1. Motivations

Statistical Mechanics. Statistical Mechanics aims to give a global description of complex systems. As physical systems developed rather complex features such as phase transitions, it was an open question, even just a century ago, whether these features may fit in the Gibbs measure formalism and more generally within the Boltzmann approach. Indeed, the very first example developed mathematically, the Ising model, gave a positive answer. The ensuing mathematical formalism, known today as the Thermodynamic Formalism, has had immense influence in Mathematics and it has become possibly the most powerful and modern tool in Dynamical Systems.

However, one should not be tricked by the names. When compared say to the Ising model, the *Dynamics* in the name Dynamical Systems would refer to translations, space translation physically speaking. An effective, rigorous and quantitative approach to time evolution in Statistical Mechanics remains an elusive goal, somehow out of reach for modern science. Several, somehow unsatisfactory, approaches have been developed, such as introducing spurious random dynamics in the evolution of (even classical, nevermind quantum) complex (high dimensional) systems and studying singular and somehow simpler systems, or oversimplifying the nature and interaction of the phase space.

Regardless of the approach, the role of phase transitions and the existence of multiple equilibrium states (probability measures on the phase space), has clearly a great influence on the dynamical behavior of a physical systems. If we think of a (finite, but large) Ising system below the critical temperature, it should be clear that the system will spend a lot of time close to the support of a pure equilibrium state, then switch to another pure state on a time-scale which becomes longer and longer as the system size diverges. This somehow complex behavior is known as metastability (yet, metastability encompasses a larger set of cases; in the Ising model with positive magnetization, a similar behavior will happen when starting the system on the -1 state: we immediately see that dynamical phase transitions are a much more complex problem, which escapes a precise mathematical characterization).

Metastability phenomena have been the subject of tremendous study in the last century. For our part, we refer to the book by Bovier and DenHollander for some mathematical tools typical of Markov processes [1]. Roughly speaking (for instance in the Ising model), in the thermodynamic limit, a system should be described by a partition of the phase space (in the case of Ising, the supports of ergodic Gibbs measures -that are mutually singular), a longer time-scale under which the transitions between partitions happens (in the case of the Ising model, sharp results are a largely open question partially solved in $d = 2$ and basically zero temperature); a limiting Markov process describing the transitions between the partitions at this longer time-scale. Formally speaking, the metastability nature of the phenomenon is hidden in the Markov nature of the process (otherwise, the projection of any process on an arbitrary measurable partition is still a stochastic process!).

Earlier results mostly treated the time-scale problem (sharp spectral gap estimates and large deviations), which usually requires a keen understanding of the underlying partition of the phase space. In the last fifteen years, mostly via potential theory, some sharper results have been established for several (somehow

---

[1] Il libro in parte glissa sui lavori di Claudio Landim e compagni, in quanto non scorre buon sangue tra FrankD.H. e Claudio.

simple) models, even if mostly limited to a sharp asymptotic of the transition time rather than the convergence of the rescaled process. These type of results include finite-dimensional diffusions (small noise limit, sharp time asymptotic via capacity, Bovier et al.), zero-range processes (thermodynamic limit, capacity and evolution semigroup convergence, Landim et al.), stochastic quantization equation (Berglund et al, sharp time asymptotic, capacity estimates uniform in the Galerkin approximation).

The elephant in the room, from the Statistical Mechanics point of view, is that most of the tools have been developed for reversible processes, and the little results concerning non-reversible processes have been limited to simple models. As well known, reversible dynamics such as Metropolis, has a great deal of applications but in the words of Ulam (the quote is certainly misattributed) *Non-reversible Statistical Mechanics is like non-elephant Biology.*

Optimization in high dimensions. The classical problem of sampling requires a fast algorithm for extracting independent random variables with a given disitribution μ, say in $\mathbb{R}^d$. Due to the curse of dimensionality, except in trivial cases (such as product measures), sampling times diverge catastrophically for d large. A standard technique is rather to simulate a Markov process X with invariant measure μ up to a time T to be chosen. The process can be started in an arbitrary (maybe smartly chosen, but computationally simple point), and the value of X_T will be the sampled variable. As d grows, simulating X_T is much faster than exactly sampling μ. Of course, the price to pay is that the law μ_T of X_T will only be close to μ, hopefully exponentially close as T diverges, but not equal. It is then of great practical interest, to find a theorical distance decay $d(\mu_T, \mu) \simeq \exp(-cT)d(\mu_0, \mu)$, so that one can tune the simulation time T depending on the wanted error in the approximation. Running this type of algorithms, known as MonteCarlo Markov Chains (MCMC), sizes a considerable percentage of the world computing power, especially in supercomputers. With machine learning on the rise, the typical dimension at which problems are run has increased dramatically.

A similar problem comes more straightly from high dimensional optimization and ML techniques. In many cases, a part of the problems consists in minimizing a data-depending function depending on many parameters. As a trivial example, suppose we want to estimate the probability that tossing an unfair coin we get tail. We model it via a binomial distribution with unknown parameter p. If the data is that we got k tails out of n tossing, we chose p as to maximize $\binom{n}{k}p^k(1-p)^{n-k}$ which of course turns out to be $p = k/n$. In this case we had one parameter and one numerical datum k. However, in many fashionable techniques to approach a problem, the probability of an outcome depends on many (thousands or millions) of parameters, and the probability that the dataset 'happened' will be a complicated function p to maximize and depending on the parameters and a vast amount of data. A rather classical technique consists in sampling a random variable from the measure $\exp(-\beta V)/Z$ for β large and $V = -p$, and claim it is close to a minimizer of V. The measure is then sampled with MCMC: as one expect a faster convergence for smaller β, usually the simulations are run with larger and larger β's, using the results of a smaller β as initial datum for a larger β.

To be precice, optimizing strictly convex functions V can be done with complexity $O(d \log d)$, which is certainly not problematic. So the described techniques have an interest in the nonconvex case. Which again leads back to metastability.

Back to MCMC, we have not discussed the elephan in the room again. What process X should one use for the simulation? Well, the naive answer would be 'the one converging the fastest to μ'. This is certainly not the case, because simulating an i.i.d. Markov chain with marginal disitribution μ would solve the question, but not the problem. One should rather fix some measure of the complexity of the simulation, for instance the amount of fluctuations in it. So a first step in optimizing the choice of X would be to optimize it within the class giving rise to the same average quadratic variation of observables. Say in continuous times, this would be the quadratic variations of Dynking martingales, which happens to be $\mathbb{E}[Lf^2 - 2fLf]$, where L is the generator of the process. At the invariant measure, this coincides with the Dirichlet form $\mathcal{E}(f, f)$ which should be fixed (otherwise, one may just consider $Y_t = X_{2t}$ and speed-up the simulation -which of course does not work). So at the end of the day, we can first try to optimize among processes with the same Dirichlet form $\mathcal{E}(f, f)$, namely the same $L + L^*$. Ironically, in this class one expects that the (unuique) reversible process is the one converging the slowest, as noted by Rey-Bellet; and yet reversible processes are widely used in MCMC out of habits.

As far as I know, there are three classes of widely used MCMC: finite state Markov chains (at the very end, on a computer everything is finite state and discrete time, but not everything is properly estimated like that - as much as one uses integrals to estimate finite sums); diffusion processes; jump processes (possibly mixed with diffusions).

2. Three projects

With the motivations above, there are three projects we may develope.

- Sharp metastability results for mean-field models. Non-trivial nearest-neighbor interaction, or in any case interactions with geometry seems to be out of reach (maybe Kac potentials?). Yet, mean-field are likely the most used class of phenomenological theories in physics, finance and macroeconomy. Reversible vs non-reversible is of great interest here.
- Comparing symmetric vs non-symmetric rate of convergence. There are very few and partial results in the sense 'non-symmetric is faster'. We have studied the case of diffusions in the small noise limit with Claudio and Insuk, I do not know much more. Rey-Bellet papers are informal. I think one can make a simple counterexample to this: symmetric vs asymmetric NN random walk on $\mathbb{Z}_n = \mathbb{Z}/(n\mathbb{Z})$ -yet the expoential rate of convergence is the same. Yet, at least for diffusions, the situation is more interesting. Let us look at

$$\dot{X} = -BX + \dot{W} \tag{2.1}$$

where B is a matrix with eigenvalues having strictly positive real part. The invariant measure is Gaussian with covariance

$$A^{-1} := \int_0^\infty \exp(-Bt)\exp(-B^\dagger t)\, dt \tag{2.2}$$

It seems to me that $\dot{X} = -AX + \dot{W}$ (the reversible process with the same symmetirc Dirichlet form) converges slower to the equilibrium, exponentially slower. A diffusion roughly speaking is a superposition of small noise

and linear processes, so maybe this is a general fact. For instance, one may investigate when is it true that

$$\lim_t \tfrac{1}{t} \log \int dm(x) H(\nu_t^x|m) \le \lim_t \tfrac{1}{t} \log \int dm(x) \log H(\mu_t^x|m) \qquad (2.3) \qquad \boxed{\texttt{e:asfd}}$$

where H is the relative entropy, ν_t^x is the law of a Markov process X_t started at x, and μ_t is the law of the Markov process generated by the symmetric part of the the generator of X in $L_2(m)$.

- A potential theory for Markov processes has been developed for diffusions and partly for finite state Markov chains. A general theory for non-reversible processes is lacking.

The equation

$$\dot{X} = b(X) + \sqrt{2/\beta}\dot{W} \tag{2.4}$$

$\boxed{\texttt{e:sa}}$

has invariant measure $m = \exp(-V)dx$ if V solves

$$\tfrac{1}{\beta}|\nabla V|^2 + b \cdot \nabla V = \tfrac{1}{\beta}\Delta V + \operatorname{div}(b) \tag{2.5}$$

$\boxed{\texttt{e:visco}}$

If $b(x) = -Ax$, so that $V = \tfrac{1}{2}Sx \cdot x$ where S solves for $\beta = 1$ for all $x \in \mathbb{R}^d$

$$Sx \cdot Sx - Ax \cdot Sx = \operatorname{Tr}(S - A) \tag{2.6}$$

$\boxed{\texttt{e:fad}}$

So S solves iff $\operatorname{Tr}(S) = \operatorname{Tr}(A)$ and $SSx \cdot x = S^\dagger A$ namely

$$2S^2 = SA + A^\dagger S \tag{2.7}$$

$\boxed{\texttt{e:fasfd}}$

3. Symmetric vs Asymmetric

See the Mathematica file from the example described above.

If X_t is a Feller process with generator L, then

$$M_t^f(X) := f(X_t) - f(X_0) - \int_0^t (Lf)(X_s)\, ds \qquad (3.1) \quad \boxed{\texttt{e:fa}}$$

More in general, one can build exponential martingales as

$$E_t^t(X) := \exp(f(X_t) - f(X_0) - \int_0^t e^{-f(X_s)}(Le^f)(X_s)\, ds) \qquad (3.2) \quad \boxed{\texttt{e:fa}}$$

Diffusions are processes $L\phi(f) = \phi'(f)Lf + \phi''(f)\Gamma(f,f)$ where

$$\Gamma(f,f) = (L(f^2) - 2fLf)/2 \qquad (3.3) \quad \boxed{\texttt{e:gamma}}$$

for smooth ϕ. So for a diffusion $E_t^f(X) = \exp(M_t^f - \int_0^t \Gamma(f,f)(X_s)ds)$.

Maybe one may want to compare processes such that

$$\mathcal{E}_L(f) = \int d\pi(x)e^{-f(x)}(Le^f)(x) \qquad (3.4) \quad \boxed{\texttt{e:comp}}$$

are the same. For instance

$$Lf(x) = \int c(x,dy)(f(y) - f(x)) \qquad (3.5) \quad \boxed{\texttt{e:fa}}$$

$$\mathcal{E}_L(f) = \int d\pi(x)c(x,dy)(e^{f(y)-f(x)} - 1) \qquad (3.6) \quad \boxed{\texttt{e:comp2}}$$

Given L and L' pure jump generators with the same π, if $\mathcal{E}_L(f) = \mathcal{E}_{L'}(f)$ then for each $\varepsilon > 0$

$$\int d\pi(x)c(x,dy)(e^{\varepsilon f(y)-\varepsilon f(x)} - 1) = \int d\pi(x)c'(x,dy)(e^{\varepsilon f(y)-\varepsilon f(x)} - 1) \qquad (3.7) \quad \boxed{\texttt{e:comp3}}$$

which seems to mean that c and c' have the same symmetric part.

4. Mean Field

See the other file. Start from the last section of the file.

5. Potential Theory

Let (E, d) be a Polish space endowed with its Borel σ-algebra and a reference probability measure π. Denote by $L^2 = L^2(\pi)$ the space of square-integrable, real-valued functions defined in E. The norm of L^2 is represented by $\| \cdot \|$. Consider a Markov semigroup with generator $L \colon D(L) \to L^2(\pi)$, with domain $D(L) \subset L^2(\pi)$, see [**?**, Definition 1.8].

Denote by $\mathcal{C}$ a core for the generator L and assume that $\mathcal{C}$ is closed by multiplication. Denote by L^* the adjoint of L, and assume that $\mathcal{C}$ is also a core for L^*.

Denote by $C(E, D(L))$ the space of continuous functions $\xi \colon E \to D(L)$. For a function $\xi \in C(E, D(L))$, we represent $\xi(x) \in D(L)$, $x \in E$, by ξ_x and $\xi_x(y) \in E$, $y \in E$, by $\xi(x, y)$.

We assume some sort of sector condition: For each function $f \in D(L)$, there exists a finite constant C_f such that for all $\varphi \in C(E, D(L))$,

$$\left(\int \pi(dx) \, f(x) \, (L \, \varphi_x)(x) \right)^2 \leq C_f \int \pi(dx) \, (L \, \varphi_x^2)(x) \, . \tag{5.1}$$

03

If $\varphi(x, y) = g(y) - g(x)$, then this condition is nothing but the sector condition with $C_f = c_K D(f, f)$.

If $L = \Delta$ is the Laplacian on $\mathbb{R}^d$, then define $V(x) = (\nabla_y \varphi)(x, x)$. Then

$$V_i(x) = (\partial_{y_i} \varphi)(x, x)$$

$$\partial_{x_i} V_i(x) = (\partial_{x_i} \partial_{y_i} \varphi)(x, x) + (\partial_{y_i} \partial_{y_i} \varphi)(x, x)$$

$$\operatorname{div}(V)(x) = \sum_i \partial_{x_i} V_i(x) = (\Delta_y \varphi)(x, x) + \sum_i (\partial_{x_i} \partial_{y_i} \varphi)(x, x) = (L\varphi_x)(x)$$

$$(L \varphi_x^2)(x) = 2V(x) \cdot V(x)$$

Then the sector-like condition is

$$\left(\int dx \, f(x) \operatorname{div}(V(x)) \right)^2 \leq C_f \int dx \, 2V(x) \cdot V(x)$$

which holds since integrating by parts this is equivalent to

$$\left(\int dx \, (\nabla f)(x) \, V(x) \right)^2 \leq 2C_f \int dx \, V(x) \cdot V(x)$$

and therefore it is enough to take $C_f = D(f, f)$.

$$\Gamma(f, g) = L(fg) - fLg - gLf$$

$$D^s(f, g) = \int d\pi(x) \, \Gamma(g, f)$$

$$\mathbf{\Gamma}(\varphi, \psi)(x) = L(\varphi_x \psi_x)(x)$$

Notice that $\mathbf{\Gamma}(df, dg) = \Gamma(f, g)$ If $L = \Delta$, then $\mathbf{\Gamma}(\varphi, \psi) = 2V_\varphi V_\psi$

If $Lf = b \cdot df$, then $\mathbf{\Gamma}(f, g) = 0$. Then the condition

$$\left(\int \pi(dx) \, f(x) \, (L \varphi_x)(x) \right)^2 \leq C_f \int \pi(dx) \, (L \varphi_x^2)(x)$$

writes as

$$\left(\int \pi(dx) \, f(x) \, (\mathbf{L}\varphi)(x) \right)^2 \leq C_f \int \pi(dx) \, \mathbf{\Gamma}(\varphi, \varphi)(x)$$

If we restrict to the case $\varphi = dg$ then the last condition becomes

$$\left(\int \pi(dx) \, f(x) \, (Lg)(x) \right)^2 \leq C_f \int \pi(dx) \, \Gamma(g, g)(x) = 2C_f D(g, g)$$

which is implied by the sector condition.

Let us try to prove

$$\left(\int \pi(dx)\, f(x)\, (\mathbf{L}\varphi)(x) \right)^2 \leq C_f \int \pi(dx)\, \mathbf{\Gamma}(\varphi,\varphi)(x)$$

assuming L is self-adjoint in $L^2(d\pi)$. By self-adjointness

$$2 \int \pi(dx)\, f(x)\, (\mathbf{L}\varphi)(x) = -\int d\pi(x)\mathbf{\Gamma}(df,\varphi) \leq \sqrt{\int d\pi(x)\mathbf{\Gamma}(df,df)}\sqrt{\int d\pi(x)\mathbf{\Gamma}(\varphi,\varphi)}$$

The first equality should follow as

$$\int \pi(dx)\, f(x)\, (\mathbf{L}\varphi)(x) = \int \pi(dx) \lim_{t\to 0} \mathbb{E}_x \varphi(X_t, X_0) f(X_0)/t = \lim_{t\to 0} \mathbb{E}_\pi \varphi(X_t, X_0) f(X_0)/t$$

By self-adjointness the last quantity also equals

$$\lim_{t\to 0} \mathbb{E}_\pi \varphi(X_0, X_t) f(X_t)/t = -\lim_{t\to 0} \mathbb{E}_\pi \varphi(X_t, X_0) f(X_t)/t = -\lim_{t\to 0} \tfrac{1}{2}\mathbb{E}_\pi \varphi(X_t, X_0)(f(X_t)-f(X_0))/t$$

$$= -\tfrac{1}{2} \int \pi(dx)\, \mathbf{\Gamma}(df,\varphi)(x)$$

Probably it is enough to do it for pure jump processes. Namely, is it true that

$$\left(\int \pi(dx)c(x,dy)f(x)\varphi(x,y) \right)^2 \leq C_f \int \pi(dx)\, c(x,dy)\varphi(x,y)^2$$

if we know

$$\left(\int \pi(dx)c(x,dy)f(x)(g(y)-g(x)) \right)^2 \leq c\, D(f,f)D(g,g)$$

?

Given a Markov process, for each $t \geq 0$ it is defined a measurable $p_t \colon E \to \mathcal{P}(E)$, which we denote $p_t(x, dy)$, given by

$$\int p_t(x, dy) f(y) = \mathbb{E}_x[f(X_t)]$$

Moreover

$$\int_{y \in E} p_t(x, dy) p_s(y, dz) = p_{t+s}(x, dz)$$

Given a Markov process, with transition probability p_t, one can weakly approximate it, with a pure jump process with generator $L_\varepsilon f = \varepsilon^{-1} \int_y p_\varepsilon(x, dy)(f(y) - f(x))$. Notice that if the original Markov process admits a generator L, then

$$L f(x) = \lim_{t \to 0} \mathbb{E}_x[f(X_t) - f(X_0)]/t = \lim_{t \to 0} \int p_t(x, dy)(f(y) - f(x))/t = \lim_{\varepsilon \to 0} L_\varepsilon f(x)$$

Notice $c_\varepsilon(x, dy) = \varepsilon^{-1} p_\varepsilon(x, dy)$ has TV ε^{-1}

Puro salto: siano $k(dx, dy) = \pi(dx) c(x, dy)$ e $k^\dagger(dx, dy) = k(dy, dx)$, $k^s(dx, dy) = (k(dx, dy) + k^\dagger(dx, dy))/2$. In particolare $k \leq 2 \, k^s$. Sia

$$q(x, y) = \frac{k - k^\dagger}{k^s}(x, y) \in [-2, 2]$$

$$\left(\int \pi(dx) c(x, dy) f(x) \varphi(x, y) \right)^2 = \left(\int k(dx, dy) f(x) \varphi(x, y) \right)^2$$

$$= \left(\tfrac{1}{2} \int k(dx, dy) f(x) \varphi(x, y) - k(dy, dx) f(y) \varphi(x, y) \right)^2$$

$$= \left(\tfrac{1}{2} \int k^s(dx, dy)(f(x) - f(y)) \varphi(x, y) - \tfrac{1}{2} \int (k(dx, dy) - k(dy, dx)) f(x) \varphi(x, y) \right)^2$$

$$\leq 2 D(f, f) \mathbb{D}(\varphi, \varphi) + 4 \left(\int k^s(dx, dy) q(x, y) f(x) \varphi(x, y) \right)^2$$

Esempio: $E = \mathbb{T}_N$, $k^s(dx, dy) = 1/(2N)(\delta_{y-1}(dy) + \delta_{y+1}(dy))$, δ

Let $\mathcal{U}_0$, $\mathcal{A}$ be the set given by

$$\mathcal{U}_0 \;=\; \big\{\, \xi \in C(E, D(L)) : \xi_x(x) = 0 \ \; \forall\, x \in E \,\big\} \,.$$

$$\mathcal{A} \;=\; \big\{\, \xi \in \mathcal{U}_0 : \xi_x(y) = -\,\xi_y(x) \ \; \forall\, x \in E,\ \xi_x \in \mathcal{C} \,\big\} \,.$$

Define the operator $\mathbb{L} : \mathcal{U}_0 \to L^2(\pi)$ by

$$(\mathbb{L}\,\xi)(x) \;=\; (L\,\xi_x)(x)\,, \quad x \in E \,.$$

Denote by $\langle\!\langle\, \cdot\, ,\, \cdot\, \rangle\!\rangle$ the scalar product on $\mathcal{U}_0$ defined by

$$\langle\!\langle\, \xi\, ,\, \xi' \,\rangle\!\rangle \;:=\; \frac{1}{2} \int \pi(dx)\,(\mathbb{L}\,\xi\,\xi') \;=\; \frac{1}{2} \int \pi(dx)\,(L\,\xi_x\,\xi_x')(x) \,.$$

Note that $\langle\!\langle\, \xi\, ,\, \xi \,\rangle\!\rangle \ge 0$ because, since ξ vanish on the diagonal,

$$\langle\!\langle\, \xi\, ,\, \xi \,\rangle\!\rangle \;=\; \lim_{t\downarrow 0} \frac{1}{2} \int \pi(dx)\, \frac{(P_t\,\xi_x^2)(x) - \xi_x^2(x)}{t}$$

$$=\; \lim_{t\downarrow 0} \frac{1}{2} \int \pi(dx)\, \frac{(P_t\,\xi_x^2)(x)}{t} \;\ge\; 0 \,.$$

Denote by $\|\cdot\|$ the pre-norm associated to this scalar product and by $\sim$ the equivalence relation in $\mathcal{U}_0$ given by $\xi \sim \xi'$ whenever $\|\,\xi - \xi'\,\| = 0$. Let $\mathcal{U}$, $\mathcal{H}$ be the completions of $\mathcal{U}_0$, $\mathcal{A}$, respectively, with respect to the scalar product $\langle\, \cdot\, ,\, \cdot\, \rangle$: $\mathcal{U} = \overline{\mathcal{U}_0\big|_{\sim}}$, $\mathcal{H} = \overline{\mathcal{A}\big|_{\sim}}$.

For every function $f \in L^2(\pi)$, denote by $df : E \times E \to \mathbb{R}$ the function defined by $[(df)(x)](y) = f(y) - f(x)$. Clearly, if $f \in D(L)$, then df belongs to $\mathcal{A}$. Note that $\mathbb{L}(df) = Lf$ and that for every g in $D(L)$,

$$\langle\!\langle\, df\, ,\, dg \,\rangle\!\rangle \;=\; \frac{1}{2} \int \pi(dx)\,\big\{ L\,f\,g \,-\, f\,Lg \,-\, g\,Lf \,\big\}(x) \,.$$

As π is the stationary state, $\int \pi(dx)\,(L\,f\,g)(x) = 0$ so that

$$\langle\!\langle\, df\, ,\, dg \,\rangle\!\rangle \;=\; \frac{1}{2}\,\mathcal{D}(f,g) \,+\, \frac{1}{2}\,\mathcal{D}(g,f) \,. \tag{5.2}$$

10

In the particular case where $f = g$,

$$\big\|\, df \,\big\|_{\mathcal{H}}^2 \;=\; -\int \pi(dx)\,f(x)\,(Lf)(x) \;=\; \mathcal{D}(f,f) \,. \tag{5.3}$$

05

5.1. Markov Flows.

Denote by $\mathcal{F}$ the dual of $\mathcal{H}$. We refer to $\mathcal{F}$ as the space of flows.

For a function $f : E \to \mathbb{R}$, let $\Phi_f : \mathcal{H} \to \mathbb{R}$ be the linear functional defined by

$$\Phi_f(\chi) \;=\; -\int \pi(dx)\,f(x)\,(\mathbb{L}\,\chi)(x) \;=\; -\int \pi(dx)\,f(x)\,(L\,\chi_x)(x) \,.$$

By (??), for each $f \in D(L)$, Φ_f is a bounded functional and therefore belongs to $\mathcal{F}$.

Notice that for every functions f, g in $D(L)$,

$$\Phi_g(df) \;=\; -\int \pi(dx)\,g(x)\,(\mathbb{L}\,df)(x) \;=\; -\int \pi(dx)\,g(x)\,(L\,f)(x) \;=\; \mathcal{D}(f,g) \,. \tag{5.4}$$

phig

For a function f in $D(L)$. let Ψ_f the element of $\mathcal{F}$ defined by

$$\Psi_f(\chi) \;=\; \langle\!\langle\, df\,,\,\chi\,\rangle\!\rangle\,,\quad \chi\in\mathcal{H}\,. \tag{5.5}$$ 06

As df belongs to $\mathcal{H}$, Ψ_f is an element of $\mathcal{F}$ and

$$\|\Psi_f\|_{\mathcal{F}}^2 \;=\; \|df\|_{\mathcal{H}}^2 \;=\; \mathcal{D}(f,f)\,, \tag{5.6}$$ 07

where the last identity follows from (**??**). Indeed, by definition,

$$\|\Psi_f\|_{\mathcal{F}}^2 \;:=\; \sup_{\chi}\big\{\,2\,\Psi_f(\chi)\,-\,\|\chi\|_{\mathcal{H}}^2\,\big\} \;=\; \sup_{\chi\in\mathcal{H}}\big\{\,2\,\langle\!\langle\, df\,,\,\chi\,\rangle\!\rangle\,-\,\|\chi\|_{\mathcal{H}}^2\,\big\}\,,$$

and this expression is equal to $\|df\|_{\mathcal{H}}^2$. Moreover, for every function g in $D(L)$,

$$\Psi_f(dg) \;=\; \langle\!\langle\, df\,,\,dg\,\rangle\!\rangle \tag{5.7}$$ 09

Define the codifferential d^* on $\mathcal{F}$ as the dual operator of d, namely

$$(d^*\Phi)(f) \;:=\; -\,\Phi(df)\,.$$

Fix two disjoint, non-empty subsets A, B of E: $A\neq\varnothing$, $B\neq\varnothing$, $A\cap B=\varnothing$. Let $C_{\alpha,\beta}$, $\alpha,\beta\in\mathbb{R}$, be the subspace of real continuous functions given by

$$C_{\alpha,\beta} \;:=\; \big\{\, f\in C(E)\,:\, f(x)\,=\,\alpha\ \text{for } x\in A \text{ and } f(y)\,=\,\beta\ \text{for } y\in B,\, f\in D(\mathcal{D}^s)\,\big\}\,.$$

Let $\mathcal{F}_\gamma$, $\gamma\in\mathbb{R}$, be given by

$$\mathcal{F}_\gamma \;:=\; \big\{\,\Phi\in\mathcal{F}\,:\, d^*\Phi(f)\,=\,-\gamma\ \text{for all } f\in C_{1,0}\,\big\}\,. \tag{5.8}$$ fgamma

5.2. **Capacity.** A closed set $A\subset E$ is recurrent if for all $x\in E$, $\mathbb{P}_x(\tau_A<\infty)=1$.

def1 **Definition 5.1.** *Two closed recurrent sets A, $B\subset E$ have* finite capacity *if the function $h(x)\equiv h_{A,B}(x)=\mathbb{P}_x(\tau_A<\tau_B)$ is in the domain of $\mathcal{D}(\cdot)$ or equivalently $dh\in\mathcal{H}$. In such a case we define*

$$\mathrm{cap}(A,B)=\mathcal{D}(h,h) \tag{5.9}$$ 08

102 **Lemma 5.2.** *Fix two disjoint, non-empty subsets A, B of E.*

$$\mathrm{cap}(A\,,\,B) \;=\; \mathrm{cap}^\dagger(A\,,\,B)\,.$$

101 **Lemma 5.3.** *Fix two disjoint, non-empty subsets A, B of E. For each $\alpha,\gamma\in\mathbb{R}$, $f\in C_{\alpha,0}$ and $\Phi\in\mathcal{F}_\gamma$,*

$$\big[\,\Phi_f\,-\,\Phi\,\big](dh_{A,B}) \;=\; \alpha\,\mathrm{cap}(A,B)\,-\,\gamma\,. \tag{5.10}$$ 04

Proof. From (**??**), we have that

$$\Phi_f(dh) \;=\; -\int \pi(dx)\,f(x)\,(L\,h)(x)\,.$$

Since LH vanishes on $\Omega=(A\cup B)^c$ and $f=\alpha, h$ 0 on $A\cup B$, we may replace in the previous formula, f by $\alpha\,h$ to get that $\Phi_f(dh)=\alpha\,\mathcal{D}(h,h)=\alpha\,\mathrm{cap}(A,B)$.

On the other hand, by the definition (**??**) of $\mathcal{F}_\gamma$ and since h belongs to $C_{1,0}$,

$$\Phi(dh) \;=\; \gamma\,.$$

This proves the lemma. $\qquad\qquad\square$

prop1 **Proposition 5.4** (Dirichlet Principle)**.** *Fix two disjoint, non-empty subsets A, B of E. We have that*

$$\mathrm{cap}(A.B) \;=\; \inf_{f\in C_{1,0}}\,\inf_{\Phi\in\mathcal{F}_0}\,\|\Phi_f-\Phi\|_{\mathcal{F}}^2\,.$$

Proof. Fix f in $C_{1,0}$ and Φ in $\mathcal{F}_0$. By the previous lemma and by Schwarz inequality,

$$\mathrm{cap}(A,B)^2 \;=\; \left\{ \left[\, \Phi_f \,-\, \Phi \,\right] (dh_{A,B}) \right\}^2 \;\leq\; \left\| \Phi_f - \Phi \right\|_{\mathcal{F}}^2 \left\| dh \right\|_{\mathcal{H}}^2 .$$

By (**??**) and (**??**), the last term is equal to $\mathrm{cap}(A,B)$, so that $\mathrm{cap}(A,B) \leq \| \Phi_f - \Phi \|_{\mathcal{F}}^2$ for every f in $C_{1,0}$ and Φ in $\mathcal{F}_0$:

$$\mathrm{cap}(A,B) \;\leq\; \inf_{f \in C_{1,0}} \inf_{\Phi \in \mathcal{F}_0} \| \Phi_f - \Phi \|_{\mathcal{F}}^2 .$$

Let $f_\star = (h + h^\dagger)/2$ and $\Phi_\star = \Phi_{f_\star} - \Psi_h$ so that $\Phi_{f_\star} - \Phi_\star = \Psi_h$. By (**??**) and (**??**), $\mathrm{cap}(A,B) = \mathcal{D}(h,h) = \| \Psi_h \|_{\mathcal{F}}^2 = \| \Phi_{f_\star} - \Phi_\star \|_{\mathcal{F}}^2$. Hence, to prove that

$$\inf_{f \in C_{1,0}} \inf_{\Phi \in \mathcal{F}_0} \| \Phi_f - \Phi \|_{\mathcal{F}}^2 \;\leq\; \mathrm{cap}(A,B) ,$$

it remains to check that $f_\star$ belongs to $C_{1,0}$, and $\Phi_\star$ to $\mathcal{F}_0$.

Clearly, $f_\star = (h + h^\dagger)/2 \in C_{1,0}$. To shows that $\Phi_\star \in \mathcal{F}_0$, fix g in $C_{1,0}$. By definition, $\Phi_\star(dg) = \Phi_{f_\star}(dg) - \Psi_h(dg)$. By definition of $f_\star$, (**??**), (**??**) and (**??**), this expression is equal to

$$\frac{1}{2} \left\{ \mathcal{D}(h\,,g) + \mathcal{D}(h^\dagger,g) - \mathcal{D}(h\,,g) - \mathcal{D}(g\,,h) \right\} \;=\; \frac{1}{2} \left\{ \mathcal{D}(h^\dagger,g) - \mathcal{D}(g\,,h) \right\} .$$

As g belongs to $C_{1,0}$, g coincides with h amd $h^\dagger$ on $A \cup B$. Hence, as $Lh = L^\dagger h^\dagger = 0$ on Ω, the previous expression is equal to

$$\frac{1}{2} \left\{ - \langle L^\dagger h^\dagger, h^\dagger \rangle + \langle Lh\,, h \rangle \right\} \;=\; \frac{1}{2} \left\{ \mathcal{D}(h^\dagger, h^\dagger) - \mathcal{D}(h\,, h) \right\} \;=\; 0$$

because both quantities are equal to $\mathrm{cap}(A,B)$. $\qquad\square$

Proposition 5.5 (Thomson Principle)**.** *Fix two disjoint, non-empty subsets A, B of E. We have that*

$$\frac{1}{\mathrm{cap}(A.B)} \;=\; \inf_{\Phi \in \mathcal{F}_1} \inf_{f \in C_{0,0}} \| \Phi_f - \Phi \|_{\mathcal{F}}^2 .$$

Proof. Fix f in $C_{0,0}$ and Φ in $\mathcal{F}_1$. By the previous lemma and by Schwarz inequality,

$$1 \;=\; \left\{ \left[\, \Phi_f - \Phi \,\right] (dh_{A,B}) \right\}^2 \;\leq\; \left\| \Phi_f - \Phi \right\|_{\mathcal{F}}^2 \left\| dh \right\|_{\mathcal{H}}^2 .$$

By (**??**) and (**??**), the last term is equal to $\mathrm{cap}(A,B)$, so that $1/\mathrm{cap}(A,B) \leq \| \Phi_f - \Phi \|_{\mathcal{F}}^2$ for every f in $C_{0,0}$ and Φ in $\mathcal{F}_1$:

$$\frac{1}{\mathrm{cap}(A.B)} \;\leq\; \inf_{\Phi \in \mathcal{F}_1} \inf_{f \in C_{0,0}} \| \Phi_f - \Phi \|_{\mathcal{F}}^2 .$$

To complete the proof of the theorem, it remains to find g in $C_{0,0}$ and Φ in $\mathcal{F}_1$ such that $\| \Phi_g - \Phi \|^2 = 1/\mathrm{cap}(A,B)$. Let $g_\star = (h^\dagger - h)/2\,\mathrm{cap}(A,B)$ and $\Psi_\star = \Phi_{g_\star} + \Psi_{h_0}$, where $h_0 = h/\mathrm{cap}(A,B)$, so that $\Phi_{g_\star} - \Psi_\star = - \Psi_{h_0}$. By (**??**) and (**??**), $1/\mathrm{cap}(A,B) = \mathcal{D}(h_0,h_0) = \| \Psi_{h_0} \|_{\mathcal{F}}^2 = \| \Phi_{g_\star} - \Psi_\star \|_{\{}^2$. Hence, to prove that

$$\inf_{f \in C_{1,0}} \inf_{\Phi \in \mathcal{F}_0} \| \Phi_f - \Phi \|_{\mathcal{F}}^2 \;\leq\; \mathrm{cap}(A.B) ,$$

it remains to check that $g_\star$ belongs to $C_{0,0}$, and $\Psi_\star$ to $\mathcal{F}_1$.

Clearly, $g_\star = (h^\dagger - h)/2\,\mathrm{cap}(A,B) \in C_{0,0}$. To shows that $\Psi_\star \in \mathcal{F}_1$, fix f in $C_{1,0}$. By definition, $\Psi_\star(df) = \Phi_{g_\star}(df) + \Psi_{h_0}(df)$. By definition of $g_\star$, (**??**), (**??**) and (**??**), this expression is equal to

$$\frac{1}{2\,\mathrm{cap}(A,B)}\left\{ \mathcal{D}(f\,,\,h^\dagger) \,-\, \mathcal{D}(f,h) \,+\, \mathcal{D}(h\,,\,f) \,+\, \mathcal{D}(f\,,\,h) \right\}$$

$$= \frac{1}{2\,\mathrm{cap}(A,B)}\left\{ \mathcal{D}(f\,,\,h^\dagger) \,+\, \mathcal{D}(h\,,\,f) \right\}\,.$$

As f belongs to $C_{1,0}$, f coincides with h and $h^\dagger$ on $A\cup B$. Hence, as $Lh = L^\dagger h^\dagger = 0$ on Ω, the previous expression is equal to

$$\frac{1}{2\,\mathrm{cap}(A,B)}\left\{ -\,\langle L^\dagger h^\dagger\,,\,h^\dagger\rangle \,-\, \langle Lh\,,\,h\rangle \right\}$$

$$= \frac{1}{2\,\mathrm{cap}(A,B)}\left\{ \mathcal{D}(h^\dagger\,,\,h^\dagger) \,+\, \mathcal{D}(h\,,\,h) \right\} \,=\, 1$$

because both quantities are equal to $\mathrm{cap}(A,B)$. $\qquad\square$

5.3. **Examples. Discrete spaces.**

The operator $\mathbb{L} : \mathcal{A} \to E$ is given by

$$(\mathbb{L}\,\xi)(x) \,=\, (L\,\xi_x)(x)\,, \quad x \in E$$

$$= \sum_{y \in E} r(x,y)\,[\xi_x(y) - \xi_x(x)]$$

$$= \sum_{y \in E} r(x,y)\,\xi(x,y)\,.$$

The operator $L^\eta : D(L^\eta) \to B$ given by

$$(L^\eta f)(x) \,=\, [\,\mathbb{L}\,(e^\eta\,df)\,](x);= \sum_{y \in E} r(x,y)\,e^{\eta(x,y)}\,[f(y) - f(x)]\,.$$

$$\langle\!\langle\,\xi\,,\,\xi'\,\rangle\!\rangle \,:=\, \frac{1}{2}\int_{E\times E}\pi(dx)\,R(x,dy)\,\xi(x,y)\,\xi'(x,y) \,=\, \frac{1}{2}\int\pi(dx)\,(\mathbb{L}\xi\,\xi')\,.$$

$$\langle\,\xi\,,\,\xi'\,\rangle \,=\, \frac{1}{2}\sum_{x,y}\pi(x)r_s(x,y)\,\xi(x,y)\,\xi'(x,y)\,.$$

$$\Phi_f(\chi) \,=\, \int\pi(dx)\,f(x)\,(\mathbb{L}\chi)(x)\,.$$

5.4. **Diffusions.**

Fix $d \geq 1$, and denote by $\mathbb{T}^d = [0,1)^d$ the d-dimensional torus of length 1. Denote by $a(x)$ a uniformly positive-definite matrix whose entries $a_{i,j}$ are smooth functions: There exist $c_0 > 0$ such that for all $x \in \mathbb{T}^d$, $y \in \mathbb{R}^d$,

$$y \cdot a(x)\,y \,\geq\, c_0\,\|y\|^2\,, \tag{5.11}$$

19

where $y \cdot z$ represents the scalar product in $\mathbb{R}^d$.

Generator. Denote by $\mathcal{L}$ the generator given by

$$\mathcal{L}f \,=\, \nabla\cdot(a\,\nabla f) \,+\, b\cdot\nabla f\,, \tag{5.12}$$

2-12

where $b : \mathbb{T}^d \to \mathbb{R}^d$ is a smooth vector field. By modifying the drift b we could assume the matrix a to be symmetric. We will not assume this condition for reasons which will become below. There exists a unique Borel probability measure μ on $\mathbb{T}^d$

such that $\mu\mathcal{L} = 0$. This measure is absolutely continuous, $\mu(dx) = m(x)dx$, where m is the unique solution to

$$\nabla \cdot (a^\dagger \nabla m) \;-\; \nabla \cdot (b\,m) \;=\; 0 \;, \qquad (5.13)$$

where $a^\dagger$ stands for the transpose of a. For existence, uniqueness and regularity conditions of solutions of elliptic equations, we refer to [?]. Let $V(x) = -\log m(x)$, so that $m(x) = e^{-V(x)}$.

We may rewrite the generator $\mathcal{L}$ introduced in (??) as

$$\mathcal{L}f \;=\; e^V \nabla \cdot \left(e^{-V} a \nabla f\right) \;+\; c \cdot \nabla f \;,$$

where $c = b + a^\dagger \nabla V$. It follows from (??) that

$$\nabla \cdot (e^{-V} c) \;=\; 0 \;. \qquad (5.14)$$

This implies that the operator $c \cdot \nabla$ is skew-adjoint in $L_2(\mu)$: for any smooth functions $f,\, g : \mathbb{T}^d \to \mathbb{R}$,

$$\int f\, c \cdot \nabla g\, d\mu \;=\; -\int g\, c \cdot \nabla f\, d\mu \;. \qquad (5.15)$$

In view of (??), the adjoint of $\mathcal{L}$ in $L_2(\mu)$, represented by $\mathcal{L}^*$, is given by

$$\mathcal{L}^* f \;=\; e^V \nabla \cdot \left(e^{-V} a^\dagger \nabla f\right) \;-\; c \cdot \nabla f \;,$$

while the symmetric part, denoted by $\mathcal{L}^s$, $\mathcal{L}^s = (1/2)(\mathcal{L} + \mathcal{L}^*)$, takes the form

$$\mathcal{L}^s f \;=\; e^V \nabla \cdot \left(e^{-V} a_s \nabla f\right) \;. \qquad (5.16)$$

where a_s stands for the symmetrization of the matrix a: $a_s = (1/2)[a + a^\dagger]$.

Recall from (??), (??), the definition of the spaces $C(\mathbb{T}^d, D(L))$, $\mathcal{A}$ and the operator $\mathbb{L}$. In this context, for any smooth element φ of $C(\mathbb{T}^d, D(L))$,

$$\langle\!\langle \varphi\,,\, \varphi \rangle\!\rangle \;=\; \int_{\mathbb{T}^d} \mu(dx)\, (\nabla_y \varphi)(x,x) \cdot a(x)\, (\nabla_y \varphi)(x,x) \;. \qquad (5.17)$$

Hence, as the matrix a is strictly elliptic, $\|\varphi\| = 0$ if and only if $(\nabla_y \varphi)(x,x) = 0$ for all $x \in \mathbb{T}^d$.

Consider a smooth, conservative vector field $\mathfrak{v} : \mathbb{T}^d \to \mathbb{R}^d$. Recall that conservative means that its line integral over closed paths vanishes or, equivalently, that its line integral is path independent. Denote by $\varphi_\mathfrak{v} : \mathbb{T}^d \times \mathbb{T}^d \to \mathbb{R}$ the function defined by

$$\varphi_\mathfrak{v}(x,y) \;=\; \int_\gamma \mathfrak{v}(\ell)\, d\ell \;,$$

where γ is a path from x to y. The function $\varphi_\mathfrak{v}$ is well defined because $\mathfrak{v}$ is conservative. The same property yields that $\varphi_\mathfrak{v}(x,y) = -\varphi_\mathfrak{v}(y,x)$ so that $\varphi_\mathfrak{v} \in \mathcal{A}$. Furthermore, $(\nabla_y \varphi_\mathfrak{v})(x,x) = \mathfrak{v}(x)$ so that, by (??),

$$\langle\!\langle \varphi_\mathfrak{v}\,,\, \varphi_\mathfrak{v} \rangle\!\rangle \;=\; \int_{\mathbb{T}^d} \mu(dx)\, \mathfrak{v}(x) \cdot a(x)\, \mathfrak{v}(x) \;. \qquad (5.18)$$

Denote by $\mathcal{V}$ the closure of this set in $\mathcal{A}$.

We claim that $\mathcal{H} = \mathcal{V}$. Indeed, since $\mathcal{V}$ is contained in $\mathcal{H}$, it is enough to show that for any smooth φ in $\mathcal{A}$ there exists a conservative vector field $\mathfrak{v}$ such that $(\nabla_y \varphi)(x,x) = (\nabla_y \varphi_\mathfrak{v})(x,x)$ for all $x \in \mathbb{T}^d$. Fix such a smooth function φ in $\mathcal{A}$, define the vector field $\mathfrak{v}$ by $\mathfrak{v}(x) = (\nabla_y \varphi)(x,x)$. It is clear that $(\nabla_y \varphi)(x,x) = (\nabla_y \varphi_\mathfrak{v})(x,x)$, which proves the claim. Thus, in the context of diffusions, the space $\mathcal{H}$ can be identified with the space of vector fields.

Jump processes.

Consider a family of jump rates $r_\epsilon(x, dy)$ on $\mathbb{R}^d$. Denote by $\pi_\epsilon(dx)$ the invariant measure. Let $\lambda_\epsilon(x) = r_\epsilon(x, \mathbb{R}^d) \in [0, \infty]$.

$$\langle\!\langle \xi, \xi' \rangle\!\rangle = \frac{1}{2} \int \pi(dx)\, (\mathbb{L}\xi\, \xi') := \frac{1}{2} \int_{E \times E} \pi(dx)\, R(x, dy)\, \xi(x, y)\, \xi'(x, y) .$$

$$\text{cap}(A, B) = \inf$$

$$\Phi_f(\xi) = \int \pi(dx)\, f(x)\, (\mathbb{L}\,\xi)(x)$$

$$= -\int \pi(dx)\, r(x, dy)\, f(x)\, \xi(x, y)$$

$$= \int \pi(dx)\, r_s(x, dy) \left(\frac{f(y) - f(x)}{2} \right) \xi(x, y)$$

$$- \int \pi(dx)\, r_a(x, dy) \left(\frac{f(y) + f(x)}{2} \right) \xi(x, y)$$

Now

$$r_a(x, dy) = \rho(x, y)\, r_s(x, dy)$$

last line

$$- \int \pi(dx)\, r_s(x, dy)\, \rho(x, y) \left(\frac{f(y) + f(x)}{2} \right) \xi(x, y)$$

$$(L_{\epsilon,\beta} f)(x) = \int \mu_\epsilon(x, dy)\, \frac{1}{2} [\, f(y) - f(x)\,]\, e^{-\beta q(x, y)} ,$$

where $q(x, y) = V(y) - V(x) + c(x, y)$ and c is symmetric and $\mu_\epsilon(x, dy) = [\mu(x + \epsilon dy) + \mu(x - \epsilon dy)]/2$.

5.5. **remains.** Let $\mathcal{B}$ be the set given by

$$\mathcal{B} \;=\; \left\{\eta \in C(E, D(L)) : e^{\eta} - 1 \in \mathcal{A}\right\}.$$

For $\eta \in \mathcal{B}$, define the operator $L^{\eta} : D(L^{\eta}) \to B$ given by

$$(L^{\eta} f)(x) \;=\; [\,\mathbb{L}\,(e^{\eta}\, df\,)\,](x)\,.$$

$$\eta \sim \eta' \quad if \quad L^{\eta} = L^{\eta'}$$