Lecture notes

Particle systems, large-deviation and variational approaches to generalised gradient flows

Manh Hong Duong
Mathematics Institute, University of Warwick
Email: m.h.duong@warwick.ac.uk

January 4, 2017

Abstract

In 1998, Jordan-Kinderleher-Otto [JKO98] proved a remarkable result that the diffusion equation can be seen as a *gradient flow* of the *Boltzmann entropy* with respect to the *Wasserstein distance*. This result has sparked off a large body of research in the field of partial differential equations and others in the last two decades. Many evolution equations have been proved to have a Wasserstein gradient flow structure such as the convection and nonlinear diffusion, the Cahn-Hilliard equation, the thin-film equation and finite Markov chains, just to name a few. Not only revealing physical nature of a PDE, a Wasserstein gradient flow structure can also be exploited to prove its well-posedness, to characterise long-time behaviour and to study multi-scale analysis. Recently, Adams-Dirr-Peletier-Zimmer [ADPZ11] has established an intriguing connection between the JKO-Wasserstein gradient flow structure of the diffusion equation with large-deviation principle of many Brownian motions showing that the former can be derived from the latter. This result explains, among other things, the microscopic origin of the combination of the Wasserstein metric and the Boltzmann entropy that appeared in the JKO-scheme. In [DLR13, MPR14, EMR15] this result has been generalised to other systems including the Fokker-Planck and general Markov process with detailed balance.

However, the Wasserstein gradient flow theory is only applicable to dissipative systems. In nature and applied sciences, there exist many non-dissipative systems. A typical example is the Kramers (or kinetic Fokker Planck equation) that has been used extensively in statistical mechanics and chemistry. In fact, the GENERIC (General Equation for Non-Equilibrium Reversible-Irreversible Coupling [Ött05]) framework covers a large class of evolution equations that consist both conservative and dissipative dynamics. A nature question is whether one can we generalise [JKO98] and [ADPZ11] to such systems. In a series of papers [DPZ13a, DPZ13b, DLPS16, DLP$^+$16] we address this question. More specially, we show that a GENERIC structure of the Vlasov-Fokker-Planck equation is ultimately related to a large-deviation principle of an underlying stochastic process [DPZ13a, DPZ13b]. Based on this connection, we introduce new technique for coarse-graining (multi-scale analysis), both qualitative and quantitative, of conservative-dissipative systems [DLPS16, DLP$^+$16].

The aim of this crash course is to introduce these recent developments. That is to (1) provide a brief introduction to generalised Wasserstein gradient flows and large-deviation principles, (2) present connections between the two theories and (3) discuss about applications in multi-scale analysis of PDEs. This notes is mainly based on my PhD thesis carried out at the Eindhoven University of Technology. However, basic knowledge, mostly in Section 1, is included. Due to the limitation of time, only main heuristic ideas and main steps of proofs are presented; all details can be found in [Duo14, DPZ13a, DPZ13b, DLPS16, DLP$^+$16].

This mini-course was given at Department of Probability and Statistics, Hanoi Institute of Mathematics in Summer 2015.

Contents

1 Wasserstein gradient flows

The main aim of this section is to introduce the main result of [JKO98] that proves that the diffusion equation is a *gradient flow* of the *Boltzmann entropy* with respect to the *Wassersetin metric*. In order to state this result, we need to introduce thee relevant concepts first. We will review the notion of a gradient flow in a finite dimensional space, its approximation scheme and weak formulations. Then we will see how this is generalised to a probabilistic space endowed with the Wasserstein metric. Having all these concepts, we finally will be able to state and discuss the main result of [JKO98].

1.1 Gradient flows in $\mathbf{R}^d$

To begin with let us recall that the gradient flow of a smooth functional $\mathcal{E}\colon \mathbf{R}^d \to \mathbf{R}$ is a map $x\colon [0,\infty) \to \mathbf{R}^d$ which solves the differential equation

$$\begin{aligned} \frac{d}{dt}x(t) &= -\nabla\mathcal{E}(x(t)) \quad \text{for} \quad t>0, \\ x(0) &= x_0. \end{aligned} \tag{1}$$

1.2 Discrete approximation of gradient flows

To numerically solve the equation (1) we often use the implicit Euler scheme as follows. Take $h>0$ as a time step.

1. Step 1: $x_0^h = x_0$,
2. Step 2: Assume that $x_k^h = x(kh)$ for $k = 0,1,...,n$ are known. Find x_{n+1}^h from the following equation

$$\frac{x_{n+1}^h - x_n^h}{h} + \nabla\mathcal{E}(x_{n+1}^h) = 0. \tag{2}$$

Note that a solution of (2) is also a minimizer of the minimization problem

$$\min_{x\in\mathbf{R}^d} \frac{\|x-x_n^h\|^2}{2h} + \mathcal{E}(x) - \mathcal{E}(x_n^h), \tag{3}$$

where $\|\cdot\|$ is the Euclidean metric. Hence solving (2) can be achieved by finding x_{n+1}^h that minimizes over $x \in \mathbf{R}^d$

$$\mathcal{K}(x) = \frac{\|x-x_n^h\|^2}{2h} + \mathcal{E}(x) - \mathcal{E}(x_n^h). \tag{4}$$

Therefore the gradient flow (1) can be viewed as steepest descent of the functional $\mathcal{K}(x)$ with respect to the Euclidean distance.

1.3 Gradient flows on Riemannian manifolds

Let $\mathcal{M}$ be a Riemannian manifold, i.e. $\mathcal{M}$ is a real differentiable manifold, and for each $x \in \mathcal{M}$ there is an inner product g_x on the tangent space $T_x\mathcal{M}$.

Let $\mathcal{E}\colon \mathcal{M} \to \mathbf{R}$ be differentiable, $x \in \mathcal{M}$ and v be a tangent vector at x. The directional derivative of $\mathcal{E}$ at x along v, $\delta_v\mathcal{E}(x)$, is defined as follows. Let $\gamma(t)$ be a differentiable curve in $\mathcal{M}$ with $\gamma(0) = x$ and $\dot{\gamma}(0) = v$. Then

$$\delta_v\mathcal{E}(x) := \frac{d}{dt}\mathcal{E}(\gamma(t))\Big|_{t=0}. \tag{5}$$

The gradient of $\mathcal{E}$, denoted by $\nabla\mathcal{E}$, is the vector field defined by

$$\begin{aligned} \nabla\mathcal{E}\colon \mathcal{M} &\to T\mathcal{M} \\ x &\mapsto \nabla\mathcal{E}(x) \in T_x\mathcal{M}, \end{aligned}$$

such that

$$g_x(\nabla\mathcal{E}(x), v) = \delta_v\mathcal{E}(x) \text{ for all } v \in T_x\mathcal{M}. \tag{6}$$

The gradient flow of $\mathcal{E}$ is a curve $x\colon [0,\infty) \to \mathcal{M}$ which solves the differentiable equation

$$\frac{d}{dt}x(t) = -\nabla\mathcal{E}(x(t)) \text{ in } T_{x(t)}\mathcal{M}. \tag{7}$$

1.4 Weak formulation of the gradient flows

An alternative way to formulate the gradient flow in (1) relies on using its weak form. Let $x\colon [0,\infty) \to \mathbf{R}^d$ be any differentiable map. We always have

$$\begin{aligned} \frac{d}{dt}\mathcal{E}(x(t)) &= \langle \nabla\mathcal{E}(x(t)), \dot{x}(t)\rangle \\ &\geq -\|\nabla\mathcal{E}(x(t))\| \cdot \|\dot{x}(t)\| \\ &\geq -\frac{1}{2}\|\nabla\mathcal{E}(x(t))\|^2 - \frac{1}{2}\|\dot{x}(t)\|^2. \end{aligned}$$

The inequality becomes equality if and only if $\dot{x}(t)$ has opposite direction and same length as $\nabla\mathcal{E}(x(t))$, i.e., $\dot{x}(t) = -\nabla\mathcal{E}(x(t))$. Hence we can reformulate the gradient flow in (1) as follows.

Theorem 1.1 (Rayleigh principle)**.** *Assume that $\mathcal{E} \in C^1(\mathbf{R}^d)$. A curve $x \in C^1([0,T];\mathbf{R}^n)$ is gradient flow of $\mathcal{E}$ if and only if for each $t \in [0,T]$*

$$\frac{d}{dt}\mathcal{E}(x(t)) \leq -\frac{1}{2}\|\nabla\mathcal{E}(x(t))\|^2 - \frac{1}{2}\|\dot{x}(t)\|^2, \tag{8}$$

or equivalently

$$\mathcal{E}(x(T)) - \mathcal{E}(x(0)) + \frac{1}{2}\int_0^T \left[\|\nabla\mathcal{E}(x(t))\|^2 + \|\dot{x}(t)\|^2\right] dt \leq 0. \tag{9}$$

This formulation can be generalized to the gradient flows on Riemannian manifold in the previous section. Then (9) becomes

$$\mathcal{E}(x(T)) - \mathcal{E}(x(0)) + \frac{1}{2}\int_0^T \left[g_{x(t)}(\nabla\mathcal{E}(x(t)), \nabla\mathcal{E}(x(t))) + g_{x(t)}(\dot{x}(t), \dot{x}(t))\right] dt \le 0. \tag{10}$$

1.5 Wasserstein distance

Let $\mathcal{P}_2(\mathbf{R}^d)$ be the set of probability measures with finite second moment, i.e.,

$$\mathcal{P}_2(\mathbf{R}^d) = \left\{\rho(dx) \,\Big|\, \int_{\mathbf{R}^d} \rho(dx) = 1 \text{ and } \int_{\mathbf{R}^d} |x|^2 \rho(dx) < \infty\right\}. \tag{11}$$

Given $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbf{R}^d)$, we denote by $\Gamma(\mu_0, \mu_1)$ the set of probability measures in $\mathbf{R}^d \times \mathbf{R}^d$ having μ_0, μ_1 as its marginals, i.e.,

$$\Gamma(\mu_0, \mu_1) = \{\gamma \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d) \,\Big|\, \gamma(A \times \mathbf{R}^d) = \mu_0(A), \gamma(\mathbf{R}^d \times A) = \mu_1(A) \text{ for all Borel sets } A \subset \mathbf{R}^d\}. \tag{12}$$

The 2-Wasserstein distance between μ_0 and $\mu_1 \in \mathcal{P}_2(\mathbf{R}^d)$ is defined via

$$d(\mu_0, \mu_1)^2 = \inf_{\gamma \in \Gamma(\mu_0, \mu_1)} \int_{\mathbf{R}^d \times \mathbf{R}^d} |x - y|^2 \gamma(dxdy). \tag{13}$$

Example 1.1. In several cases, the Wasserstein distance can be computed explicitly.

(1) $W_2(\delta_a, \delta_b) = |a - b|$.

(2) $W_2\left(\frac{1}{n}\sum_{i=1}^n \delta_{x_i}, \frac{1}{n}\sum_{i=1}^n \delta_{y_i}\right)^2 = \min_{\sigma \in S_n} \frac{1}{n}\sum_{i=1}^n |x_i - y_{\sigma(i)}|^2$.

(3) $W_2(\mathcal{N}(a, \sigma_1^2), \mathcal{N}(b, \sigma_2^2))^2 = (a - b)^2 + (\sigma_1 - \sigma_2)^2$.

1.6 $\mathcal{P}_2(\mathbf{R}^d)$ as a manifold

Remind that in a Riemannian manifold $\mathcal{M}$ the distance between two points x_0 and x_1 is defined via

$$d(x_0, x_1)^2 = \inf\{\int_0^1 g_{x(t)}(\dot{x}(t), \dot{x}(t))dt \,\Big|\, x \in C^1([0,1], \mathcal{M}), x(0) = x_0, x(1) = x_1\}. \tag{14}$$

Brenier and Benamou [BB00] have shown a similar formula for the Wasserstein distance. Let $\rho_0(dx) = \rho_0(x)dx, \rho_1(dx) = \rho_1(x)dx \in \mathcal{P}_2(\mathbf{R}^d)$. Then

$$\begin{aligned} d(\rho_0, \rho_1)^2 &= \inf\{\int_0^1 \int_{\mathbf{R}^d} |\nabla u|^2 \, d\rho(t) \, dt \,\Big|\, \dot{\rho}(t) = -\nabla \cdot (\rho \nabla u), \rho(0) = \rho_0, \rho(1) = \rho_1\} \quad (15) \\ &= \inf\{\int_0^1 g_{\rho(t)}(\dot{\rho}(t), \dot{\rho}(t)) \, dt \,\Big|\, \dot{\rho}(t) = -\nabla \cdot (\rho \nabla u), \rho(0) = \rho_0, \rho(1) = \rho_1\}, \quad (16) \end{aligned}$$

where

$$g_{\rho(t)}(\dot\rho(t),\dot\rho(t)) = \int_{\mathbf{R}^d} |\nabla u|^2 d\rho(t). \tag{17}$$

The similarity between (14) and (16) suggests that we can formally view $\mathcal{P}_2(\mathbf{R}^d)$ as a Riemannian manifold and that we can identify the tangent space at $\rho(t)$, $T_{\rho(t)}\mathcal{P}_2(\mathbf{R}^d)$, with the family of functions u satisfying the continuity equation in (16)

$$T_{\rho(t)}\mathcal{P}_2(\mathbf{R}^d) = \left\{u(t)\middle| u \in C^2([0,1];\mathbf{R}^d) \text{ such that } \dot\rho(t) = -\nabla\cdot(\rho\nabla u)\right\}, \tag{18}$$

and the Riemannian metric $g_{\rho(t)}(\dot\rho(t),\dot\rho(t))$ is computed as in (17). By generalizing this we can define the tangent space and the Riemannian metric at any point $\rho \in \mathcal{P}_2(\mathbf{R}^d)$. This interpretation indeed was done first by Otto in [Ott01] and intensively later by Ambrosio, Gigli and Savaré in the second part of the book [AGS08]. The appendix D.5 of [FK06] also discusses on this.

In [Ott01] the tangent space $T_\rho\mathcal{P}_2(\mathbf{R}^d)$ of ρ is defined by

$$s \in T_\rho\mathcal{P}_2(\mathbf{R}^d) \Leftrightarrow s = -\nabla\cdot(\rho\nabla u), \tag{19}$$

and the inner product by

$$g_\rho(s,s) = \int_{\mathbf{R}^d} |\nabla u|^2 d\rho. \tag{20}$$

Hence for $s_1, s_2 \in T_\rho\mathcal{P}_2(\mathbf{R}^d)$ we have

$$g_\rho(s_1,s_2) = \frac{1}{4}[g_\rho(s_1+s_2) - g_\rho(s_1-s_2)] = \int_{\mathbf{R}^d} \nabla u_1\cdot\nabla u_2 d\rho, \tag{21}$$

where $s_1 = -\nabla\cdot(\rho\nabla u_1)$ and $s_2 = -\nabla\cdot(\rho\nabla u_2)$.

In [AGS08] the relationship in (19) is made precise as

$$T_\rho\mathcal{P}_2(\mathbf{R}^d) = \overline{\{\nabla\varphi\colon \varphi \in C_c^\infty(\mathbf{R}^d)\}}^{L^2_\rho(\mathbf{R}^d)}. \tag{22}$$

In [FK06] $T_\rho\mathcal{P}_2(\mathbf{R}^d)$ is identified with the space $H_{-1,\rho}(\mathbf{R}^d)$

$$H_{-1,\rho(\mathbf{R}^d)} = \left\{u \in \mathcal{D}'(\mathbf{R}^d)\colon \|u\|_{-1,\rho} < \infty\right\}, \tag{23}$$

where $\mathcal{D}'(\mathbf{R}^d)$ is the space of Schwartz distributions on $\mathbf{R}^d$ and

$$\|u\|^2_{-1,\rho} = \sup_{\varphi\in C_c^\infty(\mathbf{R}^d)} \left\{2\langle u,\varphi\rangle - \int_{\mathbf{R}^d} |\nabla\varphi|^2 d\rho\right\}. \tag{24}$$

With this identification in [FK06] the gradient of a functional on $\mathcal{P}_2(\mathbf{R}^d)$ is defined as follows

Definition 1.2. ([FK06, Definition 9.36]) Let $\mathcal{E}\colon \mathcal{P}_2(\mathbf{R}^d) \to [-\infty, +\infty]$ and $\rho \in \mathcal{P}_2(\mathbf{R}^d)$. We say that gradient of $\mathcal{E}$ at ρ, denoted by $\nabla\mathcal{E}(\rho)$ exists if it can be identified as the unique element in $\mathcal{D}'(\mathbf{R}^d)$ such that for each $\varphi \in C_c^\infty(\mathbf{R}^d)$ and each $\rho(t)\colon [0,\infty) \to \mathcal{P}_2(\mathbf{R}^d)$ satisfying the continuity equation

$$\partial_t\rho(t) + \operatorname{div}(\rho(t)\nabla\varphi) = 0,\ \rho(0) = \rho \ \text{ in } \mathcal{D}'(\mathbf{R}^d). \tag{25}$$

we have

$$\lim_{t\to 0^+} \frac{\mathcal{E}(\rho(t)) - \mathcal{E}(\rho)}{t} =: \langle \nabla\mathcal{E}(\rho), \varphi\rangle \tag{26}$$

1.7 The result of Jordan-Kinderlehrer-Otto 1998

In 1998 Jordan-Kinderlehrer and Otto [JKO98] have made an important discovery that the heat equation $\partial_t\rho = \Delta\rho$ can be expressed as steepest descent of the entropy functional $\mathcal{E}(\rho) = \int \rho\log\rho$ with respect to the so-called Wasserstein distance in the space of probability measures on $\mathbf{R}^d$. We now will recall the result in [JKO98].

Consider the heat equation

$$\partial_t\rho = \Delta\rho. \tag{27}$$

(In [JKO98] the authors actually considered a more general equation $\partial_t\rho = \Delta\rho + \operatorname{div}(\nabla\Phi(x)\rho)$. (27) is a special case of this equation when $\Phi(x) \equiv 0$).

Define the entropy $\mathcal{E}(\rho) = \int_{\mathbf{R}^d} \rho\log\rho$. The main result in [JKO98] is the following.

Theorem 1.3. *[JKO98] Let $\rho_0 \in \mathcal{P}_2(\mathbf{R}^d)$ satisfy $\mathcal{E}(\rho_0) < \infty$, and for a given $h > 0$, let ρ_h^k be the solution of the following scheme.*

- $\rho_h^0 = \rho_0$,
- *Determine ρ_h^k that minimizes over $\rho \in \mathcal{P}_2(\mathbf{R}^d)$*

$$K_h(\rho) = \frac{1}{2h} d(\rho_h^{k-1}, \rho)^2 + \mathcal{E}(\rho) - \mathcal{E}(\rho_h^{k-1}). \tag{28}$$

Define $\rho_h : (0,\infty) \times \mathbf{R}^d \to [0,\infty)$ by

$$\rho_h(t) = \rho_h^k \text{ for } t \in [kh, (k+1)h). \tag{29}$$

Then $\rho_h(t) \rightharpoonup \rho(t)$ weakly in $L^1(\mathbf{R}^d)$ for all $t \in (0,\infty)$ where $\rho \in C^\infty((0,\infty)\times\mathbf{R}^d)$ is the unique solution of the heat equation.

We notice that (4) and (28) have the same form. Hence the result above generalizes the implicit Euler scheme to the infinite dimensional space $\mathcal{P}_2(\mathbf{R}^d)$ equipped with the 2-Wasserstein distance instead of the Euclidean one.

The theorem above has sparked off a large body of research in the field of partial differential equations and others in the last two decades. Many evolution equations have been proved to have a Wasserstein gradient flow structure such as the convection and nonlinear diffusion, the Cahn-Hilliard equation, the thin-film equation and finite Markov chains, just to name a few. See [AGS08, Vil03] for excellent expositions about this topic.

However, in the theorem above it is not clear why the Wasserstein metric and the Boltzmann entropy appear, why their combination gives rise to the diffusion equation. In [ADPZ11], the author made the first attempt to answer this question. In order to introduce this result, we need to introduce another concept, large-deviation principles of stochastic processes, in the next section.

2 Large deviation principle

In this section, we review relevant knowledge on the theory of large-deviation principle. We refer to [DZ87, FK06] for the full treatment of this theory.

Definition 2.1. Let X be a complete separable metric space. A sequence of $X-$valued random variables X_n is said to satisfy the large deviation principle with a rate functional $I\colon X \to [0,\infty)$ if

1. $\liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}(X_n \in A) \geq -\inf_{x\in A} I(x)$ for all open Borel subsets $A \subset X$,
2. $\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(X_n \in B) \leq -\inf_{x\in B} I(x)$ for all closed Borel subsets $B \subset X$.

The rate functional I is good if its sub-level sets $\big\{x \in X \big| I(x) \leq a\big\}$ are compact for all $a \geq 0$.

2.1 Inverse of Varadhan's lemma

Lemma 2.2. *Let $\{X_n\}$ be a sequence of $\mathcal{S}$-valued RVs. Suppose that the sequence $\{X_n\}$ is exponentially tight and that the limit*

$$\Lambda(f) = \lim_{n\to\infty} \frac{1}{n} \log E[e^{nf(X_n)}]$$

exists for each $f \in C_b(\mathcal{S})$. Then $\{X_n\}$ satisfies the LDP with a rate functional

$$I(x) = \sup_{f\in C_b(\mathcal{S})} \{f(x) - \Lambda(f)\}.$$

2.2 Contraction principle

Lemma 2.3 (contraction principle)**.** *Let X_n be an $\mathcal{S}$-RVs with a tight distribution. Suppose that X_n satisfies a LDP with a good rate functional I. Let $(\mathcal{S}', d')$ be a metric space and* *suppose that $F : \mathcal{S} \to \mathcal{S}'$ is a measurable and continuous at $x \in \mathcal{S}$ for each x with $I(x) < \infty$. Define $Y_n = F \circ X_n$. Then Y_n also satisfies a LDP with a good rate functional*

$$I'(y) = \inf\{I(x) : F(x) = y\}.$$

2.3 Examples of LDP

Example 1 (Cramer theorem): X_n are i.i.d real RVs with common generating function $M(\theta) = E(\exp(\theta X_1))$, and $I(a) = \sup_\theta\{\theta a - \log M(\theta)\}$ is the Legendre transform of M. Then $\frac{1}{n}\sum_i X_i$ satisfies a LDP with rate functional I.

Example 1.1 (coin tossing) $P(X_n) = 0 = P(X_n) = 1 = \frac{1}{2}$. Then $M(\theta) = \frac{1}{2}(1+e^\theta)$, and

$$\begin{aligned} I(a) &= \sup\{\theta a - \log \frac{1}{2}(1+e^\theta)\} \\ &= \begin{cases} a\log a + (1-a)\log(1-a) + \log 2 & \text{if} \quad 0 \le a \le 1, \\ +\infty, \quad \text{otherwise.} \end{cases} \end{aligned}$$

Example 2 (Sanov theorem): X_n are i.i.d with common distribution ν. Then $\frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ satisfies a LDP with the rate $H(\cdot||\nu)$, which is the relative entropy w.r.t ν.

$$H(\mu||\nu) = \begin{cases} \int d\mu \log \frac{d\mu}{d\nu}, & \text{if} \quad d\mu \ll d\nu, \\ +\infty, \quad \text{others} \end{cases}$$

3 Connection between gradient flows and Large Deviation Principles

In this section, we introduce the main result in [ADPZ11] that establishes an intriguing connection between Wasserstein gradient flow structure of the diffusion equation with large-deviation principle of many Brownian motions.

3.1 The result of [ADPZ11]: Connection via Gamma convergence

Let us first remind the definition of Gamma convergence in a metric space and narrow convergence in a probability space.

Definition 3.1. [Bra02] Let X be a metric space. We say that a sequence $f_n\colon X \to \overline{\mathbf{R}}$ $\Gamma-$converges in X to $f\colon X \to \overline{\mathbf{R}}$, denoted by $f_n \xrightarrow{\Gamma} f$, if for all $x \in X$ we have

- For every sequence x_n converging to x

$$\liminf_{n\to\infty} f_n(x_n) \geq f(x), \tag{30}$$

- There exists a sequence x_n converging to x such that

$$\lim_{n\to\infty} f_n(x_n) = f(x). \tag{31}$$

Definition 3.2. We say that a sequence $\rho_n \in \mathcal{P}(\mathbf{R}^d)$ narrowly converges to $\rho \in \mathcal{P}(\mathbf{R}^d)$, denoted by $\rho_n \Rightarrow \rho$, if for all $f \in C_b(\mathbf{R}^d)$ we have

$$\int_{\mathbf{R}^d} f d\rho_n \to \int_{\mathbf{R}^d} f d\rho. \tag{32}$$

In [ADPZ11] the authors consider a family of n independent Brownian particles $X_i(t) \in \mathbf{R}, t \geq 0$ and they examine the empirical measure

$$L_n^t := \frac{1}{n}\sum_{i=1}^n \delta_{X_i(t)}. \tag{33}$$

Let $\mathbb{P}_n = \bigotimes_{i=1}^n \mathbb{P}_{\rho_0}$ where $\mathbb{P}_{\rho_0}$ is the probability measure under which a particle starts with initial distribution ρ_0. Let $h, \delta > 0$ and $\rho_0 \in \mathcal{P}(\mathbf{R})$ be given. Denote by $B_\delta(\rho_0)$ the open ball of radius δ with respect to the Levy metric on $\mathcal{P}(\mathbf{R})$. The authors showed that the sequence $\mathbb{P}_n \circ (L_n^h)^{-1}$ satisfies under the condition that $L_n^0 \in B_\delta(\rho_0)$ a large deviation principle with rate functional $J_{h,\delta}(\rho, \rho_0)$ and $J_{h,\delta}(\rho, \rho_0) \xrightarrow{\Gamma} J_h(\rho, \rho_0)$ in $\mathcal{P}(\mathbf{R})$ where

$$J_h(\rho, \rho_0) = \inf_{q \in \Gamma(\rho, \rho_0)} H(q|q_0), \tag{34}$$

with $q_0(dxdy) = \rho_0(dx)\frac{1}{\sqrt{4\pi h}}e^{-\frac{(y-x)^2}{4h}}dy$ and

$$H(q|q_0) = \begin{cases} \int_{\mathbf{R}\times\mathbf{R}} \frac{dq}{dq_0} \log \frac{dq}{dq_0} q_0(dxdy) & \text{if } q \ll q_0 \\ +\infty & \text{else.} \end{cases} \tag{35}$$

Moreover the rate functional $J_h(\rho, \rho_0)$ is closely related to the entropy functional as stated in the main theorem in [ADPZ11].

Theorem 3.3. *[ADPZ11] Let $L > 0$ be fixed. There exists $\delta > 0$ such that for each $\rho_0 \in A_\delta \cap C([0, L])$, where $A_\delta = \{\rho \in L^\infty(0, L) : \int_0^L \rho(dx) = 1$ and $\|\rho - L^{-1}\| < \delta\}$,*

$$J_h(\cdot; \rho_0) - \frac{1}{4h}d(\cdot, \rho_0)^2 \xrightarrow{\Gamma} \frac{1}{2}E(\cdot) - \frac{1}{2}E(\rho_0) \text{ as } h \to 0 \text{ in the set } A_\delta. \tag{36}$$

This means that

1. *For each sequence ρ^h converging narrowly to ρ in A_δ we have*

$$\liminf_{h\to 0} J_h(\rho^h, \rho_0) - \frac{1}{4h}d(\rho^h, \rho_0)^2 \geq \frac{1}{2}E(\rho^h) - \frac{1}{2}E(\rho_0). \tag{37}$$

2. *For each $\rho \in A_\delta$, there exists a sequence $(\rho^h) \subset A_\delta$ with $\rho^h \Rightarrow \rho$ such that*

$$\lim_{h\to 0} J_h(\rho^h, \rho_0) - \frac{1}{4h}d(\rho^h, \rho_0)^2 = \frac{1}{2}E(\rho^h) - \frac{1}{2}E(\rho_0). \tag{38}$$

3.2 Weak formulation of gradient flow and the rate functional: a direct connection

In this subsection we will analyze the example 14 in the book [FK06] to see a direct connection between the gradient flow and large deviation principle. Let $\Psi, \Phi \in C^2(\mathbf{R}^d)$ and consider the equation

$$\frac{\partial}{\partial t}\rho = \nabla \cdot (\rho \nabla(\Psi + \rho * \Phi)) + \frac{1}{2}\Delta\rho. \tag{39}$$

where $\rho * \Phi(x) = \int_{\mathbf{R}^d} \Phi(x-y)\rho(dy)$. Let $\rho \in \mathcal{P}_2(\mathbf{R}^d)$ such that $\rho(dx) = \rho(x)dx$. Define

$$\mathcal{E}(\rho) = \frac{1}{2}\log Z + \frac{1}{2}\int_{\mathbf{R}^d} \rho(x)\log\rho(x)\,dx + \int_{\mathbf{R}^d} \Psi(x)dx + \frac{1}{2}\int_{\mathbf{R}^d\times\mathbf{R}^d} \Phi(x-y)\rho(x)\rho(y)dxdy, \tag{40}$$

where Z is a constant. It is shown in chapter 9 (see also theorem D.28) in [FK06] that the gradient of $\mathcal{E}$ at ρ is

$$\nabla\mathcal{E}(\rho) = -\frac{1}{2}\Delta\rho - \nabla \cdot (\rho\nabla(\Psi + \rho * \Phi)). \tag{41}$$

and that the equation (39) is the gradient flow of $\mathcal{E}(\rho)$. Hence we can rewrite this equation in the weak form as in (10) as follows

$$\mathcal{E}(\rho(T)) - \mathcal{E}(\rho(0)) + \frac{1}{2}\int_0^T \left(\|\frac{\partial\rho}{\partial t}\|^2_{-1,\rho(t)} + \|\nabla\mathcal{E}(\rho(t))\|^2_{-1,\rho(t)}\right)dt \le 0. \tag{42}$$

Now let consider a stochastic process

$$dX_{i,n}(t) = -\nabla\Psi(X_{i,n}(t))dt - \frac{1}{n}\sum_{j=1}^n \nabla\Phi(X_{i,n}(t) - X_{j,n}(t))dt + dW_i(t). \tag{43}$$

for $i = 1, \ldots, n$, where $\{W_i : i = 1, \ldots, n\}$ are independent $\mathbf{R}^d$-valued Brownian motions. Let consider the empirical process

$$\rho_n(t, dx) = \frac{1}{n}\sum_{i=1}^n \delta_{X_{i,n}(t)}(dx). \tag{44}$$

Theorem 3.4. *[FK06, Theorem 13.37] Under some certain conditions $\rho_n(t,dx)$ satisfies a large deviation principle in the path space $C_{\mathcal{P}_2(\mathbf{R}^d)}[0,\infty)$ with a good rate functional $I\colon C_{\mathcal{P}_2(\mathbf{R}^d)}[0,\infty) \to [0,\infty]$ with*

$$I(\rho) = I_0(\rho(0)) + \frac{1}{2}\int_0^\infty \|\frac{\partial}{\partial t}\rho - \frac{1}{2}\Delta\rho - \nabla\cdot(\rho\nabla(\Psi + \rho * \Phi))\|^2_{-1,\rho(t)}dt. \tag{45}$$

where $I_0(\rho(0))$ is the rate functional for the initial process $\rho_n(0)$.

Let $T > 0$. By the remark 8.15 in [FK06] if we restrict to a bounded time interval $[0, T]$ then for each $\rho \in C_{\mathcal{P}_2(\mathbf{R}^d)}[0, T]$ the rate functional becomes

$$I^T(\rho) = \frac{1}{2}\int_0^T \|\frac{\partial}{\partial t}\rho - \frac{1}{2}\Delta\rho - \nabla\cdot(\rho\nabla(\Psi + \rho * \Phi))\|^2_{-1,\rho(t)}dt. \tag{46}$$

By (41) we can rewrite the rate functional $I^T(\rho)$ as follows

$$\begin{aligned} I^T(\rho) &= \frac{1}{2}\int_0^T \|\frac{\partial}{\partial t}\rho + \nabla\mathcal{E}(\rho(t))\|^2_{-1,\rho(t)}dt \\ &= \frac{1}{2}\int_0^T \left(\|\frac{\partial\rho}{\partial t}\|^2_{-1,\rho(t)} + \|\nabla\mathcal{E}(\rho(t))\|^2_{-1,\rho(t)} + 2\langle\frac{\partial\rho}{\partial t}, \nabla\mathcal{E}(\rho(t))\rangle_{-1,\rho(t)}\right)dt \\ &= \mathcal{E}(\rho(T)) - \mathcal{E}(\rho(0)) + \frac{1}{2}\int_0^T \left(\|\frac{\partial\rho}{\partial t}\|^2_{-1,\rho(t)} + \|\nabla\mathcal{E}(\rho(t))\|^2_{-1,\rho(t)}\right)dt. \end{aligned} \tag{47}$$

The rate functional is nothing but the left hand side of (42). We see that the rate functional in the large deviation principle is directly related to the weak formulation of the gradient flows.

3.3 Summary

We summarize the main points of the section.

- The Fokker-Planck equation $\partial_t\rho = \mathrm{div}(\rho\nabla\Psi) + \Delta\rho$ is a Wasserstein gradient flow of the free energy that can be approximated via the JKO-scheme using the functional $\mathcal{K}_h$.
- The Fokker-Planck equation is the thermodynamic limit of the particle systems

  $$dX_i(t) = -\nabla\Psi(X_i(t))dt + \sqrt{2}dW_i(t),$$

  and the empirical process satisfies a large-deviation principle with a rate functional J_h.
- The two functional $\mathcal{K}_h$ and J_h are equivalent in the Gamma-convergence sense: $\mathcal{K}_h \approx \frac{1}{2}J_h$ as $h \downarrow 0$.
- The connection between Wasserstein gradient flow and large-deviation principle can also be seen in the continuous setting.

4 GENERIC and large deviation principle

In this section, we generalise the result in [ADPZ11] to the Kramers equation that consists of both conservative (Hamiltonian flows) and dissipative (gradient flow) effects. We also discuss about a more general class of evolution equations, the GENERIC framework [DPZ13b].

4.1 Langevin dynamics and Kramers equation

Langevin dynamics is an important model in molecular dynamics and statistical physics which used to describe the dynamics of a molecular system. It is obtained as an application of Newtwon's second law where there are three forces acting on the system: an external force, a friction and a random force.

$$m\ddot{Q}(t) = -\nabla V(Q(t)) - \gamma \dot{Q}(t) + \sqrt{2\gamma k_B T} R(t),$$

Here $Q(t)$ is the position of all atoms at time t, m: mass, V: external potential, γ friction coefficient, k_B: Boltzmann constant, T: absolute temperature, $R(t)$ Gaussian noise. More precisely, the system should be written as a stochastic differential equation

$$\begin{aligned} dQ(t) &= \frac{P(t)}{m} dt, \\ dP(t) &= -\nabla V(Q(t)) dt - \gamma \frac{P(t)}{m} dt - \sqrt{2\gamma k_B T} dW(t). \end{aligned}$$

Kramers' equation is the evolution equation for the distribution ρ_t at time t, position q and momentum p

$$\partial_t \rho_t = -\operatorname{div}(\rho_t JH) + \operatorname{div}_p(\rho_t \frac{p}{m}) + \gamma k_B T \Delta_p \rho_t.$$

Properties of Kramers equation:

- Not a Hamiltonian, neither a gradient flow. But a combination of them.
- A typical example of a GENERIC (General Equation for Non-Equilibrium Reversible-Irreversible Coupling)
$$\dot{z} = L(z) \frac{\delta E(z)}{\delta z} + M(z) \frac{\delta S(z)}{\delta z}.$$

A GENERIC equation (General Equation for Non-Equilibrium Reversible-Irreversible Coupling [Ött05]) for an unknown z in a state space Z is a mixture of both reversible and dissipative dynamics:

$$\partial_t \mathsf{z} = \mathsf{L}\, \mathsf{dE} + \mathsf{M}\, \mathsf{dS}. \tag{48}$$

Here

- $\mathsf{E}, \mathsf{S} \colon \mathsf{Z} \to \mathbf{R}$ are interpreted as energy and entropy functionals,
- dE, dS are appropriate derivatives of E and S (such as either the Fréchet derivative or a gradient with respect to some inner product);
- $\mathsf{L} = \mathsf{L}(\mathsf{z})$ is for each z an antisymmetric operator satisfying the Jacobi identity
$$\{\{\mathsf{F}_1, \mathsf{F}_2\}_\mathsf{L}, \mathsf{F}_3\}_\mathsf{L} + \{\{\mathsf{F}_2, \mathsf{F}_3\}_\mathsf{L}, \mathsf{F}_1\}_\mathsf{L} + \{\{\mathsf{F}_3, \mathsf{F}_1\}_\mathsf{L}, \mathsf{F}_2\}_\mathsf{L} = 0, \tag{49}$$
for all functions $\mathsf{F}_i \colon \mathsf{Z} \to \mathbf{R},\ i = 1, 2, 3$, where the Poisson bracket $\{\cdot, \cdot\}_\mathsf{L}$ is defined via
$$\{\mathsf{F}, \mathsf{G}\}_\mathsf{L} := \mathsf{dF} \cdot \mathsf{L}\, \mathsf{dG} \tag{50}$$

- $\mathsf{M} = \mathsf{M}(\mathsf{z})$ is symmetric and positive semidefinite.

Moreover, the building blocks $\{\mathsf{L}, \mathsf{M}, \mathsf{E}, \mathsf{S}\}$ are required to fulfill the *degeneracy conditions*: for all $\mathsf{z} \in \mathsf{Z}$,

$$\mathsf{L}\,\mathsf{dS} = 0, \quad \mathsf{M}\,\mathsf{dE} = 0. \tag{51}$$

As a consequence of these properties, energy is conserved along a solution, and entropy is non-decreasing:

$$\begin{aligned}
\frac{d\mathsf{E}(\mathsf{z}(t))}{dt} &= \mathsf{dE} \cdot \frac{d\mathsf{z}}{dt} = \mathsf{dE} \cdot (\mathsf{L}\,\mathsf{dE} + \mathsf{M}\,\mathsf{dS}) = 0, \\
\frac{d\mathsf{S}(\mathsf{z}(t))}{dt} &= \mathsf{dS} \cdot \frac{d\mathsf{z}}{dt} = \mathsf{dS} \cdot (\mathsf{L}\,\mathsf{dE} + \mathsf{M}\,\mathsf{dS}) = \mathsf{dS} \cdot \mathsf{M}\,\mathsf{dS} \geq 0.
\end{aligned}$$

A GENERIC system is then fully characterized by $\{\mathsf{Z}, \mathsf{E}, \mathsf{S}, \mathsf{L}, \mathsf{M}\}$. Note that a gradient flow is a special case of GENERIC when $\mathsf{E} = 0$.

Can we derive GENERIC structure from microscopic particle models?

In this section, we will generalise the result in Chapter 1 to Kramers equation showing that the GENERIC structure of Kramers equation arises from large-deviation principle of a particle model.

4.2 Particle model

The particle system is constructed as follows. For $i = 1, \dots, n$

$$\begin{aligned}
dQ_i(t) &= \frac{P_i(t)}{m} dt, \\
dP_i(t) &= -\nabla V(Q_i(t))dt - \gamma \frac{P_i(t)}{m} dt - \sqrt{2\gamma k_B T} dW_i(t).
\end{aligned}$$

We consider the empirical process

$$\rho^n(t, dq, dp) = \frac{1}{n} \sum_{i=1}^{n} \delta_{(Q_i(t), P_i(t))}(dq, dp).$$

The above SDE models a system of particles in interaction with a heat bath, and this interaction causes fluctuations of the natural energy (the Hamiltonian) of the particle system,

$$H_n(Q_1, \dots, Q_n, P_1, \dots, P_n) := \frac{1}{n} \sum_{i=1}^{n} \left[\frac{P_i^2}{2m} + V(Q_i) \right] + \frac{1}{2n^2} \sum_{i,j=1}^{n} \psi(Q_i - Q_j). \tag{52}$$

Indeed, using Itô's lemma the derivative of the expression above is

$$-\frac{1}{n} \sum_{i=1}^{n} \left[\frac{\gamma}{m^2} P_i^2 \, dt - \frac{\gamma \theta d}{m} \, dt + \frac{\sqrt{2\gamma\theta}}{m} P_i \, dW_i \right],$$

which has no reason to vanish. We add a single scalar unknown e_n and define its evolution by the negative of the above, leading to the extended particle system

$$dQ_i = \frac{P_i}{m}\,dt, \tag{53a}$$

$$dP_i = -\nabla V(Q_i)\,dt - \sum_{j=1}^{n} \nabla\psi(Q_i - Q_j) - \frac{\gamma}{m}P_i\,dt + \sqrt{2\gamma\theta}\,dW_i, \tag{53b}$$

$$de_n = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\gamma}{m^2}P_i^2\,dt - \frac{\gamma\theta d}{m}\,dt + \frac{\sqrt{2\gamma\theta}}{m}P_i\,dW_i\right], \tag{53c}$$

with which $H_n + e_n$ becomes deterministically constant. Note that e_n can be interpreted as the energy of the heat bath; the flow of energy between the particle system and the heat bath is described by the flow of energy between H_n and e_n.

4.3 From LDP to Kramers' equation

In [DPZ13b], we show that

Theorem 4.1.

(1) The particle system above satisfies a large-deviation principle,

(2) The GENRERIC structure of the extended Kramers equation can be derived from the large-deviation rate functional.

4.4 Summary

- Connection between PDEs and stochastic processes via LDP,
- The rate functional plays a key role: Finding limits of PDEs is equivalent to that of rate functionals,
- Variational formulation of the rate functional gives rise to variational technique,
- Multi-scale analysis, coarse-graining are performed simultaneously,
- Both qualitative and quantitative multi-scale analysis,
- Potential use for a large class of PDEs.

5 Multi-scale analysis of PDEs

In this section, we discuss about a variational technique introduced in [DLPS16, DLP$^+$16] to study multi-scale analysis of PDEs. This variational technique is based on the connection between PDEs and large-deviation principle obtained in previous sections. The advantage of this connection is that it allows us to study multi-scale analysis of non-dissipative systems that existing methods could not treat. An abstract framework will be introduced first. Then several examples, including the overdamped limit of the Kramers equation and the small noise limit of a perturbed Hamiltonian, will be illustrated.

5.1 General framework

Suppose that $\rho^\varepsilon\colon [0,T]\to\mathcal{P}(\mathcal{X}) \quad (\mathcal{X}:=\mathbf{R}^N)$, solves the ε-dependent problem

$$(\mathsf{P}_\varepsilon): \qquad \begin{cases}\partial_t\rho^\varepsilon=\mathcal{L}_\varepsilon^*\,\rho^\varepsilon,\\ \rho^\varepsilon(0)=\rho_0^\varepsilon.\end{cases} \tag{54}$$

The aim is to derive an ε-independent problem P that can be considered as an approximation (in a suitable sense) of (P_ε) as $\varepsilon\to 0$,

$$(\mathsf{P}): \qquad \begin{cases}\partial_t\rho=\mathcal{L}^*\,\rho,\\ \rho(0)=\rho_0.\end{cases} \tag{55}$$

Here $\rho\colon[0,T]\to\mathcal{P}(\mathcal{X}_0)$, where $\mathcal{X}_0$ is some Euclidean space.

Coarse-graining is a technique for such purpose. It consists of two steps. The first one is to transform the problem (P_ε) to a *coarse-grained problem* $(\hat{\mathsf{P}}^\varepsilon)$ defined on $\mathcal{P}(\mathcal{Y})$, where $\mathcal{Y}$ is some *coarse-grained Euclidean space*, via a *coarse-graining map* $\Pi_\varepsilon\colon\mathcal{X}\to\mathcal{Y}$. The coarse-grained space is often of dimension less than the original space and as a consequence the coarse-grained map is non-injective. The coarse-grained problem $(\hat{\mathsf{P}}^\varepsilon)$ describes the evolution of the coarse-grained profile $\hat\rho^\varepsilon$ which is the push-forward of ρ^ε under Π_ε, $\hat\rho^\varepsilon=\Pi_{\varepsilon\#}\rho^\varepsilon\colon[0,T]\to\mathcal{P}(\mathcal{Y})$,

$$(\hat{\mathsf{P}}^\varepsilon): \qquad \begin{cases}\partial_t\hat\rho^\varepsilon=\hat{\mathcal{L}}_\varepsilon^*(\rho^\varepsilon)\,\hat\rho^\varepsilon,\\ \hat\rho^\varepsilon(0)=\hat\rho_0^\varepsilon.\end{cases} \tag{56}$$

Note that the coarse-grained generator $\hat{\mathcal{L}}_\varepsilon^*(\rho^\varepsilon)$ depends on ρ^ε, therefore it is not Markovian in general.

The second step is to derive (P) from $(\hat{\mathsf{P}}^\varepsilon)$. The success of the technique relies on whether one can define an appropriate coarse-grained problem. Usually one also has to rescale the temporal and/or the spatial variables appropriately depending on the effects that one wishes to observe.

5.2 Coarse-graining from large-deviation principle

For fixed ε the equation (P_ε) can be derived from the rate functional of the large deviation principle of the empirical process of an underlying particle system X_i^ε. More precisely,

$$\rho^\varepsilon \text{ is a solution to } (\mathsf{P}_\varepsilon) \text{ iff } I^\varepsilon(\rho^\varepsilon) = 0,$$

where the rate functional $I^\varepsilon(\rho^\varepsilon)$ is given by

$$I^\varepsilon(\rho^\varepsilon) = \sup_{f \in C_c^\infty(\mathbf{R}\times\mathcal{X})} G^\varepsilon(\rho^\varepsilon, f). \tag{57}$$

The functional $G^\varepsilon(\rho^\varepsilon, f)$ has the following form

$$G^\varepsilon(\rho^\varepsilon, f) = \int_{\mathcal{X}} \big[f_T\, d\rho_T^\varepsilon - f_0\, d\rho_0^\varepsilon\big] - \int_0^T \int_{\mathcal{X}} \big[(\partial_t + \mathcal{L}_\varepsilon) f_t\big] d\rho_t^\varepsilon dt - \frac{1}{2}\int_0^T \int_{\mathcal{X}} A\nabla f_t \cdot \nabla f_t \, d\rho_t^\varepsilon \, dt,$$

where A is the diffusion matrix. In order to study the asymptotic behavior of ρ^ε, we study Gamma-convergence of the functional I^ε instead. If one is only interested in convergence of the solutions, one only needs to prove the liminf inequality in the Gamma-convergence provided that the limiting functional is non-negative. In this chapter, we introduce a new method for coarse-graining using the rate functional. The core idea of our method can be summarized in the following four steps.

Step 1. **Choose a special class of test functions:** By taking $f = g \circ \Pi_\varepsilon$, where $g \in C_c^\infty(\mathbf{R} \times \mathcal{Y})$, we obtain

$$I^\varepsilon(\rho^\varepsilon) \geq \sup_{g \in C_c^\infty(\mathbf{R}\times\mathcal{Y})} G^\varepsilon(\rho^\varepsilon, g \circ \Pi_\varepsilon). \tag{58}$$

Note that $g \circ \Pi_\varepsilon$ may not have compact support. Therefore, some approximation argument may be required to ensure that $g \circ \Pi_\varepsilon$ is admissible.

Step 2. **Compactness property for ρ^ε and $\hat{\rho}^\varepsilon$.** In this step, one needs to prove that ρ^ε and $\hat{\rho}^\varepsilon$ possess appropriate compactness property. Assume that $\rho^\varepsilon \xrightarrow{\sigma} \rho$, $\hat{\rho}^\varepsilon \xrightarrow{\hat{\sigma}} \hat{\rho}$, where σ and $\hat{\sigma}$ denote appropriate topologies.

Step 3. **Prove that, up to an $o(1)$ term, $G^\varepsilon(\rho^\varepsilon, g \circ \Pi_\varepsilon)$ depends only on g and the coarse-grained variable $\hat{\rho}^\varepsilon$.** We denote by $\hat{G}^\varepsilon(\hat{\rho}^\varepsilon, g)$ the dominating term in $G^\varepsilon(\rho^\varepsilon, g \circ \Pi_\varepsilon)$. In addition, suppose that we can pass to the limit, with respect to the topology $\hat{\sigma}$, in the functional $\hat{G}^\varepsilon(\hat{\rho}^\varepsilon, g)$ for any fixed g. If this assumptions hold, we may define

$$G(\hat{\rho}, g) := \lim_{\varepsilon\to 0} \hat{G}^\varepsilon(\hat{\rho}^\varepsilon, g) \quad \text{for fixed } g, \tag{59}$$

and also

$$I(\hat{\rho}) := \sup_g G(\hat{\rho}, g). \tag{60}$$

Step 4. **Derive the limiting system** as the law $\hat{\rho} \in C([0,T],\mathcal{P}(X_0))$ that uniquely satisfies $I(\hat{\rho}) = 0$.

We now apply this method to derive two limiting systems: the overdamped (high friction) limit of the Kramers equation and the small-noise limit of a perturbed Hamiltonian system.

5.3 From a perturbed Hamiltonian system to diffusion on a graph

We now describe the small-noise limit. We consider the following stochastically perturbed Hamiltonian system and the time is rescaled $t \mapsto t/\varepsilon$),

$$dQ_\varepsilon = \frac{1}{\varepsilon}P_\varepsilon, \tag{61a}$$

$$dP_\varepsilon = -\frac{1}{\varepsilon}\nabla V(Q_\varepsilon) + \sqrt{2}\,dW. \tag{61b}$$

The probability density ρ^ε of $(Q^\varepsilon, P^\varepsilon)$ satisfies the following equation,

$$(\mathsf{P}_\varepsilon) \qquad \partial_t\rho^\varepsilon = -\frac{1}{\varepsilon}\operatorname{div}(\rho^\varepsilon J\nabla H) + \Delta_p\rho^\varepsilon,$$

where $H(q,p) = \frac{p^2}{2} + V(q)$. The asymptotic behavior of this equation as $\varepsilon \downarrow 0$ was first studied by Freidlin and Wentzell [FW94]. They showed that the limiting system can be described as a diffusion on a graph: over $O(\varepsilon)$ time the solution follows level sets of H, while at $O(1)$ time scale, it performs a biased Brownian motion between level sets.

In this section, we re-prove this result as an illustration of our method. The associated rate functional is as follows

$$I^\varepsilon(\rho^\varepsilon) = \sup_{f\in C_c^\infty(\mathbf{R}\times\mathbf{R}^2)} \left\{ \int_{\mathbf{R}^{2d}} \left[f_T\, d\rho_T^\varepsilon - f_0\, d\rho_0^\varepsilon\right] - \int_0^T \int_{\mathbf{R}^2} \left[\partial_t f + \frac{1}{\varepsilon}J\nabla H\cdot\nabla f_t + \Delta_p f_t\right] d\rho_t^\varepsilon dt \right.$$
$$\left. -\frac{1}{2}\int_0^T\int_{\mathbf{R}^2} |\nabla_p f_t|^2]\, d\rho_t^\varepsilon dt \right\}. \tag{62}$$

We now discuss the simplest case: $d = 1$ and V is a single-well potential (i.e. strictly convex). In this case, the coarse-graining map is the Hamiltonian.

Theorem 5.1. *[DLPS16] Assume that*

(S1) The rate functional and the initial data are uniformly bounded,

$$\sup_{\varepsilon>0}\left[\int H\rho_0^\varepsilon + I^\varepsilon(\rho^\varepsilon)\right] < C. \tag{63}$$

(S2) V is strictly convex, bounded from below and satisfies $\lim_{|q|\to\infty} V(q) = \infty$. Without loss of generality, we assume that $V \geq 0$.

(S3) (Growth conditions on H) There exist constant C such that

$$\max\{|\nabla H|, |\Delta H|\} \leq C(1+H). \tag{64}$$

Then the following hold

(1) (compactness properties) ρ_t^ε and the push-forward $\hat{\rho}^\varepsilon := H_\# \rho^\varepsilon$ satisfy,

$$\sup_{t\in[0,T]} \sup_{\varepsilon>0} \int_{\mathbf{R}^2} H\rho_t^\varepsilon < C, \quad \textit{for some } C > 0, \tag{65}$$

and

$$\hat{\rho}^\varepsilon \longrightarrow \hat{\rho} \quad \textit{in } C([0,T], \mathcal{P}(\mathbf{R})) \textit{ for some } \hat{\rho}. \tag{66}$$

(2) (local equilibrium property) $\rho_t(dx)$ is "constant on level sets" in the sense that,

$$\rho_t(dx) = \hat{\rho}_t(H(x)) \frac{1}{T(H(x))} dx, \tag{67}$$

where T is defined in (69).

(3) (liminf inequality) I^ε satisfies the following liminf-inequality

$$\liminf_{\varepsilon\to 0} I^\varepsilon(\rho^\varepsilon) \geq I(\hat{\rho}), \tag{68}$$

where

$$I(\hat{\rho}) = \sup_{g\in C_c^\infty(\mathbf{R}\times\mathbf{R})} \left[\int_{\mathbf{R}} g_T d\hat{\rho}_T - \int_{\mathbf{R}} g_0 d\hat{\rho}_0 - \int_0^T \int_{\mathbf{R}} \left(\partial_t g(h) + b(h)g'(h) + a(h)g''(h) + \frac{1}{2}a(h)(g'(h))^2 \right) \hat{\rho}_t(dh)\, dt \right],$$

with

$$T(h) = \int_{H^{-1}(H(h))} \frac{1}{|\nabla H(x)|} \mathscr{H}^1(dx), \quad (\mathscr{H}^1 \textit{ is the 1-d Hausdorff measure}), \tag{69}$$

$$a(h) = \frac{1}{T(h)} \int_{H^{-1}(h)} \frac{|\nabla_p H(x)|^2}{|\nabla H(x)|} \mathscr{H}^1(dx), \tag{70}$$

$$b(h) = \frac{1}{T(h)} \int_{H^{-1}(h)} \frac{\Delta_p H(x)}{|\nabla H(x)|} \mathscr{H}^1(dx). \tag{71}$$

(4) (The limiting system) The limiting system can be written as

$$\partial_t \hat{\rho} = \partial_h(a(h)\partial_h \hat{\rho}) - \partial_h(b(h)\hat{\rho}). \tag{72}$$

Proof. We sketch the main steps of the proof, details can be found in [DLPS16].

Since $I^\varepsilon(\rho^\varepsilon) < \infty$ there exists $h^\varepsilon_t \in \mathbb{L}^2(0,T;\mathbb{L}^2_\nabla(\rho^\varepsilon_t))$ such that

$$\partial_t \rho^\varepsilon_t = -\frac{1}{\varepsilon}\operatorname{div}(\rho^\varepsilon J\nabla H) + \Delta_p \rho^\varepsilon - \operatorname{div}_p(h^\varepsilon_t \rho^\varepsilon_t).$$

The rate functional $I^\varepsilon(\rho^\varepsilon)$ can be expressed in terms of h^ε as

$$I^\varepsilon(\rho^\varepsilon) = \frac{1}{2}\int_0^T |h^\varepsilon_t|^2 \rho^\varepsilon_t \, dt. \tag{73}$$

Therefore, for $t \in [0,T]$ and $f \in C^2_c(\mathbf{R}^{2d})$, we have

$$\frac{d}{dt}\int_{\mathbf{R}^2} f(x)\rho^\varepsilon_t(x)dx = \int_{\mathbf{R}^2} f(x)\partial_t\rho^\varepsilon_t(x)dx \tag{74}$$

$$= \int_{\mathbf{R}^2}\left(\frac{1}{\varepsilon}J\nabla H\cdot\nabla f + \Delta_p f + \nabla_p f\cdot h^\varepsilon_t\right)\rho^\varepsilon_t. \tag{75}$$

Substituting $f = H$ in (75) we have the following formal calculation,

$$\begin{aligned}\frac{d}{dt}\int_{\mathbf{R}^{2d}} H\rho^\varepsilon_t &= \int_{\mathbf{R}^{2d}}\left(\frac{1}{\varepsilon}J\nabla H\cdot\nabla H + \Delta_p H + \nabla_p H\cdot h^\varepsilon_t\right)\rho^\varepsilon_t \\ &\le \int_{\mathbf{R}^{2d}}\left(\Delta_p H + \frac{1}{2}\left[|\nabla_p H|^2 + |h^\varepsilon_t|^2\right]\right)\rho^\varepsilon_t \\ &\overset{(64),(73)}{\le} C\int_{\mathbf{R}^{2d}}(1+H)\rho^\varepsilon_t + I^\varepsilon(\rho^\varepsilon).\end{aligned}$$

Using (63) and a Gronwall-type estimate, we obtain

$$\int_{\mathbf{R}^{2d}} H\rho^\varepsilon_t < C.$$

To make these calculations rigorous we define for each $m \in \mathbb{N}$, $\psi_m \in C^\infty_c(\mathbf{R})$ with $0 \le \psi_m \le 1$ such that $|\psi'_m| \le \psi_m/m$ and $|\psi''_m| \le \psi_m/m^2$. We make the choice $f_m(x) = H(x)\psi_m(H(x))$, where note that $f_m \in C^2_c(\mathbf{R}^{2d})$. Proceeding as in calculation above and using Gronwall type estimates we arrive at

$$\int_{\mathbf{R}^{2d}} f_m\rho^\varepsilon_t \le C.$$

Using monotone convergence theorem we obtain (65).

To prove (66), we will use [CL12, Theorem 3] which is an extension of the classical compactness results of Simon [Sim86] to the case of semi-normed spaces. The spatial compactness of

$\hat\rho^\varepsilon$ is a direct consequence of (65) and the coercivity of V. To prove the time compactness, we define three spaces

$$X_1 = (\mathcal{M}_+(\mathbf{R}), \|\cdot\|_{1BL}), \quad X_2 = \overline{(\mathcal{M}(\mathbf{R}), \|\cdot\|_{BL})}, \quad X_3 = (C_0^2(\mathbf{R}))^*,$$

where

$$\|\mu\|_{1BL} = \|\mu\|_{BL} + \int |x|\, d\mu, \quad \|\mu\|_{BL} := \sup_{f\in \mathrm{BL}(\mathbf{R}),\ \|f\|_{BL}\le 1} \left\{ \Big| \int f\, d\mu \Big| \right\}.$$

Here $\mathrm{BL}(\mathbf{R})$ denotes the space of bounded Lipschitz functions on $\mathbf{R}$. Note that $\|\cdot\|_{BL}$ metrizes the narrow topology. Then X_1 is a seminormed nonnegative cone in X_2. Moreover, $X_1 \hookrightarrow\hookrightarrow X_2 \hookrightarrow X_3$. Take $\varphi \in C_0^2(\mathbf{R})$, we have

$$\begin{aligned}
\int_{\mathbf{R}^{2d}} \varphi \hat\rho_\tau^\varepsilon \Big|_{\tau=t}^{\tau=t+s} &= \int \varphi(H)\rho_\tau^\varepsilon \Big|_{\tau=t}^{\tau=t+s} \\
&= \int_t^{t+s} \int_{\mathbf{R}^{2d}} \left(\frac{1}{\varepsilon} J\nabla H(x)\cdot \nabla\varphi(H(x)) + \Delta_p \varphi(H(x)) + \nabla_p\varphi(H(x)) h_\tau^\varepsilon \right) \rho_\tau^\varepsilon\, d\tau.
\end{aligned}$$

The first term inside the integral above equals to 0. Using the argument as in the proof of (65), we find that

$$\left| \int_{\mathbf{R}^{2d}} \varphi\hat\rho_{t+s}^\varepsilon - \int_{\mathbf{R}^{2d}} \varphi\hat\rho_t^\varepsilon \right| \le Cs.$$

By [CL12, Theorem 3], $\hat\rho^\varepsilon$ is relatively compact in $C([0,T],\mathcal{P}(\mathbf{R}))$.

Now we prove (67). From (62), we have for every $f\in C_c^\infty(\mathbf{R}\times\mathbf{R}^2)$

$$\begin{aligned}
\int_0^T \int_{\mathbf{R}^2} J\nabla H\cdot\nabla f\, d\rho_t^\varepsilon dt &\le \varepsilon \left[\int_{\mathbf{R}^2} \left[f_0\rho_0^\varepsilon - f_T\rho_T^\varepsilon \right] + \int_0^T \int_{\mathbf{R}^2} \left(\partial_t + \Delta_p f_t + \frac{1}{2}|\nabla_p f_t|^2 \right) d\rho_t^\varepsilon dt + I^\varepsilon(\rho^\varepsilon) \right] \\
&\le C\varepsilon.
\end{aligned}$$

Substituting f by $-f$, we obtain the opposite inequality. This and together with (65) we get

$$\int_0^T \int_{\mathbf{R}^2} J\nabla H\cdot\nabla f\, \rho_t(dx)dt = 0, \quad \text{for all } f\in C_c^\infty(\mathbf{R}\times\mathbf{R}^2). \tag{76}$$

In particular, for each fixed $t\in[0,T]$ and $f\in C_c^2(\mathbf{R}^2)$,

$$\int_{\mathbf{R}^2} J\nabla H(x)\cdot\nabla f(x)\rho_t(dx) = 0. \tag{77}$$

Choosing $f(x) = \zeta(H(x))\psi(x)$, where $\psi\in C_c^2(\mathbf{R}^2)$ is a spatial cutoff function, and applying Disintegration Theorem [AGS08, Theorem 5.3.1], we get

$$\begin{aligned}
0 &= \int_{\mathbf{R}^2} J\nabla H(x)\cdot \Big(\zeta(H(x))\nabla\psi(x) \Big) \rho_t(dx) \\
&= \int_{\mathbf{R}} \zeta(h)\hat\rho_t(dh) \int_{H^{-1}(h)} \nabla\psi(x)\cdot \frac{J\nabla H(x)}{|\nabla H(x)|} |\nabla H(x)| \tilde\rho_t(dx|h).
\end{aligned}$$

We denote $\tau := \frac{J\nabla H}{|\nabla H|}$. Since $\tau \perp \nabla H, |\tau| = 1$, τ is the tangential vector of the level set $H^{-1}(h)$. Since the choice of ζ is arbitrary, we conclude

$$\text{for } \hat{\rho}_t\text{-a.e. } h \in \mathbf{R}, \int_{H^{-1}(h)} |\nabla H(x)| \partial_\tau \psi(x) \tilde{\rho}_t(dx|h) = 0. \tag{78}$$

Since ψ is arbitrary, the above equality implies that $|\nabla H| \tilde{\rho}_t(dx|h)$ is constant on $H^{-1}(h)$. This means that $\tilde{\rho}_t(dx|h) = \frac{c(h)}{|\nabla H|}$ where $c(h)$ depends only on h but not x. Since $\tilde{\rho}_t(dx|h)$ is a probability measure on $H^{-1}(h)$, the function $c(h)$ can be found by

$$1 = c(h) \int_{H^{-1}(h)} \frac{1}{|\nabla H(x)|} \mathscr{H}^1(dx),$$

or equivalently,

$$c(h) = \frac{1}{T(h)},$$

where $T(h)$ is defined in (69).

As a consequence, we get

$$\tilde{\rho}_t(dx|h) = \frac{\mathscr{H}^1(dx)}{T(h)|\nabla H(x)|}, \quad \text{for } \hat{\rho}_t\text{-a.e. } h \in \mathbf{R}. \tag{79}$$

To obtain (67) we use the following co-area formula. The proof can be found in [MSZ03].

Lemma 5.2 (co-area formula for Sobolev mappings)**.** *Let $H \in W^{1,p}_{loc}(\Omega, \mathbf{R})$ where $\Omega \subset \mathbf{R}^{2d}$ is an open subset such that $\nabla H(x) \neq 0$ a.e. and $g \in L^1(\mathbf{R}^{2d})$. Then,*

$$\int\limits_\Omega g(x) dx = \int\limits_{\mathbf{R}} dh \Big(\int\limits_{H^{-1}(h) \cap \Omega} \frac{g(x)}{|\nabla H(x)|} \mathcal{H}^{2d-1}(dx) \Big). \tag{80}$$

Applying this lemma, on one hand, we have

$$\int_{\mathbf{R}^2} f(x) \rho_t(x)\, dx = \int_{\mathbf{R}} dh \int_{H^{-1}(h)} \frac{f(x)\rho_t(x)}{|\nabla H(x)|} \mathscr{H}^1(dx). \tag{81}$$

On the other hand, from (79), we have for any $f \in C^2_c(\mathbf{R}^2)$,

$$\int_{\mathbf{R}^2} f(x) \rho_t(dx) = \int_{\mathbf{R}} \hat{\rho}_t(dh) \int_{H^{-1}(h)} f(x) \tilde{\rho}_t(dx|h) = \int_{\mathbf{R}} \frac{\hat{\rho}_t(dh)}{T(h)} \int_{H^{-1}(h)} \frac{f(x)}{|\nabla H(x)|} \mathcal{H}^1(dx). \tag{82}$$

Comparing (81) and (82) gives (67).

Next, we prove (68).

We take $f(t,x) = g(t, H(x))$ for $g \in C_c^\infty(\mathbf{R}\times\mathbf{R})$ and pass to the limit in the rate functional (62). We compute derivatives of f,

$$\partial_t f = \partial_t g(H), \quad \nabla f = g'(H)\nabla H, \quad \nabla_p f = g'(H)\nabla_p H, \quad \Delta_p f = g''(H)|\nabla_p H|^2 + g'(H)\Delta_p g.$$

The first three terms are straightforward,

$$\int_{\mathbf{R}^2} \big[f_T\, d\rho_T^\varepsilon - f_0\, d\rho_0^\varepsilon\big] = \int_{\mathbf{R}^2} \big[g_T \circ H\, d\rho_T^\varepsilon - f_0 \circ H\, d\rho_0^\varepsilon\big] = \int_{\mathbf{R}^2} \big[g_T\, d\hat\rho_T^\varepsilon - f_0\, d\hat\rho_0^\varepsilon\big],$$

$$\int_0^T \int_{\mathbf{R}^2} \partial_t f\, d\rho_t^\varepsilon dt = \int_0^T \int_{\mathbf{R}} \partial_t g\, d\hat\rho_t^\varepsilon dt.$$

The fourth term vanishes since J is anti-symmetric,

$$\int_0^T \int_{\mathbf{R}^2} J\nabla H \cdot \nabla f_t\, d\rho_t^\varepsilon dt = \int_0^T \int_{\mathbf{R}^2} g'(H) J\nabla H \cdot \nabla H\, d\rho_t^\varepsilon dt = 0.$$

To transform the last two terms we need to use (67) and (82). We have

$$\begin{aligned}\liminf_{\varepsilon\to 0} \int_0^T \int_{\mathbf{R}^2} \Delta_p f d\rho_t^\varepsilon dt &= \liminf_{\varepsilon\to 0} \int_0^T \int_{\mathbf{R}^2} \big[g''(H(x))|\nabla_p H(x)|^2 + g'(H(x))\Delta_p H(x)\big]\rho_t^\varepsilon(x) dx dt \\ &= \int_0^T \int_{\mathbf{R}^2} \big[g''(H(x))|\nabla_p H(x)|^2 + g'(H(x))\Delta_p H(x)\big]\rho_t(x) dx dt \\ &\overset{(82)}{=} \int_0^T \int_{\mathbf{R}} \Big(a(h)g''(h) + b(h)g'(h)\Big)\hat\rho_t(dh) dt,\end{aligned} \tag{83}$$

and

$$\begin{aligned}\liminf_{\varepsilon\to 0} \int_0^T \int_{\mathbf{R}^2} |\nabla_p f|^2 d\rho_t^\varepsilon dt &= \liminf_{\varepsilon\to 0} \int_0^T \int_{\mathbf{R}^2} (g'(H(x)))^2 |\nabla_p H|^2 d\rho_t^\varepsilon dt \\ &= \int_0^T \int_{\mathbf{R}^2} (g'(H(x)))^2 |\nabla_p H|^2 d\rho_t dt \\ &\overset{(82)}{=} \int_0^T \int_{\mathbf{R}} a(h)(g'(h))^2 \hat\rho_t(dh) dt,\end{aligned} \tag{84}$$

where $a(h),\ b(h)$ are defined in (70)-(71).

Combining all these terms we have,

$$\begin{aligned} I^\varepsilon(\rho^\varepsilon) \ge \sup_{g\in C_c^\infty(\mathbf{R}\times\mathbf{R})} \bigg[&\int_{\mathbf{R}} g_T d\hat\rho_T - \int_{\mathbf{R}} g_0 d\hat\rho_0 \\ &- \int_0^T \int_{\mathbf{R}} \Big(\partial_t g(h) + b(h)g'(h) + a(h)g''(h) + \frac{1}{2}a(h)(g'(h))^2\Big)\,\hat\rho_t(dh)\, dt\bigg] \\ =:\ & I(\hat\rho).\end{aligned}$$

Note that by choosing $g = 0$, we always have $I(\hat\rho) \ge 0$. The limiting system (72) then follows from the form of the rate functional $I(\hat\rho)$.

□

5.4 From the Kramers equation to the Fokker-Planck equation

In this section, we derive the Fokker-Planck equation as the overdamped (high friction) limit of the Kramers equation. The overdamped limit was derived formally first in [Kra40] and has been extensively studied in the literature from different point of view such as asymptotic expansions or probabilistic methods, see for instance [Nel67, Wil76, GPK12] and references therein. We reprove this result to illustrate our method.

We recall the Kramers equation

$$\partial_t\rho = -\operatorname{div}_q\left(\frac{p}{m}\rho\right) + \operatorname{div}_p\Big(\nabla_q V(q)\rho\Big) + \gamma\left[\operatorname{div}_p\Big(\frac{p}{m}\rho\Big) + \theta\Delta_p\rho(t,q,p)\right], \tag{85}$$

where m, γ, θ are positive constants. For simplicity we set $\theta = 1$. The overdamped limit corresponds to the limit $\gamma \to \infty$ in (85).

Rescaling time appropriately (speeding up by $1/\gamma$) we arrive at

$$\partial_t\rho = -\gamma\operatorname{div}_q\left(\frac{p}{m}\rho\right) + \gamma\operatorname{div}_p\Big(\nabla_q V(q)\rho\Big) + \gamma^2\left[\operatorname{div}_p\Big(\frac{p}{m}\rho\Big) + \Delta_p\rho\right].$$

The large-deviation rate functional associated to this equation is (see (57))

$$\begin{aligned} I^\gamma(\rho) = \sup_{f\in C_c^\infty(\mathbf{R}\times\mathbf{R}^{2d})}\Bigg[\int_{\mathbf{R}^{2d}}(f_T d\rho_T - f_0 d\rho_0) - \int_0^T\int_{\mathbf{R}^{2d}}\Big(\partial_t f + \gamma\frac{p}{m}\cdot\nabla_q f - \gamma\nabla_q V\cdot\nabla_p f - \gamma^2\frac{p}{m}\cdot\nabla_p f \\ + \gamma^2\Delta_p f\Big)d\rho_t dt - \frac{\gamma^2}{2}\int_0^T\int_{\mathbf{R}^{2d}}|\nabla_p f|^2\,d\rho_t dt\Bigg]. \end{aligned} \tag{86}$$

The rate functional can be written in a more general form in terms of the generator $\mathcal{L}$ as,

$$\begin{aligned} I^\gamma(\rho) &= \sup_{f\in C_c^\infty(\mathbf{R}\times\mathbf{R}^{2d})}\Big\{\int_{\mathbf{R}^{2d}}\big[f_T\rho_T - f_0\rho_0\big] - \int_0^T\int_{\mathbf{R}^{2d}}(\partial_t f_t + (J-A)\nabla H\cdot\nabla f_t + \Delta_p f_t \\ &\qquad + \frac{1}{2}|\nabla_p f_t|^2)\,d\rho_t\,dt\Big\} \\ &= \sup_{f\in C_c^\infty(\mathbf{R}\times\mathbf{R}^{2d})}\Big\{\int_{\mathbf{R}^{2d}}\big[f_T\rho_T - f_0\rho_0\big] - \int_0^T\int_{\mathbf{R}^{2d}}(\partial_t f_t + \mathcal{L}f_t + \frac{1}{2}(\nabla f_t)^T A\nabla f_t)\,d\rho_t\,dt\Big\}, \end{aligned} \tag{87}$$

where

$$\mathcal{L}f = (J-A)\nabla H\cdot\nabla f + \operatorname{div}(A\nabla f), \quad J = \gamma\begin{pmatrix}0 & I\\ -I & 0\end{pmatrix}, \quad A = \gamma^2\begin{pmatrix}0 & 0\\ 0 & I\end{pmatrix}.$$

We recall definition of the relative entropy and the relative Fisher information. Let $\mu(dx) = Z^{-1}\exp(-H(x))dx$ be the invariant measure. The relative entropy $\mathscr{H}(\nu|\mu)$ and the relative Fisher information $RF(\nu|\mu)$ of a measure ν with respect to μ are respectively given by

$$\mathscr{H}(\nu|\mu) = \begin{cases} \int_{\mathbf{R}^{2d}} \frac{d\nu}{d\mu}\log\frac{d\nu}{d\mu}\,d\mu & \text{if} \quad \nu \ll \mu, \\ \infty \quad \text{otherwise.} \end{cases} \tag{88}$$

$$RF(\nu|\mu) = \begin{cases} \int_{\mathbf{R}^{2d}} \frac{A\nabla\frac{d\nu}{d\mu}\cdot\nabla\frac{d\nu}{d\mu}}{\frac{d\nu}{d\mu}}\,d\mu & \text{if} \quad d\nu = \nu(x)dx, \nabla\frac{d\nu}{d\mu} \in L^1_{\text{loc}}(\mathbf{R}^{2d}), \\ \infty \quad \text{otherwise.} \end{cases} \tag{89}$$

We define the coarse-graining map as follows,

$$\begin{aligned} \Pi_\gamma \colon \mathbf{R}^{2d} &\to \mathbf{R}^d \\ (q,p) &\mapsto \Pi_\gamma(q,p) = q + \frac{p}{\gamma}. \end{aligned}$$

Theorem 5.3. *[DLPS16] Assume that*

(B1) The rate functional and the initial data are uniformly bounded

$$\sup_{\gamma>0}\left[I^\gamma(\rho^\gamma) + \int_{\mathbf{R}^{2d}} \rho_0^\gamma \log\rho_0^\gamma + H\rho_0^\gamma\right] < \infty.$$

(B2) V is bounded from below and satisfies $\lim_{|q|\to\infty} = \infty$ and $\|\nabla^2 V\|_\infty < \infty$. Without loss of generality, we assume $V \geq 0$.

Then the following hold

1. *(compactness properties) ρ^γ and the push-forward $\hat{\rho}^\gamma := \Pi_{\gamma\#}\rho^\gamma$ satisfy,*

$$\sup_{t\in[0,T]}\sup_{\gamma>0}\int_{\mathbf{R}^{2d}} H(q,p)\rho_t^\gamma(dqdp) < \infty, \tag{90}$$

and

$$\hat{\rho}^\gamma \to \sigma \quad \text{in} \quad C([0,T],\mathcal{P}(\mathbf{R}^d)) \quad \text{for some} \quad \sigma. \tag{91}$$

2. *(local equilibrium statement)*

$$\rho^\gamma \rightharpoonup Z^{-1}\exp\left(-\frac{p^2}{2m}\right)\sigma \quad \text{in} \quad \mathcal{P}([0,T]\times\mathbf{R}^{2d}). \tag{92}$$

3. *(liminf inequality) $I^\gamma(\rho^\gamma)$ satisfies the following liminf inequality*

$$\liminf_{\gamma\to\infty} I^\gamma(\rho^\gamma) \geq I(\sigma), \tag{93}$$

where

$$I(\sigma) := \sup_{g\in C_c^\infty(\mathbf{R}\times\mathbf{R}^d)} \Bigg[\int\limits_{\mathbf{R}^d} g_T d\sigma_T - \int\limits_{\mathbf{R}^d} g_0 d\sigma_0 - \int\limits_0^T\int\limits_{\mathbf{R}^d}(\partial_t g - \nabla V\cdot\nabla g + \Delta g)d\sigma_t dt \\ -\frac{1}{2}\int\limits_0^T\int\limits_{\mathbf{R}^d} |\nabla g|^2\, d\sigma_t dt\Bigg].$$

4. (the limiting system) σ satisfies the Fokker-Planck equation

$$\partial_t\sigma = \operatorname{div}(\nabla V\sigma) + \Delta\sigma. \tag{94}$$

We now show the main steps of the proof; all details can be found in [DLPS16]. A crucial step is to establish a priori estimate on the relative entropy and the relative Fisher information.

A priori estimate (upper bound for the relative entropy and the relative Fisher information).

Claim 1: It holds that

$$\mathscr{H}(\rho_T^\gamma|\mu) + \frac{1}{2}\int_0^T RF(\rho_t^\gamma|\mu)\,dt \le I(\rho^\gamma) + \mathscr{H}(\rho_0^\gamma|\mu). \tag{95}$$

As a consequence,

$$\sup_{t\in[0,T]}\sup_{\gamma>0}\int_{\mathbf{R}^{2d}} H(q,p)\rho_t^\gamma(dqdp) < \infty. \tag{96}$$

We use the following variational formulation for the relative entropy and the Fisher information [FK06, Chapter 9 and Appendix D6],

$$\mathscr{H}(\rho|\mu) = \sup_{\psi\in C_c^\infty(\mathbf{R}^{2d})}\Big\{\int\psi\rho - \log\int e^\psi\,d\mu\Big\},$$
$$\frac{1}{2}RF(\rho|\mu) = \sup_{\varphi\in C_c^\infty(\mathbf{R}^{2d})}\Big\{\int\Big(-\operatorname{div}(A\nabla\varphi) + A\nabla\varphi\cdot\nabla H - \frac{1}{2}(\nabla\varphi)^T A\nabla\varphi\Big)\rho\Big\}.$$

Given φ and ψ, we take f such that

$$\partial_t f_t + \mathcal{L}f_t + \frac{1}{2}(\nabla f_t)^T A\nabla f_t = \operatorname{div}(A\nabla\varphi) - A\nabla\varphi\cdot\nabla H + \frac{1}{2}(\nabla\varphi)^T A\nabla\varphi, \quad f_T = \psi. \tag{97}$$

For the Kramers, $\mathcal{L}f = -\gamma\frac{p}{m}\cdot\nabla_q f + \gamma\nabla V(q)\cdot\nabla_p f - \gamma^2\frac{p}{m}\cdot\nabla_p f + \gamma^2\Delta_p f$ and the equation above becomes

$$\partial_t f - \gamma\frac{p}{m}\cdot\nabla_q f + \gamma\nabla V(q)\cdot\nabla_p f - \gamma^2\frac{p}{m}\cdot\nabla_p f + \gamma^2\Delta_p f + \frac{\gamma^2}{2}|\nabla_p f|^2 \\ = \gamma^2\Delta_p\varphi - \gamma^2\nabla_p\varphi\cdot\frac{p}{m} + \frac{\gamma^2}{2}|\nabla_p\varphi|^2. \tag{98}$$

Set $F = \exp(f/2)$, then $f = 2\log F$. F satisfies the following equation

$$\begin{cases}\partial_t F - \gamma\frac{p}{m}\cdot\nabla_q F + \gamma\nabla V(q)\cdot\nabla_p F - \gamma^2\frac{p}{m}\cdot\nabla_p F + \gamma^2\Delta_p F = \frac{\gamma^2 F}{4}\left[|\nabla_p\varphi|^2 + 2\Delta_p\varphi\right],\\ F_T = \exp(\psi/2).\end{cases} \tag{99}$$

Assumption 5.1. Assume that we can take f as a test function in the variational formulation of the rate functional (57).

Then we have

$$\begin{aligned}
&\mathscr{H}(\rho_T|\mu) + \frac{1}{2}\int_0^T RF(\rho_t|\mu)\,dt\\
&\quad= \sup_{\psi,\varphi}\Big\{\int\rho_T\psi - \log\int e^{\psi}\,d\mu - \int_0^T\int_{\mathbf{R}^{2d}}\Big[\mathrm{div}(A\nabla\varphi) - A\nabla\varphi\cdot\nabla H + \frac{1}{2}(\nabla\varphi)^T A\nabla\varphi\Big]\rho_t\,dt\Big\}\\
&\quad\overset{(97)}{=} \sup_{\psi,\varphi}\Big\{\int\rho_T f_T - \int_0^T\int_{\mathbf{R}^{2d}}\Big[\partial_t f_t + \mathcal{L}f_t + \frac{1}{2}(\nabla f_t)^T A\nabla f_t\Big]\rho_t\,dt - \log\int e^{\psi}\,d\mu\Big\}\\
&\quad\le I(\rho) + \sup_{\psi,\varphi}\Big\{\int_{\mathbf{R}^{2d}} f_0\rho_0 - \log\int_{\mathbf{R}^{2d}} e^{f_T}\,d\mu\Big\}\\
&\quad\le I(\rho) + \mathscr{H}(\rho_0|\mu) + \sup_{\psi,\varphi}\Big\{\log\frac{\int_{\mathbf{R}^{2d}} e^{f_0}\,d\mu}{\int_{\mathbf{R}^{2d}} e^{f_T}\,d\mu}\Big\}.
\end{aligned}$$

Now we prove that $\int_{\mathbf{R}^{2d}} e^{f_0}\,d\mu \le \int_{\mathbf{R}^{2d}} e^{f_T}\,d\mu$. This will be proven if we show that $t\mapsto \int_{\mathbf{R}^{2d}} e^{f_t}\,d\mu$ is an increasing function. Indeed, we compute its derivative with respect to time,

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbf{R}^{2d}} e^{f_t}\,d\mu\\
&\quad\overset{(97)}{=}\int_{\mathbf{R}^{2d}}\Big(-\mathcal{L}f_t - \frac{1}{2}(\nabla f_t)^T A\nabla f_t + \mathrm{div}(A\nabla\varphi) - A\nabla\varphi\cdot\nabla H + \frac{1}{2}(\nabla\varphi)^T A\nabla\varphi\Big)e^{f_t}\,d\mu.
\end{aligned}$$

Since

$$\begin{aligned}
-\int_{\mathbf{R}^{2d}} e^{f_t-H}\mathcal{L}f_t &= \int_{\mathbf{R}^{2d}}[-b(x)\cdot\nabla f_t - \mathrm{div}(A\nabla f_t)]e^{f_t-H}\\
&= \int_{\mathbf{R}^{2d}}(-J+A)\nabla H\cdot\nabla f_t\,e^{f_t-H} + \int_{\mathbf{R}^{2d}} A\nabla f_t\cdot\nabla(f_t-H)\,e^{f_t-H}\\
&= -\int_{\mathbf{R}^{2d}} e^{-H}J\nabla H\cdot\nabla(e^{f_t}) + \int_{\mathbf{R}^{2d}}(\nabla f_t)^T A\nabla f_t\,e^{f_t-H}\\
&= \int_{\mathbf{R}^{2d}} e^{f_t}\,\mathrm{div}[e^{-H}J\nabla H] + \int_{\mathbf{R}^{2d}}(\nabla f_t)^T A\nabla f_t\,e^{f_t-H}\\
&= \int_{\mathbf{R}^{2d}} A\nabla f_t\cdot\nabla f_t\,e^{f_t-H}\quad\text{(since } J \text{ is anti-symmetric)},
\end{aligned}$$

and

$$\int_{\mathbf{R}^{2d}} \operatorname{div}(A\nabla\varphi)\, e^{f_t - H} = -\int_{\mathbf{R}^{2d}} A\nabla\varphi \cdot \nabla(f_t - H)\, e^{f_t - H},$$

it follows that

$$\begin{aligned}\frac{d}{dt}\int_{\mathbf{R}^{2d}} e^{f_t}\, d\mu &= \int_{\mathbf{R}^{2d}} \left[\frac{1}{2}A\nabla f_t \cdot \nabla f_t + \frac{1}{2}A\nabla\varphi\cdot\nabla\varphi - A\nabla\varphi\cdot\nabla f_t\right] e^{f_t - H}\\ &= \frac{1}{2}\int_{\mathbf{R}^{2d}} A\nabla(f_t-\varphi)\cdot\nabla(f_t-\varphi) \geq 0.\end{aligned}$$

Therefore, $t \mapsto \int e^{f_t}\mu$ is an increasing function. Thus we obtain

$$\int e^{f_0}\, d\mu \leq \int e^{f_T}\, d\mu.$$

The assertion (95) then follows. It is more helpful to use its explicit form as follows

$$\begin{aligned}&\sup_{\gamma>0}\left\{\sup_{t\in[0,T]} \mathscr{H}(\rho_t^\gamma\big|\mu) + \frac{1}{2}\gamma^2\int_0^T\int_{\mathbf{R}^{2d}} \frac{1}{\rho_\gamma^t}\left|\frac{p}{m}\rho_t^\gamma + \nabla_p\rho_t^\gamma\right|^2 dqdpdt\right\}\\ &\quad\leq \sup_{\gamma>0} I^\gamma(\rho_t^\gamma) + \mathscr{H}(\rho_0\big|\mu) < C. \qquad (100)\end{aligned}$$

Now we prove (65). It follows from the above estimate that

$$\sup_{t\in[0,T]}\sup_{\gamma>0}\int_{\mathbf{R}^{2d}} \rho_t^\gamma \log\rho_t^\gamma\, dqdp + \int_{\mathbf{R}^{2d}} H(q,p)\rho_t^\gamma\, dqdp < \infty. \qquad (101)$$

Let $0<\alpha<1$. We have

$$0 \leq \mathscr{H}(\rho_t^\gamma\big|Z_\alpha^{-1}\exp(-\alpha H)) = \int_{\mathbf{R}^{2d}} \rho_t^\gamma\log\rho_t^\gamma\, dqdp + \alpha\int_{\mathbf{R}^{2d}} H(q,p)\rho_t^\gamma\, dqdp + \log Z_\alpha.$$

It implies that

$$\int_{\mathbf{R}^{2d}} \rho_t^\gamma\log\rho_t^\gamma\, dqdp \geq -\alpha\int_{\mathbf{R}^{2d}} H(q,p)\rho_t^\gamma\, dqdp - \log Z_\alpha.$$

Substituting the above inequality into (101), we get

$$\sup_{t\in[0,T]}\sup_{\gamma>0}\int_{\mathbf{R}^{2d}} H(q,p)\rho_t^\gamma\, dqdp < \infty.$$

Verify the conjecture: We now show how the conjecture can be deduced from (95) and (96).

1. Estimate (90) has been already proved in (96). Similarly as in the proof of part (1) of Theorem 5.1, the compactness properties of ρ^γ and $\hat{\rho}^\gamma$ follows directly from (96).

2. The local equilibrium statement is a consequence of the vanishing of the relative Fisher information obtained from (100).

3. Now we prove the liminf inequality (93). In (86) by taking $f(t,q,p) = g(t,\Pi_\gamma(q,p))$, and using

$$\partial_t f = (\partial_t g)\circ \Pi_\gamma, \quad \nabla_q f = (\nabla g)\circ\Pi_\gamma, \quad \nabla_p f = \frac{1}{\gamma}(\nabla g)\circ\Pi_\gamma, \quad \Delta_p f = \frac{1}{\gamma^2}(\Delta g)\circ\Pi_\gamma,$$
$$\nabla V(q) = (\nabla V)\circ\Pi_\gamma + \nabla V(q) - \nabla V\big(q + \frac{1}{\gamma}p\big),$$

we get

$$I^\gamma(\rho^\gamma) \geq \int\limits_{\mathbf{R}^d} g_T d\hat\rho_T^\gamma - \int\limits_{\mathbf{R}^d} g_0 d\hat\rho_0^\gamma - \int\limits_0^T\int\limits_{\mathbf{R}^d} (\partial_t g - \nabla V\cdot\nabla g + \Delta g)d\hat\rho_t^\gamma dt$$
$$- \frac{1}{2}\int\limits_0^T\int\limits_{\mathbf{R}^d} |\nabla g|^2\, d\hat\rho_t^\gamma dt + \int_0^T\int_{\mathbf{R}^{2d}} \left[\nabla V(q) - \nabla V\left(q+\frac{1}{\gamma}p\right)\right]\cdot\nabla g(t,q+\frac{1}{\gamma}p)d\rho_t^\gamma dt. \tag{102}$$

In order to pass to the limit, we need to control the last term in (102). Since

$$\left|\nabla V(q) - \nabla V\left(q+\frac{1}{\gamma}p\right)\right| \leq \frac{1}{\gamma}\|\nabla^2 V\|_\infty |p|, \quad |\nabla g(t,q+\frac{1}{\gamma}p)| \leq \|\nabla g\|_\infty,$$

we have

$$\left|\int\limits_0^T\int\limits_{\mathbf{R}^{2d}} \left[\nabla V(q) - \nabla V\left(q+\frac{1}{\gamma}p\right)\right]\cdot\nabla g(t,q+\frac{1}{\gamma}p)d\rho_t^\gamma dt\right|$$
$$\leq \frac{1}{\gamma}\int\limits_0^T\int\limits_{\mathbf{R}^{2d}} \|\nabla^2 V\|_\infty \,\|\nabla g\|_\infty\, |p|\, d\rho_t^\gamma dt. \tag{103}$$

Due to (96), the right hand side of (103) vanishes as $\gamma\to\infty$. Therefore

$$\liminf_{\gamma\to\infty} I^\gamma(\rho^\gamma) \geq I(\sigma),$$

where

$$I(\sigma) := \sup_{g\in C_c^\infty(\mathbf{R}\times\mathbf{R}^d)} \left[\int\limits_{\mathbf{R}^d} g_T d\sigma_T - \int\limits_{\mathbf{R}^d} g_0 d\sigma_0 - \int\limits_0^T\int\limits_{\mathbf{R}^d} (\partial_t g - \nabla V\cdot\nabla g + \Delta g)d\sigma_t dt \right.$$
$$\left. -\frac{1}{2}\int\limits_0^T\int\limits_{\mathbf{R}^d} |\nabla g|^2\, d\sigma_t dt\right].$$

4. It follows from the structure of I that the limiting system is the Fokker-Planck equation. In addition, according to [DG87], I is the rate functional of the large-deviation principle for the empirical process

$$\sigma_n(t, dx) := \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i(t)},$$

where $dX_i(t) = -\nabla V(X_i(t))\, dt + \sqrt{2}\, dW_i(t)$ and $W_i, i = 1, \dots, n$ are independent Wiener processes.

To make the argument rigorous, we need to justify that the functions we used in Assumption 5.1 and in Step 3 are indeed admissible. In Assumption 5.1, it is not straightforward to see whether or not $f = 2 \log F$, where F is a solution of (99), is bounded and has sufficient regularity. We expect that this difficulty can be overcome by using the fact that (97) is a hypoelliptic equation. In Step 3, the function $f = g \circ \Pi_\gamma$ does not has compact support. Hence, we need to approximate these two functions by a sequence of smooth functions with compact support. Some modification of the argument in the proof of Lemma 4.11 in [DG87] might be required.

References

[ADPZ11] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage. *Communications in Mathematical Physics*, 307:791–815, 2011.

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics. ETH Zürich. Birkhauser, Basel, 2nd edition, 2008.

[BB00] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[Bra02] A. Braides. *Gamma convergence for beginners.* Oxford University Press, Oxford, 2002.

[CL12] X. Chen and J. G. Liu. Two nonlinear compactness theorems in $L^p(0, T; B)$. *Applied Mathematics Letters*, 25(12):2252–2257, 2012.

[DG87] D. A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[DLP+16] M. H. Duong, A. Lamacz, M. A. Peletier, A. Schlichting, and U. Sharma. Quantification of coarse-graining error in langevin and overdamped langevin dynamics. *In preparation*, 2016.

[DLPS16] M. H. Duong, A. Lamacz, M. A. Peletier, and U. Sharma. Variational approach to coarse-graining of generalized gradient fows. *Accepted for publication in Calculus of Variations and Partial Differential Equations*. See also at Arxiv preprints `http://arxiv.org/abs/1507.03207`, 2016.

[DLR13] M. H. Duong, V. Laschos, and D. R. M. Renger. Wasserstein gradient flows from large deviations of many-particle limits. *ESAIM Control Optim. Calc. Var.*, 19(4):1166–1188, 2013.

[DPZ13a] M. H. Duong, M. A. Peletier, and J. Zimmer. Conservative-dissipative approximation schemes for a generalized kramers equation. *Mathematical Methods in the Applied Sciences, To appear*, 2013.

[DPZ13b] M. H. Duong, M. A. Peletier, and J. Zimmer. GENERIC formalism of a Vlasov-Fokker-Planck equation and connection to large-deviation principles. *Nonlinearity*, 26(11):2951–2971, 2013.

[Duo14] M. H. Duong. *Large deviation and variational approaches to generalised gradient flows*. PhD thesis, Eindhoven University of Technoligy, 2014.

[DZ87] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic modelling and applied probability*. Springer, New York, NY, USA, 2nd edition, 1987.

[EMR15] M. Erbar, J. Maas, and M. Renger. From large deviations to wasserstein gradient flows in multiple dimensions. *Electron. Commun. Probab.*, 20:12 pp., 2015.

[FK06] J. Feng and T. G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.

[FW94] M. I. Freidlin and A. D. Wentzell. Random perturbations of hamiltonian systems. *Mem. Amer. Math. Soc.*, 109 (523), 1994.

[GPK12] B. D. Goddard, G. A. Pavliotis, and S. Kalliadasis. The overdamped limit of dynamic density functional theory: rigorous results. *Multiscale Model. Simul.*, 10(2):633–663, 2012.

[JKO98] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the fokker-planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

[Kra40] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.

[MPR14] A. Mielke, M.A. Peletier, and D.R.M. Renger. On the relation between gradient flows and the large-deviation principle, with applications to markov chains and diffusion. *Potential Analysis*, 41(4):1293–1327, 2014.

[MSZ03] J. Malý, D. Swanson, and W. P. Ziemer. The co-area formula for Sobolev mappings. *Trans. Amer. Math. Soc.*, 355(2):477–492 (electronic), 2003.

[Nel67] E. Nelson. *Dynamical theories of Brownian motion*. Princeton University Press, Princeton, N.J., 1967.

[Ott01] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[Ött05] H. C. Öttinger. *Beyond equilibrium thermodynamics*. Wiley-Interscience, 1st edition, 2005.

[Sim86] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Annali di Matematica Pura ed Applicata*, 146(1):65–96, 1986.

[Vil03] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Wil76] G. Wilemski. On the derivation of smoluchowski equations with corrections in the classical theory of brownian motion. *Journal of Statistical Physics*, 14(2):153–169, 1976.