The handle http://hdl.handle.net/1887/20110 holds various files of this Leiden University dissertation.

**Author:** Wang, Feijia  
**Title:** Dynamical Gibbs-non-Gibbs transitions: a study via coupling and large deviations  
**Issue Date:** 2012-11-07
Large deviations for the trajectory of the empirical distribution and empirical measure

4.1 Introduction

In [3] we started investigating how Gibbs-non-Gibbs transitions in lattice spin systems can be related to a bifurcation phenomenon in the nature of the optimal trajectories of the empirical measure. In [9] such a formalism was developed in the mean-field context, i.e., for the trajectory of the magnetization. The idea to study the trajectory of the empirical measure is that one conditions to arrive at time $T > 0$ at a given empirical measure and at time zero one gives a certain cost to each starting measure. This cost is determined by the choice of the initial Gibbs measure, i.e., equals the relative entropy density w.r.t. the initial Gibbs measure $\mu$. Uniqueness for every conditioning of the empirical measure at time $T > 0$ or non-uniqueness for a particular conditioning of the empirical measure at time $T$, correspond (roughly speaking) to Gibbsianness or non-Gibbsianness of the distribution $\mu_T$ at time $T$. The total cost to arrive at time $T > 0$ at a given empirical measure is the sum of the initial cost and a path cost, determined by the Markovian dynamics. This path cost is usually of the form of a Lagrangian action. This means, informally written, that the probability of a trajectory of the empirical measure, where one averages shifts of the point mass of the lattice-spin configuration over the box $[-N, N]^d$, is expected to behave as

$$P \left( (\mathbb{L}_N(\sigma_t))_{0 \leq t \leq T} \approx (\mu_t)_{0 \leq t \leq T} \right) \approx \exp \left( -(2N + 1)^d \int_0^T \Xi(\mu_s, \mu_s) \, ds \right)$$

The Lagrangian $\Xi(\mu_s, \mu_s)$ is the object we are after in the present paper.

More precisely, we consider two cases in the present paper.
First, in the context of independent Markov processes on a general state space $E$, we study the Lagrangian of the associated to the large deviations of the trajectory of the empirical distribution

$$\mathcal{L}_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i^t}$$

which is a random probability measure on $E$. We compute explicitly the Hamiltonian and provide some information about the associated Hamiltonian trajectories for finite-state space Markov chains. For diffusion processes, the Lagrangian is a natural quadratic form associated to the generator. For Markov chains, the Lagrangian is less explicit (except for two-state Markov chains), but can still be characterized as a relative entropy production. Moreover, the characterization of minimal action or optimal (w.r.t. path cost) trajectories can be done using the Hamiltonian formalism as well. The study of the large deviations of the trajectory of the empirical distribution has to be considered as the intermediate step between the magnetization (studied in [9]) and the empirical measure. In particular, for finite-state space Markov chains, the empirical distribution is still a finite-dimensional object.

Second, in the context of translation invariant interacting Markov processes, we consider the trajectory of the empirical measure, and compute explicitly the Hamiltonian, both for diffusion processes and for jump processes of interacting particle type. In the context of diffusion processes, the Lagrangian is a quadratic form, while in the context of jump processes (of interacting particle systems type), the Lagrangian is less explicit, but also there a relative entropy production (density) characterization can be given.

This study is a step in the research programme proposed in [3]. Given the Hamiltonians and Lagrangians computed in the present paper, one can then characterize bifurcation phenomena, i.e., non-uniqueness of optimal trajectories for particular choices of initial costs. We leave this problem for future work and focus here on the explicit form of the Hamiltonian and Lagrangian.

Our paper is organized as follows. In section 2 we give a general computation of the Feng-Kurtz Hamiltonian for the trajectory of the empirical distribution. In section 3 we study the case of finite continuous-time Markov chains. In section 4 we consider the case of diffusion processes. In section 5 we consider the case of interacting Markov processes, both of jump type (interacting particle systems in the spirit of [12]) and of diffusion type.
4.2 The trajectory of the empirical distribution: general case

We consider \( \{X_t : 0 \leq t \leq T\} \) a (Feller) Markov process on a state space \( E \). We assume \( E \) to be a locally compact Polish space. Relevant cases for the present paper are, \( E \) a finite set (finite Markov chains), or \( E = \mathbb{R}^k \) or a compact submanifold of \( \mathbb{R}^k \) (diffusions). The computation of this section is however valid for general \( E \).

We denote by \( Q \) the generator of the process \( \{X_t : 0 \leq t \leq T\} \), i.e.,

\[
Qf(x) = \lim_{t \to 0} \frac{1}{t} (\mathbb{E}_x f(X_t) - f(x))
\]

for \( f \in D(Q) \). The corresponding semigroup is denoted by \( S_t \). For \( E \) compact \( S_t \) acts on \( C_b(E) \), the space of bounded continuous functions on \( E \), while for \( E = \mathbb{R}^d \), \( S_t \) acts on \( C_0(E) \), the space of continuous functions vanishing at infinity. We further denote \( C_b(E) \) the space of bounded continuous functions on \( E \). For \( \mu \) a finite Borel measure on \( E \) and \( f \in C_b(E) \), we denote \( \langle \mu, f \rangle = \int f d\mu \). We denote by \( \mathcal{P}(E) \) the set of probability measures on \( E \).

We now let \( \{X^i_t : 0 \leq t \leq T\} \) be independent copies of the process \( \{X_t : 0 \leq t \leq T\} \) starting at initial points \( X^i_0 = x_i \), and consider the empirical distribution

\[
M_N(t) = \frac{1}{N} \sum_{i=1}^{N} \delta_{X^i_t}
\]

(4.2.1)

This is a random probability measure on \( E \), i.e., a random element of \( \mathcal{P}(E) \), which in the limit \( N \to \infty \) converges to the solution of the Kolmogorov forward equation.

If at time zero, \( M_N(0) \to \mu \) (where \( \mu \) is a probability measure on \( E \)), then at time \( t \), \( M_N(t) \to \mu_t \), where \( \mu_t \) solves

\[
\frac{d\mu_t}{dt} = Q^* \mu_t
\]

(4.2.2)

where \( Q^* \) denotes the dual generator defined via

\[
\langle \mu, Qf \rangle = \langle Q^* \mu, f \rangle
\]

Indeed, by the law of large numbers, for all \( f \in C_b(E) \),

\[
\langle M_N(t), f \rangle = \sum_{i=1}^{N} \mathbb{E}_{x_i}f(X^i_t) \to \int f \, d\mu_t
\]

where \( \mu_t \) denotes the law of \( X_t \) when started initially from \( X_0 \) distributed according to \( \mu \).

The convergence \( M_N(t) \to \mu_t \) is a manifestation of the law of large numbers, and therefore it is natural to expect an associate large deviation principle, i.e.,

\[
\mathbb{P}(\{M_N(t) : 0 \leq t \leq T\} \approx \{\mu_t : 0 \leq t \leq T\}) \approx \exp(-N\mathbb{J}(\{\mu_t : 0 \leq t \leq T\}))
\]

(4.2.3)
where \( \approx \) has to be interpreted in the sense of the large deviation principle, in a suitable topology on the space of trajectories. By the Markov property, the rate function \( I \) has the form of a Lagrangian “action”

\[
I(\{\mu_t : 0 \leq t \leq T\}) = \int_0^T L(\mu_s, \dot{\mu}_s) \, ds
\]

where \( \dot{\mu}_s \) denotes the weak derivative of the trajectory at time \( s \), defined via

\[
\langle \dot{\mu}_s, f \rangle = \frac{d}{ds} \langle \mu_s, f \rangle
\]

Our aim here is to compute the Lagrangian \( L \).

This opens the road to an analysis of bifurcation phenomena related to Gibbs-non-Gibbs transitions, as is done on the level of the magnetization in [9], [3]. The case of the empirical distribution would correspond to Gibbs-non-Gibbs phenomena in the context of mean-field models, where the mean field interaction is a function of possibly several empirical averages (rather than only of the magnetization).

Notice that the expression of \( L \) is independent of the precise topology (on the space of trajectories of probability measures on \( E \)) in which the large deviation principle (4.2.3) holds. As usual, one then first considers the weakest topology which is product topology (pointwise convergence at every time), and if one wants to strengthen the topology to e.g. uniform topology, one proves exponential tightness in that topology. In this paper we focus on the computation of the lagrangian \( L \) with the scheme of Feng and Kurtz [5], explained e.g. in [11].

In our context this means that we first compute the non-linear generator. To explain this, we need some more notation. First notice that \((X_1^1, X_2^1, \ldots, X_N^N)\) is a Markov process with generator

\[
Q_N f(x_1, \ldots, x_n) = \sum_{i=1}^N Q_i f
\]

where \( Q_i \) denotes the generator \( Q \) applied to the \( i \)-th coordinate. Next, for \( N \) points \( x_1, \ldots, x_N \in E \), and a function \( F : \mathcal{P}(E) \to \mathbb{R} \) we use the notation \( F(M_N) = F \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) \).

The first computation in the Feng-Kurtz scheme is then the non-linear generator

\[
HF(\mu) = \lim_{N \to \infty, M_N(x_1, \ldots, x_N) \to \mu} \frac{1}{N} \left( e^{-NF(M_N)} Q_N e^{NF(M_N)} \right)
\]

If \( \mathcal{H} \) is of the form \( \mathcal{H}(\mu, \nabla F) \), with \( \mathcal{H} \) a strictly convex function in the second variable, then we call \( \mathcal{H}(\mu, f) \) the Feng-Kurtz Hamiltonian, and the Lagrangian is then given by the Legendre transform of \( \mathcal{H} \):

\[
\mathcal{L}(\mu, \alpha) = \sup_{f \in \mathcal{C}(E)} \left( \int f \, d\alpha - \mathcal{H}(\mu, f) \right)
\]
4.3 Finite-state space continuous-time Markov chains

The interpretation of the “gradient” $\nabla F$ is straightforward when we are in the context of finite-state space Markov chains, because the set $\mathcal{P}(E)$ is then finite-dimensional. In the context of diffusion processes or more general Markov processes, the gradient will be a (context dependent) functional derivative.

The second variable of the Lagrangian (4.2.7) is the velocity variable, which in our context is a signed measure of total mass zero.

The Hamiltonian $\mathcal{H}(\mu, f)$ can then be obtained as follows:

$$\mathcal{H}(\mu, f) = \lim_{N \to \infty, M_N \to \mu} \frac{1}{N} e^{-N(\mathcal{M}_N, f)} Q_N e^{N(\mathcal{M}_N, f)}$$

(4.2.8)

Notice here that for a given $f \in \mathcal{C}(E)$, the function $e^{N(\mathcal{M}_N, f)} = e^{\sum_{i=1}^{N} f(x_i)}$ is a function from $E^N$ to $\mathbb{R}$, on which the generator $Q_N$ can act, i.e., the notation in $Q_N e^{N(\mathcal{M}_N, f)}$ makes sense.

The $\mu$ variable is interpreted as the “position” and the $f$ variable as the “momentum” (dual to the velocity variable in the Lagrangian formalism).

By the form (4.2.5) of the independent generator, the Hamiltonian can be computed:

$$\mathcal{H}(\mu, f) = \lim_{N \to \infty, M_N \to \mu} \frac{1}{N} e^{-N(\mathcal{M}_N, f)} Q_N e^{N(\mathcal{M}_N, f)}$$

$$= \lim_{N \to \infty, M_N \to \mu} \frac{1}{N} \sum_{i=1}^{N} e^{-f(x_i)} Q e^{f(x_i)}$$

$$= \int e^{-f} Q e^{f} d\mu$$

(4.2.9)

Notice that since $\mathcal{H}(\mu, 0) = 0$, for the corresponding Lagrangian (4.2.7) we have

$$\mathcal{L}(\mu, \alpha) \geq (\langle \alpha, 0 \rangle - \mathcal{H}(\mu, 0)) = 0$$

i.e., the Lagrangian is automatically non-negative (as it should be since it is the integrand of the rate function).

4.3 Finite-state space continuous-time Markov chains

In this case $E = \{a_1, \ldots, a_k\}$ is a finite set, of which we denote the elements by $a, b, \ldots$. The continuous-time Markov chain is defined via its transition rates between states $a, b \in E$, denoted by $r(a, b)$. The generator is given by

$$Qf(a) = \sum_{b \in E} r(a, b)(f(b) - f(a))$$

(4.3.1)
and correspondingly, the Kolmogorov forward equation reads
\[
\frac{d\mu_a(t)}{dt} = \sum_b (r(b,a)\mu_b(t) - r(a,b)\mu_a(t)) \tag{4.3.2}
\]

Since a function \( f : E \to \mathbb{R} \) is identified with a column of numbers \( f_a, a \in E \), we will use both notations \( f(a) \), or \( f_a \), idem for probability measures (identified with rows \( \mu_a, a \in E \)).

The Hamiltonian \([4.2.9]\) is given by
\[
\mathcal{H}(\mu, f) = \sum_{a,b \in E} \mu_a r(a,b) (e^{f_b - f_a} - 1) \tag{4.3.3}
\]
The Lagrangian is then
\[
\mathcal{L}(\mu, \alpha) = \sup_f \left( \sum_a f_a \alpha_a - \mathcal{H}(\mu, f) \right) \tag{4.3.4}
\]
The \( f = f^*(\alpha) \) which realizes the supremum satisfies
\[
\alpha_b = \sum_a \left( \mu_a r^*(a,b) e^{f^*_b - f^*_a} - \mu_b r^*(b,a) e^{f^*_a - f^*_b} \right) \tag{4.3.5}
\]
which then gives
\[
\mathcal{L}(\mu, \alpha) = \sum_{a,b} \mu_b r^*(b,a) \left( f^*_a e^{f^*_a - f^*_b} - f^*_b e^{f^*_a - f^*_b} - (e^{f^*_a - f^*_b} - 1) \right) \tag{4.3.6}
\]
defining the “modified” rates
\[
r^*(b,a) = r(b,a) e^{f^*_a - f^*_b}
\]
the equation \(4.3.5\) reads
\[
\alpha_b = \sum_a \left( \mu_a r^*(a,b) - \mu_b r^*(b,a) \right) \tag{4.3.7}
\]
which can be interpreted as follows. The modified rates are such that they produce “velocity” \(4.3.2\) equal to \( \alpha \), when started from initial measure \( \mu \). In terms of these modified rates \( r^* \), the Lagrangian reads
\[
\mathcal{L}(\mu, \alpha) = \sum_{a,b} \mu_b r^*(b,a) \log \left( \frac{r^*(b,a)}{r(b,a)} \right) - \sum_{a,b} \mu_b (r^*(b,a) - r(b,a)) \tag{4.3.8}
\]
This can be interpreted in terms of relative entropy as follows. The Radon-Nikodym derivative of the path space measure of the process with rates \( r^* \) w.r.t. the process with rates \( r \) is given by the Girsanov formula
\[
\frac{d\mathbb{P}^{[0,T]}_{r^*}}{d\mathbb{P}^{[0,T]}_{r}}(\omega) = \exp \left( \sum_{a,b} \left( \log \left( \frac{r^*(b,a)}{r(b,a)} \right) N_{T,a}^{b,a} - (r^*(b,a) - r(b,a)) l(T,b,\omega) \right) \right) \tag{4.3.9}
\]
where $N^{(b,a)}_t$ denotes the number of transitions from $b$ to $a$ in $[0,t]$ and where $l(T, b, \omega)$ denotes the total time that the path $\omega$ spends in state $b$. The corresponding relative entropy is

$$s(P^{[0,T]}_r\|P^{[0,T]}_r) = \int dP^{[0,T]}_r \log \left( \frac{dP^{[0,T]}_r}{dP^{[0,T]}_r} \right)$$

Taking the limit $T \to 0$ in this expression, starting from initial distribution $\mu$, we find

$$\lim_{T \to 0} \frac{1}{T} s(P^{[0,T]}_r\|P^{[0,T]}_r) = \mathcal{L}(\mu, \alpha)$$

(4.3.10)

In words this means the following. In order to compute $\mathcal{L}(\mu, \alpha)$, we have to consider an auxiliary Markov process with rates that from starting from $\mu$ produce velocity (in the sense of (4.3.2)) equal to $\alpha$. The relative entropy of this process w.r.t. the original process in a small interval of time $[0,t]$ is then given by $t\mathcal{L}(\mu, \alpha) + O(t^2)$. The Lagrangian $\mathcal{L}(\mu, \alpha)$ can thus be viewed as a “relative entropy production” needed to force the process to have speed $\alpha$ when started from $\mu$. In particular for $\alpha = Q^*\mu$, the cost is zero, and we have $\mathcal{L}(\mu, Q^*\mu) = 0$. This shows that the evolution according to the Kolmogorov forward equation is of course an optimal trajectory, with zero cost.

### 4.3.1 Hamiltonian trajectories for finite Markov chains

The Hamiltonian (4.3.3) has Hamiltonian trajectories given by

$$\dot{f}_a = -\frac{\partial H}{\partial \mu_a} = -\sum_b r(a,b)(e^{f_a - f_b} - 1)$$

$$\dot{\mu}_a = \frac{\partial H}{\partial f_a} = \sum_b \left( \mu_b r(b,a) e^{f_a - f_b} - \mu_a r(a,b) e^{f_b - f_a} \right)$$

(4.3.11)

The interpretation of the second equation is the following. For a trajectory with “momentum” $f$, the motion of the probability measure is that of a Markov process with rates which are modified according to $f$ via

$$\tilde{r}(a,b) = r(a,b) e^{f_a - f_b}$$

(4.3.12)

Indeed, for the modified rates $\tilde{r}$, the second equation of (4.3.11) reads simply

$$\dot{\mu}_a = \sum_b \mu_b \tilde{r}(b,a) - \mu_a \tilde{r}(a,b)$$

which is precisely the Kolmogorov forward equation for the evolution of a probability distribution in a Markov chain with rates $\tilde{r}$. 

75
The equation for the momenta, i.e., the first equation of (4.3.11) can be rewritten using the variables $u_a = e^{f_a}$, $a \in E$:

$$\dot{u}_a = -\sum_b r(a, b)(u_b - u_a) = -(Qu)_a$$

which has the solution

$$u(t) = e^{-tQ}u(0)$$

(4.3.13)

The equation for the “position variables” $\mu_a$ is linear and reads

$$\mu(t) = M(u(t))\mu(t)$$

(4.3.14)

with $M$ a matrix depending on the solution of the momentum variables, given by

$$M_{a,b}(f) = r(b, a)\frac{u_a}{u_b} - \left(\sum_c r(a, c)\frac{u_c}{u_a}\right)\delta_{a,b}$$

(4.3.15)

This matrix has column sums equal to zero, i.e., for all $b \in E$ we have $\sum_a M_{a,b} = 0$, which corresponds to the conservation of mass $\sum_a \mu_a(t) = 1$ in the Hamiltonian evolution. More precisely, the matrix $M_{a,b}$ is precisely the adjoint of the generator corresponding to the modified rates $\tilde{r}$ defined in (4.3.12).

We thus conclude that the Hamiltonian trajectories are still Markovian, corresponding with time-dependent rates, steered by the solution of the momentum equation (4.3.13).

The solution of (4.3.14) is given by

$$\mu(t) = e^{\int_0^T M(u(s))}\mu(0)$$

(4.3.16)

which means that we have the form of the optimal trajectories, with integration constants given by $u(0)$ and $\mu(0)$. Although the form (4.3.16), (4.3.13) looks quite explicit, it is not easy in general to find explicit tractable formulas for $\mu(t)$. The action or path-space cost

$$J = \int_0^T \mathcal{L}(\mu_s, \dot{\mu}_s) \, ds$$

(4.3.17)

can be rewritten in Hamiltonian formalism as

$$J(\{\mu(s), f(s) : 0 \leq s \leq T\}) = \sum_a \int_0^T f_a(t)\dot{\mu}_a(t) \, dt - J(\mu(0), f(0))$$

(4.3.17)

This means that in order to find the optimal cost between a measure $\mu(0) = \mu$ and a measure $\mu(T) = \nu$ at time $T$, one has to plug in the solution (4.3.16), (4.3.14) into the expression (4.3.17), and determine the integration constants $\mu(0), f(0)$ by initial and final condition. This leads to a function $\Psi(\mu, \nu, T)$ which is the optimal path cost to travel from $\mu$ to $\nu$ in time $T$. In concrete situations beyond two-state Markov chains, in practice, this function is hard to obtain closed formulas for (an issue which we do not want to pursue here).
4.3 Finite-state space continuous-time Markov chains

4.3.1.1 Two-state symmetric flipping

To see an example of an explicit solution, we consider the case of two-states flipping at rate 1, which corresponds with mean-field independent spin flip dynamics, treated before in [9], [3], [11].

In that case, the state space is given by $E = \{1, 2\}$, the matrix $Q$ is given by

$$Q = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

and the matrix $M$ of (4.3.15) is given by

$$M = \begin{pmatrix} -u_2 & u_1 \\ u_1 & -u_2 \end{pmatrix}$$

where $u = (u_1, u_2)^T$ satisfies

$$\dot{u} = -Qu \quad (4.3.18)$$

The equation

$$\dot{\mu} = M\mu$$

can be differentiated w.r.t. time once more, which gives

$$\frac{d^2\mu}{dt^2} = \left( \frac{dM}{dt} + M^2 \right) \mu$$

Explicit computation, using (4.3.18) then gives

$$\frac{dM}{dt} + M^2 = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

which gives the equations

$$\frac{d^2\mu_1(t)}{dt^2} = 2\mu_1(t) - 2\mu_2(t) = -2 \frac{d^2\mu_2(t)}{dt^2}$$

Putting $\mu_1 - \mu_2 = x$ we have,

$$\frac{d^2x}{dt^2} = 4x$$

which gives $x_t = C_1 e^{2t} + C_2 e^{-2t}$ as optimal solutions, consistent with e.g. [9], or [3].

Remark 4. The fact that $\frac{dM}{dt} + M^2$ is a constant matrix is quite exceptional. Even in the two-state case, if the rates $r(1, 2) = \alpha \neq r(2, 1) = \beta$, the matrix $\frac{dM}{dt} + M^2$ is not constant and differentiating the equation (4.3.14) once more does not lead to further simplification.
4. LARGE DEVIATIONS FOR THE TRAJECTORY OF THE
EMPIRICAL DISTRIBUTION AND EMPIRICAL MEASURE

4.4 Diffusion processes

Here we consider the state space $E = \mathbb{R}^n$ and diffusion processes with generator

$$Q = \sum_i b_i(x) \partial_i + \sum_{ij} a_{ij}(x) \partial_{ij}^2$$ (4.4.1)

where $\partial_i$ denotes partial derivative w.r.t. $x_i$. Here $b_i(x), a_{ij}(x)$ are supposed to be Lipschitz and sufficiently smooth, ensuring the existence of a solution of the martingale problem associated to $Q$.

The covariance $a_{ij}(x)$ is supposed to be a positive definite matrix. Moreover, for simplicity we assume that it is bounded from below by a multiple of the identity (to avoid degeneracies).

The Feng-Kurtz Hamiltonian $\mathcal{H}(\mu, f)$ (4.2.9) can then be computed and this yields:

$$\mathcal{H}(\mu, f) = \exp - f Q e^f d\mu$$

$$= \int \left( Qf + \sum_{ij} \partial_i f(x) \partial_j f(x) a_{ij}(x) \right) d\mu(x)$$ (4.4.2)

The measures $\mu$ that we will have to consider are absolutely continuous probability measures w.r.t. Lebesgue measure, $\mu = \mu(x)dx$, where with slight abuse of notation we use the symbol $\mu$ both for the measure and for the density.

Although we are in the infinite-dimensional context here, since the Hamiltonian is quadratic, the corresponding Lagrangian can be obtained more easily than in the previous subsection.

Define the quadratic form

$$J_\mu(f, f) = \int \left( \sum_{ij} \partial_i f(x) \partial_j f(x) a_{ij}(x) \right) d\mu(x)$$ (4.4.3)

To this quadratic form corresponds a positive self-adjoint operator $A_\mu$ (linearly depending on $\mu$) such that

$$J_\mu(f, f) = \frac{1}{2} \langle f, A_\mu f \rangle$$

where $\langle f, g \rangle = \int f(x)g(x) \, dx$ is the usual $L^2$ innerproduct.

With this notation, the Hamiltonian can be written in the form

$$\mathcal{H}(\mu, f) = \langle \mu, Qf \rangle + \frac{1}{2} \langle f, A_\mu f \rangle = \langle Q^* \mu, f \rangle + \frac{1}{2} \langle f, A_\mu f \rangle$$ (4.4.4)
4.4 Diffusion processes

Then, the corresponding Lagrangian is computed

\[
\mathcal{L}(\mu, \alpha) = \sup_f \left( \langle f, \alpha \rangle - \langle Q^* \mu, f \rangle - \frac{1}{2} \langle f, A_\mu f \rangle \right) = \frac{1}{2} \langle (\alpha - Q^* \mu), A_\mu^{-1} (\alpha - Q^* \mu) \rangle \tag{4.4.5}
\]

The rigorous meaning of \( \langle f, A_\mu^{-1} f \rangle \) is \( \| A_\mu^{-1/2} f \|_2^2 \) for \( f \) in the domain of \( A_\mu^{-1/2} \). The Lagrangian is then defined to be infinite when \( (\alpha - Q^* \mu) \) is not in the domain of \( A_\mu^{-1/2} \) (cf. the abstract form of Schilder’s theorem in abstract Wiener spaces see [2]).

We see that the “typical trajectory” which follows the Kolmogorov forward equation has zero cost, since in that case \( \dot{\mu} = \alpha = Q^* \mu \), and hence \( \mathcal{L}(\mu, \alpha) = 0 \), and the Lagrangian is a quadratic expression in the deviation of the trajectory from the Kolmogorov forward equation.

To illustrate this formula, let us consider first the simplest example of the present context, i.e., dimension \( n = 1 \), drift \( b = 0 \), \( a = 1/2 \), corresponding to a one-dimensional Brownian motion. The generator is

\[
Q = \frac{1}{2} \frac{d^2}{dx^2}
\]

\( Q^* = Q \). The quadratic form (4.4.3) reads in this case

\[
J_\mu(f, f) = \frac{1}{2} \int \mu(x)(f')^2 dx
\]

and the corresponding operator

\[
A_\mu = \frac{d}{dx} \left( \mu(x) \frac{d}{dx} \right)
\]

which gives

\[
\mathcal{L}(\mu, \alpha) = \frac{1}{2} \left\langle \nabla^{-1} \left( \alpha - \frac{1}{2} \mu'' \right), \frac{1}{\mu} \nabla^{-1} \left( \alpha - \frac{1}{2} \mu'' \right) \right\rangle \tag{4.4.6}
\]

The rigorous meaning of the formal expression \( \langle \nabla^{-1} f, \nabla^{-1} g \rangle \) is the innerproduct in the space \( H^{-1} \), i.e., \( \langle (-\Delta)^{-1/2} f, (-\Delta)^{-1/2} g \rangle \), with \( \Delta = \frac{d^2}{dx^2} \), the Laplacian.

Remark 5. The rate function (4.4.6) has also been obtained in the context of the study of the hydrodynamic limit for independent Brownian particles, in [8]. In general, it is an interesting question to understand the relation between the rate functions which are computed in this paper and the rate functions for deviations of the hydrodynamic limit, see e.g. [7]. For Brownian particles, they coincide because of scale invariance of the Brownian motion.

The Lagrangian (4.4.6) can be interpreted in terms of relative entropy. A diffusion process on \( \mathbb{R} \) with drift \( b(x) \) and variance equal to one has the generator

\[
Q_b = b(x) \frac{d}{dx} + \frac{1}{2} \frac{d^2}{dx^2}
\]
4. LARGE DEVIATIONS FOR THE TRAJECTORY OF THE
EMPIRICAL DISTRIBUTION AND EMPIRICAL MEASURE

if we start this process from a measure \( \mu = \mu(x)dx \), then the infinitesimal change at time zero is given by the adjoint generator working on \( \mu \), i.e.,

\[
\frac{1}{2} \frac{d^2 \mu}{dx^2} + \frac{d}{dx}(b(x)\mu(x)) = (Q_b^* \mu)(x)
\]

In particular, for \( \alpha \), a given absolutely continuous signed measure of total mass zero, we can find the drift \( b \) that corresponds to it, i.e., solving the equation

\[
\frac{1}{2} \frac{d^2 \mu}{dx^2} + \frac{d}{dx}(b(x)\mu(x)) = \alpha(x)
\]

The process with drift \( b \) has a corresponding path space measure on Wiener space given by \( \mathbb{P}_b^{[0,T]} \), and we have the Girsanov formula

\[
\frac{d\mathbb{P}_b}{d\mathbb{P}_0} = \exp \left( \int_0^T b(W_s) dW_s - \frac{1}{2} \int_0^T b^2(W_s) ds \right)
\]

The relative entropy of the process with drift \( b \) w.r.t. the zero drift process is thus given by

\[
s(\mathbb{P}_b^{[0,T]} | \mathbb{P}_0^{[0,T]}) = \int d\mathbb{P}_b \log \left( \frac{d\mathbb{P}_b}{d\mathbb{P}_0} \right)
= \mathbb{E}_0 \left( \left( \int_0^T b(W_s) dW_s - \frac{1}{2} \int_0^T b^2(W_s) ds \right) \exp \left( \int_0^T b(W_s) dW_s - \frac{1}{2} \int_0^T b^2(W_s) ds \right) \right)
\]

where the expectation \( \mathbb{E}_0 \) is over the standard Brownian motion. Computing then

\[
\lim_{T \to 0} \frac{1}{T} s(\mathbb{P}_b^{[0,T]} | \mathbb{P}_0^{[0,T]})
\]

using a starting distribution \( \mu \) gives exactly the expression

\[
\frac{1}{2} \int b^2(x)\mu(x)dx
\]

which equals \( \mathcal{L}(\mu, \alpha) \) of (4.4.6), because by (4.4.9)

\[
\frac{d}{dx}(b(x)\mu(x)) = \alpha - \frac{1}{2} \mu''(x)
\]

Hence, as in the finite Markov chain case, we see that the Lagrangian can be interpreted as the infinitesimal relative entropy cost to produce a derivative measure \( \alpha \) when started from \( \mu \). In particular, when \( \alpha = Q^* \mu \) this cost is zero, showing once more (in this context) that the evolution according to the Kolmogorov forward equation is an optimal trajectory with zero cost.
4.5 Trajectory of the empirical measure

4.5.1 Context and notation

In the context of translation invariant interacting systems, the empirical distribution is no longer a natural object because of interactions. The natural object capturing the essential information about the time evolution, modulo translations is the empirical measure. In order to describe this context, we need some more notation. For $N \in \mathbb{N}$ we denote $V_N = \{-N, \ldots, N\}^d$ and denote by $T^N_d$ the $d$-dimensional torus, i.e., $V_N$ endowed with addition modulo $2N + 1$.

We will consider translation invariant systems on this torus which for large $N$ have to be thought of as approximations of an infinite interacting system where the individual components live on the lattice $\mathbb{Z}^d$.

The configuration space is $\Omega_N = E^{T^N_d}$, where $E$, the single-site space, is a locally compact Polish space. Further we denote $\Omega = E^{\mathbb{Z}^d}$. As in the previous sections, we mostly consider $E$ or a finite set (interacting particle systems) or $E$ a submanifold of $\mathbb{R}^n$ (diffusion processes). Elements of $\Omega_N$ are denoted $\sigma, \eta, \xi, \ldots$, and for $\sigma \in \Omega_N$, $i \in T^N_d$, $\sigma_i$ denotes the value of the configuration at site $i$. On $T^N_d$ we have the addition modulo $N$, and correspondingly, the shift $\tau_i$ defined on $\Omega_N$ via

$$\tau_i(\sigma)_j = \sigma_{j+i}$$

(4.5.1)

on functions $f : \Omega_N \to \mathbb{R}$ via $\tau_i f(\sigma) = f(\tau_i \sigma)$, and on probability measures via $\int f d(\tau_i \mu) = \int \tau_i f d\mu$. If $A$ is a linear operator on functions $f : \Omega_N \to \mathbb{R}$ then we define its shift over $i$ to be $\tau_i A \tau_{-i}$, and an operator is called translation invariant if for all $i$, $\tau_i A \tau_{-i} = Q$. A measure is translation invariant if $\tau_i \mu = \mu$. Natural translation invariant measures on $\Omega_N$ are obtained by periodizing translation invariant measures on $\Omega$, i.e., starting from $\sigma$ distributed according to a translation invariant measure on $\Omega$, we consider $\sigma_i^N = \sigma_i$, $i \in V_N$, periodically extended to the whole lattice. Conversely, if we have a probability measure on $\Omega_N$ we naturally associate to it a probability measure on the infinite configuration space $\Omega$. This justifies the fact that with slight abuse of notation we can use sometimes the same symbol $\mu_N$ for a translation invariant measure on $\Omega_N$ as well as for a translation invariant measure on $\Omega$. We denote by $\mathcal{P}_{inv}(\Omega)$ the set of translation invariant probability measures on $\Omega$.

A function $f : \Omega \to \mathbb{R}$ is called local if it depends on a finite number of coordinates, i.e., if there exists a (minimal) finite set $D_f$, called the dependence set of $f$ such that for all $\sigma, \eta \in \Omega$: $f(\sigma_{D_f} \eta_{\mathbb{Z}^d \setminus D_f}) = f(\sigma)$, i.e., the value of the function is not influenced
by changing the configuration outside $Df$. Obviously, a local function $f : \Omega \to \mathbb{R}$ can be thought of as being a function $f : \Omega_N \to \mathbb{R}$ as well, for $N$ large enough such that $V_N \supset Df$. The translation $\tau_i f$ of local function is obviously local, with dependence set $D_{\tau_i f} = Df + i = \{x + i : x \in Df\}$.

An linear operator (possibly unbounded) $A : \mathcal{D}(A) \subseteq \mathcal{C}(\Omega) \to \mathcal{C}(\Omega)$ is local if it acts only on a finite set of coordinates. As for a local function, a local operator acts naturally on functions $f : \Omega_N \to \mathbb{R}$ for $N$ large enough.

4.5.2 Translation invariant sequence of local generators

**Definition 2.** A translation invariant sequence of local generators is defined to be a sequence of generators of the form $L_N = \sum_{i \in T_N} \tau_i Q_{\tau-i}$, with $Q$ a local generator, such that the corresponding infinite-volume generator $L = \sum_{i \in \mathbb{Z}^d} \tau_i Q_{\tau-i}$ is well-defined and has a core consisting of local functions as a core. The generator $Q$ is called the “source generator”.

**Remark 6.** For $E$ discrete, the core in Definition 4 consists typically of all local functions, whereas for $E$ being an interval or non-discrete, typically the core consists of smooth local functions.

As a consequence, the corresponding processes $\{\sigma_{N,t} : t \geq 0\}$ converge weakly in path space to the infinite-volume process $\{\sigma_t : 0 \leq t \leq T\}$ with generator $L$. Moreover, for the associated semigroups we have that $S_t^N f \to S_t f$ uniformly as $N \to \infty$, for all local functions $f$.

Let us give some examples in order to make this concept more concrete.

1. **Independent Markov processes.** For $Q$ a generator of a Markov process on $E$, we define $L_N = \sum_{i \in T_N} \tau_i Q_{\tau-i}$

Under the process with generator $L_N$ different components evolve independently, as copies of the process with generator $Q$.

2. **Spin-flip dynamics.** $E$ is finite set (e.g. $E = \{-1,1\}$ for Ising spins), $\theta : E \to E$ a bijection such that $\theta(a) \neq a$ for all $a \in E$. Furthermore, a local function $r : \Omega \to \mathbb{R}^+$, with dependence set containing the origin, is given. The local generator is then defined $Q f(\sigma) = r(\sigma) f(\theta_0 \sigma) - f(\sigma)$, where $\theta_0$ means applying $\theta$ to the coordinate $\sigma_0$ and leaving all other coordinates unchanged (similarly we denote $\theta_i$). The corresponding sequence of generators is then given by

$$L_N f(\sigma) = \sum_i ((\tau_i Q_{\tau-i}) f)(\sigma) = \sum_{i \in T_N} r(\tau_i \sigma)(f(\theta_i \sigma) - f(\sigma))$$
3. **Interacting diffusions.** For \( E = \mathbb{R} \) and for a finite set \( D \subseteq \mathbb{Z}^d \), we consider the local generator

\[
Q_f(\sigma) = \left( \sum_{i \in D} \frac{\partial V(\sigma_D)}{\partial \sigma_0} \frac{\partial}{\partial \sigma_i} \right) + \frac{1}{2} \frac{\partial^2}{\partial \sigma_0^2}
\]

and the corresponding

\[
\mathcal{L}_N = \sum_i \tau_i Q_{\tau_i} f
\]

This represents a system of diffusions, interacting via the potential \( V \). E.g. for a nearest neighbor potential \( V : \mathbb{R} \to \mathbb{R} \) in \( d = 1 \), the full generator has the form

\[
\sum_i V'(|\sigma_i - \sigma_{i-1}|) \left( \frac{\partial}{\partial \sigma_i} - \frac{\partial}{\partial \sigma_{i-1}} \right) + \frac{1}{2} \frac{\partial^2}{\partial \sigma_i^2}
\]

corresponding to \( D = \{0, 1\} \), \( V(\sigma_D) = V(|\sigma_1 - \sigma_0|) \).

4. **Local interacting particle systems.** \( E \) is a finite set. For finite subsets \( D_\alpha \subseteq \mathbb{Z}^d \), a collection of \( \tau_\alpha : E^{D_\alpha} \to E^{D_\alpha} \) \( \alpha \in \{1, \ldots, k\} \), and corresponding rates \( c(\alpha, \sigma) \) we consider the local generator

\[
Q_f(\sigma) = \sum_\alpha c(\alpha, \sigma)(f(T_\alpha \sigma_{D_\alpha} \sigma_{D_\alpha^c}) - f(\sigma))
\]

the corresponding local generators then include of course the previous spin-flip case but also translation invariant spin-exchange (Kawasaki) dynamics, combination of spin-flip and spin-exchange, etc.

5. **Local averaging.** For \( 0 \in D \subseteq \mathbb{Z}^d \) finite, and \( m_D \) a probability measure on \( A^D \), consider

\[
Q_f(\sigma) = r(\sigma) \int (f(\sigma'_D \sigma_{D^c}) - f(\sigma)) m_D(d\sigma'_D)
\]

with \( r \) a local function. In words, this means that with rate \( r \), the configuration inside \( D \) is replaced by its average over the measure \( m_D \). An important example of this class is the KMP model of heat conduction.

### 4.5.3 Trajectory of the empirical measure

For a configuration \( \sigma \in \Omega_N \), its empirical measure is defined by

\[
\mathbb{L}_N(\sigma) = \frac{1}{|T_N^d|} \sum_{i \in \mathbb{Z}^d_N} \delta_{\tau_i \sigma}
\]

(4.5.4)

this is a translation invariant probability measure on \( \Omega_N \), capturing all information about \( \sigma \), modulo translations.
For a configuration on the full lattice, $\sigma \in \Omega$, with a slight abuse of notation we also denote

$$L_N(\sigma) = \frac{1}{|T^N_d|} \sum_{i \in T^N_d} \delta_{\tau_i(\sigma^N)}$$

(4.5.5)

where $\sigma^N$ is the periodized configuration obtained from $\sigma$.

If $\mu$ is a probability measure on $\Omega$, which is ergodic under translations, then, by the Birkhoff ergodic theorem, with $\mu$ probability one

$$L_N(\sigma) \to \mu$$

as $N \to \infty$, where $\to$ means weak convergence.

If $(L_N)_N$ is a translation invariant sequence of local generators, then we have the associated Markov processes $\sigma_{N,t}$ with semigroups $S^N_t = e^{tL_N}$. For a probability measure $\mu$ on $\Omega$, let us denote $\mu_t$ to be the distribution at time $t > 0$ in the infinite-volume process $\{\sigma_t : t \geq 0\}$, started at initial state distributed according to $\mu$. By locality of the generator $L$, for $\mu$ ergodic, we have that $\mu_t$ is ergodic as well and hence

$$L_N(\sigma_t) \to \mu_t$$

weakly, with probability one. Hence the random trajectory of translation invariant probability measures $\{L_N(\sigma_t) : 0 \leq t \leq T\}$ converges, as $N \to \infty$ to the deterministic trajectory $\{\mu_t : 0 \leq t \leq T\}$. This convergence of a random $\mathcal{P}_{inv}(\Omega)$-valued trajectory to a deterministic $\mathcal{P}_{inv}(\Omega)$-valued trajectory can be thought of as a law of large numbers (in an infinite-dimensional space), and therefore it is natural to ask for an associated large deviation principle. For spin-flip dynamics, this was studied in [3]. Here we treat the general case of a translation invariant sequence of local generators. This will naturally lead to a non-linear operator $K_Q$ associated to the local generator $Q$, which will be the analogue in the present context of the non-linear operator $e^{-fQ}e_f$ in the previous section.

More precisely, we want to identify the “path space Lagrangian” (which is in this section is denoted by $\Xi$) such that

$$\mathbb{P}\left(\{L_N(\sigma_t) : 0 \leq t \leq T\} \approx \{\nu_t : 0 \leq t \leq T\}\right) \approx \exp\left(-|T^N_d| \int_0^T \Xi(\nu_t, \dot{\nu}_t)\right)$$

The Lagrangian is now a function of a translation invariant probability measure and a translation invariant signed measure of total mass zero, and as before, $\approx$ has to be interpreted in the sense of the large deviation principle, in this case, in the space of trajectories of translation invariant measures.
4.5 Trajectory of the empirical measure

4.5.4 The Feng-Kurtz Hamiltonian

In this section we compute the Feng-Kurtz Hamiltonian. This Hamiltonian is now a function from \( C(\Omega) \times P_{inv}(\Omega) \) to \( \mathbb{R} \), where the first variable has to be thought of as the “position” variable, whereas the second variable functions as a momentum variable. The Hamiltonian is defined as the limit

\[
H(\mu, f) = \lim_{N \to \infty, \mathcal{L}_N(\sigma) \to \mu} \frac{1}{|T^N_d|} \left( e^{-|T^N_d|\langle L_N(\sigma) f \rangle} \mathcal{L}_N e^{|T^N_d|\langle L_N(\sigma) f \rangle} \right) \tag{4.5.6}
\]

Note that \( |T^N_d|\langle L_N(\sigma) f \rangle = \sum_{i \in T^N_d} \tau_i f(\sigma) \).

For the computation of (4.5.6), we assume \( f \) to be a local function. Because the source generator \( Q \) is local we have, that \( Q(\tau_k f) = 0 \) for all \( k \) outside the set \( D(Q, f) = \{ k : D_Q \cap D_f + k \neq \emptyset \} \). Therefore, for \( \Lambda \subseteq \mathbb{Z}^d \) finite,

\[
Q \left( \prod_{i \in \Lambda} \tau_i f \right) = \left( \prod_{i \in \Lambda \setminus D(Q, f)} \tau_i f \right) Q \left( \prod_{i \in \Lambda \cap D(Q, f)} \tau_i f \right) \tag{4.5.7}
\]

Use (4.5.7) to compute

\[
\mathcal{H}(\mu, f) = \lim e^{-\sum_i \tau_i f} \sum_j \tau_j \left( Q e^{\sum_i \tau_i f} \right) \left( Q e^{\sum_k \tau_k f} \right) \tag{4.5.8}
\]

We can now introduce the non-linear operator associated to the “source” generator \( Q \), working on local functions \( f \):

\[
\mathcal{K}_Q f = e^{-\sum_{k \in D(f, \theta)} \tau_k f} Q e^{\sum_{k \in D(f, \theta)} \tau_k f} \tag{4.5.9}
\]

Using this notation, we obtain from (4.5.8)

\[
\mathcal{H}(\mu, f) = \int \mathcal{K}_Q(f) d\mu \tag{4.5.10}
\]

This Hamiltonian has to be thought of as the analogue of (4.2.9) in the present context.

**Remark 7.** Note that we can write, informally,

\[
\mathcal{K}_Q f = e^{-\sum_{k \in \mathbb{Z}^d} \tau_k f} Q e^{\sum_{k \in \mathbb{Z}^d} \tau_k f}
\]
4. LARGE DEVIATIONS FOR THE TRAJECTORY OF THE EMPIRICAL DISTRIBUTION AND EMPIRICAL MEASURE

since the terms \( k \notin D(f, Q) \) “cancel”. This is of course not rigorous because the sum \( \sum_{k \in \mathbb{Z}^d} \tau_k f \) is divergent, but this divergence is the “same” as for a formal infinite-volume Hamiltonian, where energy differences are well defined. The advantage of this formal representation is that we clearly see that \( K \) is a translation invariant operator, i.e., \( K_Q(f) = K_Q(\tau_r f) \), and as a consequence, the Hamiltonian \( H(\mu, f) \) is translation invariant as well, both in the measure and in the function, i.e.,

\[
H(\tau_k \mu, \tau_r f) = H(\mu, f)
\]

for all \( k, r \in \mathbb{Z}^d \). Another advantage is that one clearly sees the analogy with the corresponding formula for the empirical distribution \(^{(4.2.9)}\).

The corresponding Lagrangian is then found by Legendre transformation, i.e.,

\[
\Xi(\mu, \dot{\mu}) = \sup_{f \in C(\Omega)} \left( \int f \, d\dot{\mu} - H(\mu, f) \right)
\]

where \( \dot{\mu} \) denotes a translation invariant signed measure of total mass zero, and \( \mu \) a translation invariant probability measure on \( \Omega \).

In general, an explicit expression for \( \Xi \) cannot be obtained easily. In the examples below we will compute \( \Xi \) quite explicitly for diffusion processes and show a relative entropy interpretation of \( \Xi \) both in the context of interacting particle systems (analogue of finite Markov chains in the previous section) and in the context of interacting diffusions.

4.5.5 Interacting particle systems

We now compute \( K_Q \) for some of the examples discussed before, starting with interacting particle systems. The local generator is of the form

\[
Qf = \sum_{\alpha} r_\alpha (T_\alpha f - f)
\]

where \( T_\alpha \) are local transformations, which change coordinates only in a finite set \( D_\alpha \) containing the origin. This gives

\[
K_Q f = \sum_{\alpha} r_\alpha \left( e^{D_\alpha(f)} - 1 \right)
\]

(4.5.13)

where the operator \( D_\alpha \) is defined by

\[
D_\alpha f = \sum_{k \in \mathbb{Z}^d} (T_\alpha(\tau_k f) - \tau_k f)
\]

Notice that the sum is in fact a finite sum since \( f \) is local, and the transformation \( T_\alpha \) is local as well. Let us now zoom in into two familiar examples.
4.5 Trajectory of the empirical measure

- **Independent spin-flip.** For $E = \{-1, 1\}$, and for a single transformation $T \sigma = \sigma^0$ (spin-flip), we get
  \[ \mathcal{D}(f) = \sum_{k \in -D_f} (\tau_k f(\sigma^0) - \tau_k f) \]
  for the special functions $f(\sigma) = H_A(\sigma) = \prod_{i \in A} \sigma_i$ we get
  \[ \mathcal{D}(H_A) = \sum_{k \in -A} -2H_{A+k} \]
  as we found before in [3].

- **Exclusion process.** For $E = \{0, 1\}$, $d = 1$ and $T(\sigma) = \sigma^{01}$, where $\sigma^{01}$ denotes exchange of the values at site 0 and 1, i.e., $(\sigma^{01})_j = \sigma_1 \delta_{j,0} + \sigma_0 \delta_{j,1} + \sigma_j (1 - \delta_{j,0} - \delta_{j,1})$. We have
  \[ \mathcal{D}(f)(\eta) = \sum_{k, k+D_f \cap \{0,1\} \neq \emptyset} f(\tau^k(\eta^{01})) - f(\eta) \]
  Notice that for $f = \eta_j$ we find only two terms contributing to $\mathcal{D}(f)$:
  \[ \mathcal{D}(f) = ((\tau^{-j}(\eta^{01}))_j - (\tau^{-j}(\eta))_j) + ((\tau^{-j+1}(\eta^{01}))_j - (\tau^{-j+1}(\eta))_j) \]
  \[ = \eta_1 - \eta_0 + \eta_0 - \eta_1 = 0 \]
  which corresponds to the fact that the density of particles is conserved in this process.

The Lagrangian associated with (4.5.13) is
  \[ \Xi(\mu, \dot{\mu}) = \sup_{f \in C(\Omega)} \left( \int f d\dot{\mu} - \int \left( \sum_{\alpha} r_{\alpha} (e^{\mathcal{D}_\alpha(f)} - 1) \right) d\mu \right) \]
  (4.5.14)
  This expression is reminiscent of (4.3.4) in section 3 (empirical distribution for finite Markov chains). Indeed, a similar relative entropy interpretation of this expression can be given. We will describe this rather informally, making the arguments rigorous here is however completely standard and analogous to the Girsanov formula computation of the section on finite Markov chains. First we note that for a translation invariant measure $\mu$, its “derivative at time zero” $L^* \mu$ is formally given by
  \[ (L^* \mu)(\sigma) = \sum_i \sum_{\alpha} (r_{\alpha}(\tau_i \sigma) \mu(\tau_i T_{\alpha} \tau_i \sigma) - r_{\alpha}(\tau_i \sigma) \mu(\tau_i \sigma)) \]
  Suppose now we consider modified rates $\tilde{r}_{\alpha}(\sigma) = r_{\alpha}(\sigma) e^{f(\sigma) - f(T_\alpha(\sigma))}$, and the associated modified local generator $\tilde{Q} = \sum_{\alpha} \tilde{r}_{\alpha}(T_\alpha - I)$, i.e., the same transformations are applied with other rates. Then for a given translation invariant signed measure of total mass zero,
we look for those modified rates, i.e., choice of \( f \), such that with the starting measure \( \mu \) they produce “derivative at time zero” equal to \( \dot{\mu} \), i.e.,

\[
\dot{\mu}(\sigma) = \sum_i \sum_{\alpha} (\tilde{r}_\alpha(\tau_i \sigma)\mu(\tau_i T_\alpha \tau_i \sigma) - \tilde{r}_\alpha(\tau_i \sigma)\mu(\tau_i \sigma))
\]

The Radon-Nikodym derivative of the path space measure of the finite-volume process (in \( \mathbb{T}_N^d \)) with rates \( \tilde{r} \) w.r.t. the process with rates \( r \) is given by the Girsanov formula:

\[
\frac{dP[0,T,N\tilde{r}]}{dP[0,T,Nr]} = \exp \left( \sum_{i \in \mathbb{T}_N^d} \sum_{\alpha} \left( \int_0^T \log \frac{\tilde{r}^i_\alpha(\sigma_s)}{r^i_\alpha(\sigma_s)} dN_s^{i,\alpha} - \int_0^T (\tilde{r}^i_\alpha(\sigma_s) - r^i_\alpha(\sigma_s)) ds \right) \right)
\]

where \( r^i_\alpha \), resp. \( \tilde{r}^i_\alpha \) denote the rate to flip from \( \sigma \) to \( \tau_i T_\alpha \tau_i \sigma \), i.e., to apply the transformation \( T_\alpha \) around the lattice site \( i \), and \( N_s^{i,\alpha} \) the corresponding counting process counting how many transitions \( \sigma \) to \( \tau_i T_\alpha \tau_i \sigma \) have happened in the time interval \([0,t]\).

We then find, as in (4.3.10) that the Lagrangian is equal to the limit

\[
\Xi(\mu,\dot{\mu}) = \lim_{T \to 0} \frac{1}{T} \lim_{N \to \infty} \frac{1}{|T_N^d|} s(P[0,T,N\tilde{r},Nr])
\]

which is the analogue of (4.3.10), replacing relative entropy by relative entropy density.

### 4.5.6 Diffusion processes: computation of the Lagrangian.

For diffusion processes, let us start with the simplest case of independent diffusions in \( d = 1 \). The general case will be analogous, but the quadratic forms appearing there will be less explicit. The source generator \( Q \) is thus given by

\[
Qf(\sigma) = \frac{1}{2} \partial_0^2 f(\sigma)
\]

where we abbreviated \( \partial_0 \) to denote the partial derivative w.r.t. \( \sigma_0 \). As a consequence, for a local function \( f \):

\[
\mathcal{K}_Q f = \sum_k Q(\tau_k f) + \left( \sum_k \partial_0(\tau_k f) \right)^2
\]

and, reminding that the full generator is the sum of shifts of \( Q \), we have

\[
\mathcal{G}(\mu, f) = \int \mathcal{K}_Q f d\mu = \int \mathcal{L} f d\mu + \mathcal{J}_\mu(f, f)
\]

where

\[
\mathcal{J}_\mu(f, f) = \int \left( \sum_k \partial_0(\tau_k f) \right)^2 d\mu
\]
is a $\mu$ dependent quadratic form. This quadratic form is the analogue of (4.4.3). Hence, for the Lagrangian we have

$$\Xi(\mu, \dot{\mu}) = \sup_f (\langle \dot{\mu} - \mathcal{L}^* \mu, f \rangle - \mathcal{J}_\mu (f, f)) = \mathcal{J}_\mu^* (\mu - \mathcal{L}^* \mu, \dot{\mu} - \mathcal{L}^* \mu)$$

where $\mathcal{J}_\mu^*$ is a dual quadratic form defined via

$$\mathcal{J}_\mu^* (\nu, \nu) = \sup_f (\nu, f) - \mathcal{J}_\mu (f, f))$$

(4.5.15)

for $\nu$ a signed measure of total mass zero. Notice that this indeed defines a quadratic form because for $\lambda > 0$ (and with similar derivation for $\lambda < 0$)

$$\mathcal{J}_\mu^* (\lambda \nu, \lambda \nu) = \sup_f (\lambda (\nu, f) - \mathcal{J}_\mu (f, f))$$

$$= \lambda^2 \sup_f ((\nu, f/\lambda) - \mathcal{J}_\mu (f/\lambda, f/\lambda))$$

$$= \lambda^2 \mathcal{J}_\mu^* (\nu, \nu)$$

We see in particular that $\Xi(\mu, \dot{\mu})$ is zero for a solution of the Kolmogorov forward equation, i.e., if $\dot{\mu} = \mathcal{L}^* \mu$, which shows also in the present context that the Markovian evolution of the distribution $\mu$ is a zero cost trajectory.

Finally, let us turn to the general case. We split $Q$, the source generator, into a first order part and a second order part:

$$Q = Q_1 + Q_2$$

where $Q_2$ contains all second order derivatives (variance part of the diffusion), $Q_1$ all first order derivatives (drift part). To $Q_2$ is then associated the quadratic form

$$\mathcal{J}_\mu^Q (f, f) = \int \left( e^{-\sum_k \tau_k f} Q_2 e^{\sum_k \tau_k f} - Q_2 \left( \sum_k \tau_k f \right) \right) d\mu$$

(4.5.16)

The Lagrangian is then given by

$$\Xi(\mu, \dot{\mu}) = (\mathcal{J}_\mu^Q)^* (\dot{\mu} - \mathcal{L}^* \mu, \dot{\mu} - \mathcal{L}^* \mu)$$

(4.5.17)

where $(\mathcal{J}_\mu^Q)^*$ is the dual quadratic form of $\mathcal{J}_\mu^Q$ (as in (4.5.15)).
4. LARGE DEVIATIONS FOR THE TRAJECTORY OF THE
EMPIRICAL DISTRIBUTION AND EMPIRICAL MEASURE
Bibliography


