum of the local martingale part is of order: $O(\sqrt{\mu})$.

In total, from Eq. (16), we find that as $\mu \to 0$:

$$
\mathbb{E}\left[\max_{0\le t\le T}|M_t|^2\right] = O(1).
$$

Consequently, returning to the main estimate in Eq. (15), we conclude that as $\mu \to 0$

$$
\mathbb{E}\left[\max_{0\le t\le T}|q_t^\mu - q_t|^2\right] = O(1)
$$

Unfortunately, in this case, the proof does not lead to the desired result (iii) stated in Lemma 2.2. Applying Markov's inequality yields: for all $\epsilon > 0$:

$$
\mathbb{P}\left(\max_{0\le t\le T}|q_t^\mu - q_t| > \epsilon\right) \le \frac{O(1)}{\epsilon^2}
$$

which provides no decay as $\mu \to 0$. That is, the bound does not vanish in the small-mass limit, and thus we cannot conclude convergence in probability from this estimate alone.

$\square$

$\square$

Remark. *We conclude the paper with the following observations regarding Lemma 2.2:*

1. *In the deterministic case where $\sigma(t,q) = 0$, the diffusion terms vanish. From the proof of statement (i), we observe that the convergence is in fact uniform and occurs at a linear rate:*

$$\max_{0\le t\le T} |q_t^\mu - q_t| = O(\mu).$$

2. *From statement (ii), applying Jensen's inequality yields:*

$$\max_{0\le t\le T} \left(\mathbb{E}\left[|q_t^\mu - q_t|\right]\right)^2 \le \max_{0\le t\le T} \mathbb{E}\left[|q_t^\mu - q_t|^2\right] \le O(\mu),$$

   which implies:

$$\max_{0\le t\le T} \mathbb{E}\left[|q_t^\mu - q_t|\right] \le O(\sqrt{\mu}).$$

   In the stochastic setting, the noise introduces fluctuations that persist even for small μ, slowing down the convergence compared to the deterministic case.

3. *In the statement (ii), we observe that as $\mu \to 0$ we can conclude that q_t^μ is a modification of q_t, that is, for all $0 \le t \le T$, we get $\mathbb{P}(|q_t^\mu - q_t|) = 1$.*

   Since both stochastic processes $\{q_t^\mu, 0 \le t \le T\}$ and $\{q_t, 0 \le t \le T\}$ were assumed to be continuous, in particular right-continuous, by Lemma Karatzas and Shreve (1998b) they are indistinguishable, as $\mu \to 0$:

$$\mathbb{P}\left(\max_{0\le t\le T} |q_t^\mu - q_t| > \epsilon\right) \to 0 \quad \textit{as } \mu \to 0.$$

   This implies that the two stochastic processes are pathwise unique a.s when the mass $\mu \to 0$.

References

Freidlin, M. (2004). Some remarks on the smoluchowski–kramers approximation. *Journal of Statistical Physics*, 117:617–634.

Karatzas, I. and Shreve, S. E. (1998a). *Martingales, Stopping Times, and Filtrations*, pages 1–46. Springer New York, New York, NY.

Karatzas, I. and Shreve, S. E. (1998b). *Martingales, Stopping Times, and Filtrations*, pages 1–46. Springer New York, New York, NY.

Karatzas, I. and Shreve, S. E. (1998c). *Stochastic Differential Equations*, pages 281–398. Springer New York, New York, NY.

Karatzas, I. and Shreve, S. E. (1998d). *Stochastic Integration*, pages 128–238. Springer New York, New York, NY.