A large-deviation view on dynamical Gibbs-non-Gibbs transitions

A.C.D. van Enter [1]
R. Fernández [2]
F. den Hollander [3]
F. Redig [4]

May 2, 2010

Abstract

We develop a space-time large-deviation point of view on Gibbs-non-Gibbs transitions in spin systems subject to a stochastic spin-flip dynamics. Using the general theory for large deviations of functionals of Markov processes outlined in Feng and Kurtz [11], we show that the trajectory under the spin-flip dynamics of the empirical measure of the spins in a large block in $\mathbb{Z}^d$ satisfies a large deviation principle in the limit as the block size tends to infinity. The associated rate function can be computed as the action functional of a Lagrangian that is the Legendre transform of a certain non-linear generator, playing a role analogous to the moment-generating function in the Gärtner-Ellis theorem of large deviation theory when this is applied to finite-dimensional Markov processes. This rate function is used to define the notion of "bad empirical measures", which are the discontinuity points of the optimal trajectories (i.e., the trajectories minimizing the rate function) given the empirical measure at the end of the trajectory. The dynamical Gibbs-non-Gibbs transitions are linked to the occurrence of bad empirical measures: for short times no bad empirical measures occur, while for intermediate and large times bad empirical measures are possible. A future research program is proposed to classify the various possible scenarios behind this crossover, which we refer to as a "nature-versus-nurture" transition.

MSC2010: Primary 60F10, 60G60, 60K35; Secondary 82B26, 82C22.
Key words and phrases: Stochastic spin-flip dynamics, Gibbs-non-Gibbs transition, empirical measure, non-linear generator, Nisio control semigroup, large deviation principle, bad configurations, bad empirical measures, nature versus nurture.
Acknowledgment: The authors are grateful for extended discussions with Christof Külske. Part of this research was supported by the Dutch mathematics cluster *Nonlinear Dynamics of Natural Systems*. RF is grateful to NWO (Netherlands) and CNRS (France) for financial support during his sabbatical leave from Rouen University in the academic year 2008–2009, which took place at Groningen University, Leiden University and EURANDOM. In the Fall of 2008 he was EURANDOM-chair.

[1] Johann Bernoulli Institute of Mathematics and Computer Science, University of Groningen, P.O. Box 407, 9700 AK, Groningen, The Netherlands, *A.C.D.van.Enter@rug.nl*

[2] Department of Mathematics, Utrecht University, P.O. Box 80010, 3508 TA Utrecht, The Netherlands, *R.Fernandez1@uu.nl*

[3] Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA, Leiden, The Netherlands, *denholla@math.leidenuniv.nl*

[4] IMAPP, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands, *f.redig@math.ru.nl*

1 Introduction, main results and research program

1.1 Dynamical Gibbs-non-Gibbs transitions

Since the discovery of the Griffiths-Pearce-Israel pathologies in renormalization-group transformations of Gibbs measures, there has been an extensive effort towards understanding the phenomenon that a simple transformation of a Gibbs measure may give rise to a non-Gibbs measure, i.e., a measure for which no reasonable Hamiltonian can be defined (see van Enter, Fernández and Sokal [4], Fernández [12], and the papers in the EURANDOM workshop proceedings [26]). From the start, R.L. Dobrushin was interested and involved in this development; indeed, Dobrushin and Shlosman [2], [3] proposed a programme of *Gibbsian restoration*, based on the idea that the pathological *bad configurations* of a transformed Gibbs measure (i.e., the essential points of discontinuity of some of its finite-set, e.g. single-site, conditional probabilities) are exceptional in the measure-theoretic sense (i.e., they form a set of measure zero). This has led to two extended notions of Gibbs measures: *weakly Gibbsian* measures and *almost Gibbsian* measures (see Maes, Redig and Van Moffaert [21]). Later, several refined notions were proposed, such as *intuitively weakly Gibbs* (Van Enter and Verbitskiy [9]) and right-continuous conditional probabilities.

In Van Enter, Fernández, den Hollander and Redig [5], the behavior of a Gibbs measure μ subject to a high-temperature Glauber spin-flip dynamics was considered. A guiding example is the case where we start from the low-temperature plus-phase of the Ising model, and we run a high-temperature dynamics, modeling the fast heating up of a cold system. The question of Gibbsianness of the measure μ_t at time $t > 0$ can then be interpreted as the existence of a reasonable notion of an *intermediate-time-dependent temperature* (at time $t = 0$ the temperature is determined by the choice of the initial Gibbs measure, while at time $t = \infty$ the temperature is determined by the unique stationary measure of the dynamics). For infinite-temperature dynamics, the effect of the dynamics is simply that of a single-site Kadanoff transformation, with a parameter that depends on time. The extension to high-temperature dynamics was achieved with the help of a *space-time cluster expansion* developed in Maes and Netocný [22]. The basic picture that emerged from this work was the following:

(1) μ_t is Gibbs for small t;

(2) μ_t is non-Gibbs for intermediate t;

(3) in zero magnetic field μ_t remains non-Gibbs for large t, while in non-zero magnetic field μ_t becomes Gibbs again for large t.

Further research went into several directions and, roughly summarized, gave the following results:

(a) Small-time conservation of Gibbsianness is robust: this holds for a large class of spin systems and of dynamics, including discrete spins (Le Ny and Redig [18]), continuous spins (Dereudre and Roelly [1], van Enter, Külske, Opoku and Ruszel, [15], [7], [8], [23], [6]), which can be subjected to Glauber dynamics, mixed Glauber/Kawasaki dynamics, and interacting-diffusion dynamics, not even necessarily Markovian (Redig, Roelly and Ruszel [25]), appliedto a large class of initial measures (e.g. Gibbs measures for a finite-range or an exponentially decaying interaction potential).

(b) Gibbs-non-Gibbs transitions can also be defined naturally for mean-field models (see e.g. Külske and Le Ny [14] for Curie-Weiss models subject to an independent spin-flip dynamics). In this context, much more explicit results can be obtained: transitions are sharp (i.e., in zero magnetic field there is a single time after which the measure becomes non-Gibbs and stays non-Gibbs forever, and in non-zero magnetic field there is a single time at which it becomes Gibbs again). Bad configurations can be characterized explicitly (with the interesting effect that non-neutral bad configurations can arise below a certain critical temperature). For further developments on mean-field results see also [16], [10].

(c) Gibbs-non-Gibbs transitions can also occur for continuous unbounded spins subject to independent Ornstein-Uhlenbeck processes (Külske and Redig [17]), and for continuous bounded spins subject to independent diffusions (Van Enter and Ruszel [7], [8]), even in two dimensions where no static phase transitions occur.

Bad configurations can be detected by looking at a so-called *two-layer system:* the joint distribution of the configuration at time $t = 0$ and time $t > 0$. If we condition on a particular configuration η at time $t > 0$, then the distribution at time $t = 0$ is a Gibbs measure with an η-dependent Hamiltonian H^η, which is a *random-field* modification of the original Hamiltonian H of the starting measure. If, for some η, H^η has a phase transition, then this η is a bad configuration (see Fernández and Pfister [13]).

1.2 Nature versus nurture

While these results led to a reasonably encompassing picture, we were unsatisfied with the strategy of the proofs for the following reason. All proofs rely on two fortunate facts: (1) the evolutions can be described in terms of space-time interactions; (2) these interactions correspond to well-studied models in equilibrium statistical mechanics. In particular, although the most delicate part of the analysis – the proof of the onset of non-Gibbsianness – was accomplished by adapting arguments developed in previous studies on renormalization transformations, the actual intuition that led to these results relied on entirely different arguments, based on the behavior of *conditioned trajectories*. These intuitive arguments, already stated without proof in our original work [5], can be summarized as follows:

(I) If a configuration η is *good* at time t (i.e., is a point of continuity of the single-site conditional probabilities), then the trajectory that leads to η is unique, in the sense that there is a single distribution at time $t = 0$ that leads to η at time $t > 0$. In particular, if t is small, then this trajectory stays close to η during the whole time interval $[0, t]$.

(II) If a configuration η is *bad* at time t (i.e., is a point of essential discontinuity of the single-site conditional probabilities), then there are at least two trajectories compatible with the occurrence of η at time t. Moreover, these trajectories can be selected by modifying the bad configuration η arbitrarily far away from the origin.

(III) Trajectories ending at a configuration η at time t are the result of a competition between two mechanisms:

- *Nature:* The initial configuration is close to η, which is not necessarily typical for the initial measure, and is preserved by the dynamics up to time t.

– *Nurture:* The initial configuration is typical for the initial measure and the system builds η in a short interval prior to time t.

As an illustration, let us consider the low-temperature zero-field Ising model subject to an independent spin-flip dynamics. In [5] we proved that the fully alternating configuration becomes and stays bad for large t. This fact can be understood according to the preceding paradigm in the following way. Short times do not give the system occasion to perform a large number of spin-flips. Hence, the most probable way to see the alternating configuration at small time t is when the system started in a zero-magnetization-like state and the evolution kept the magnetization zero up to time t. *This is the nature-scenario!* For larger times t, a less costly alternative is to start in a less atypical manner, and to arrive at the alternating configuration following a trajectory that stays close for as long as possible to the unconditioned dynamical relaxation. *This is the nurture-scenario!* In this situation, we can start either from a plus-like state or a minus-like state, as the difference in probabilistic cost between these two initial states is exponential in the size of boundary, and thus is negligible with respect to the volume cost imposed by a constrained dynamics. It is then possible to select between the plus-like and the minus-like trajectories by picking the alternating configuration in a large block, then picking either the all-plus or the all-minus configuration outside this block, and letting the block size tend to infinity.

We see that the above explanation relies on two facts:

(i) The existence of a *nature-versus-nurture* transition, as introduced in [5].

(ii) The existence of several possible trajectories (once the system is in the nurture regime), all starting from configurations that are typical for the initial measure (modulo an boundary-exponential cost). These trajectories evolve to the required bad configuration over a short interval prior to time t.

1.3 Large deviations of trajectories

The goal of the present paper is to put rigor into the above qualitative suggestions. We propose two novel aspects:

(1) the development of a suitable large deviation theory for *trajectories*, in order to estimate the costs of the different dynamical strategies;

(2) the use of *empirical measures* instead of *configurations*, in order to express the conditioning at time t.

For a translation-invariant spin-flip dynamics and a translation-invariant initial measure, nothing is lost by moving to the empirical measure because the bad configurations form a translation-invariant set. Instead, a lot is gained because, as we will show, the trajectory of the empirical measure satisfies a *large deviation principle* under quite general conditions on the spin-flip rates (e.g. there is no restriction to high temperature). Moreover, the question of uniqueness versus non-uniqueness of optimal trajectories (i.e., minimizers of the large deviation rate function) can be posed and tackled for a large class of dynamics, which places the dynamical Gibbs-non-Gibbs-transition into a framework where it gains more physical relevance.

Here is a list of the results presented in the sequel.

(A) *Existence of a large deviation principle for trajectories.* We apply the theory developed in Feng and Kurtz [11], Section 8.6. The rate function is the integral of the Legendre transform of the generator of the non-linear semigroup defined by the dynamics. In suitably abstract terms, this generator can be associated to a Hamiltonian, and the rate function to the integral of a Lagrangian (Sections 2–5).

(B) *Explicit expression for the generator of the non-linear semigroup of the dynamics.* These are obtained in Theorems 3.1–3.2 below (Section 3).

(C) *Rate functions for trajectories and associated optimal trajectories.* The general Legendre-transform prescription is explicitly worked out for a couple of simple examples, and optimal trajectories are exhibited (Sections 4.2–4.3).

(D) *Relation with thermodynamic potentials.* Relations are shown between the non-linear generator and the derivative of a "constrained pressure". Similarly, the rate function per unit time is related to the Legendre transform of this pressure (Section 5.2).

(E) *Definition of bad measures.* This definition, introduced in Section 6, is the transcription to our more general framework of the notion of *bad configuration* used in our original work [5]. In Section 7 we discuss the possible relations between these two notions of badness.

1.4 Future research program

The results in (A)–(E) above are the preliminary steps towards a comprehensive theory of dynamical Gibbs-non-Gibbs transitions based on the principles outlined above. Let us conclude this introduction with a list of further issues which must be addressed to develop such a theory:

- *Definition of "nature-trajectories" and "nurture-trajectories".* This is a delicate issue that requires full exploitation of the properties of the rate function for the trajectory. It must involve a suitable notion of distance between conditioned and unconditioned trajectories.

- *Relation between nature-trajectories and Gibbsianness.* It is intuitively clear that Gibbsianness is conserved for times so short that only nature-trajectories are possible. A rigorous proof of this fact would confirm our intuition and would lead to alternative and less technical proofs of short-time Gibbsianness preservation.

- *Study of nurture-trajectories.* We expect that nurture-trajectories start very close to unconstrained trajectories, and move away only shortly before the end in order to satisfy the conditioning. For the case of time-reversible evolutions, the time it takes to get to the nurture-regime should be the same as the initial relaxation time to (almost) equilibrium.

- *Study of nature-nurture transitions.* Transitions from nature to nurture should happen only once for every conditioning measure (i.e., there should be no nature-restoration). Natural questions are: Does the time at which these transitions take place depend on the conditioning measure? Is there a common time after which every trajectory becomes nurture?

- *Case studies of trajectories leading to non-Gibbsianness.* These should determine "forbidden regions" in trajectory space. Natural questions are: How do these regions evolve? Are they monotone in time?

- *Relation between nurture-trajectories and non-Gibbsianness.* While we expect that "all trajectories are nature" implies Gibbsianness of the evolved measure, we do not expect that"some trajectories are nurture" leads to non-Gibbsianness. Examples are needed to clarify this asymmetry. The case of the Ising model in non-zero field – in which Gibbsianness is eventually restored – should be particularly enlightening.

1.5 Outline

Our paper is organized as follows. In Section 2, we consider the case of independent spin-flips, as a warm-up for the rest of the paper. In Section 3, we compute the *non-linear generator* for dependent spin-flips, which plays a key rol in the large deviation principle we are after. In Sections 4 and 5, we compute the *Legendre transform* of this non-linear generator, which is the object that enters into the associated rate function, as an *action integral.* In Section 4 we do the computation for independent spin-flips, in Section 5 we extend the computation to dependent spin-flips. In Section 6, we look at *bad measures*, i.e., measures at time $t > 0$ for which the optimal trajectory leading to this measure and minimizing the rate function is non-unique. In Section 7, we use these results to develop our large-devation view on Gibbs-non-Gibbs transtions. In Appendix A we illustrate the large deviation formalism in Feng and Kurtz [11], which lies at the basis of Sections 2–5, by considering a simple example, namely, a Poisson random walk with small increments. This will help the reader not familiar with this formalism to grasp the main ideas.

2 Independent spin-flips: trajectory of the magnetization

2.1 Large deviation principle

As a warm-up, we consider the example of Ising spins on the one-dimensional torus $T_N = \{1, \dots, N\}$ subject to a rate-1 independent spin-flip dynamics. Write $\mathbb{P}_N$ to denote the law of this process. We look at the trajectory of the magnetization, i.e., $t \mapsto m_N(t) = N^{-1} \sum_{i=1}^{N} \sigma_i(t)$, where $\sigma_i(t)$ is the spin at site i at time t. A spin-flip from $+1$ to -1 (or from -1 to $+1$) corresponds to a jump of size $-2N^{-1}$ (or $+2N^{-1}$) in the magnetization, i.e., the generator L_N of the process $(m_N(t))_{t\geq 0}$ is given by

$$(L_N f)(m) = \tfrac{1+m}{2}\, N \left[f\left(m - 2N^{-1}\right) - f(m)\right] + \tfrac{1-m}{2}\, N \left[f\left(m + 2N^{-1}\right) - f(m)\right] \qquad (2.1)$$

for $m \in \{-1, -1 + 2N^{-1}, \dots, 1 - 2N^{-1}, 1\}$. If $\lim_{N\to\infty} m_N = m$ and f is C^1 with bounded derivative, then

$$\lim_{N\to\infty} (L_N f)(m_N) = (Lf)(m) \quad \text{with} \quad (Lf)(m) = -2mf'(m). \qquad (2.2)$$

This is the generator of the deterministic process $m(t) = m(0)e^{-2t}$, solving the equation $\dot{m}(t) = -2m(t)$ (the dot denotes the derivative with respect to time).

The trajectory of the magnetization satisfies a large deviation principle, i.e., for every time horizon $T \in (0,\infty)$ and trajectory $\gamma = (\gamma_t)_{t\in[0,T]}$,

$$\mathbb{P}_N\Big(\big(m_N(t)\big)_{t\in[0,T]} \approx (\gamma_t)_{t\in[0,T]}\Big) \approx \exp\left[-N\int_0^T L(\gamma_t,\dot\gamma_t)\,dt\right], \tag{2.3}$$

where the Lagrangian $t \mapsto L(\gamma_t,\dot\gamma_t)$ can be computed following the scheme of Feng and Kurtz [11], Example 1.5. Indeed, we first compute the so-called *non-linear generator* H given by

$$(Hf)(m) = \lim_{N\to\infty} (\mathcal{H}_N f)(m_N) \quad \text{with} \quad (\mathcal{H}_N f)(m_N) = \frac{1}{N}\, e^{-Nf(m_N)}\, L_N(e^{Nf})(m_N), \tag{2.4}$$

where $\lim_{N\to\infty} m_N = m$. This gives

$$(Hf)(m) = \tfrac{m+1}{2}\,(e^{-2f'(m)}-1) + \tfrac{1-m}{2}\,(e^{2f'(m)}-1), \tag{2.5}$$

which is of the form

$$(Hf)(m) = H\big(m, f'(m)\big) \tag{2.6}$$

with

$$H(m,p) = \tfrac{m+1}{2}\,(e^{-2p}-1) + \tfrac{1-m}{2}\,(e^{2p}-1). \tag{2.7}$$

Because $p \mapsto H(m,p)$ is convex, we have

$$H(m,p) = \sup_{q\in\mathbb{R}} \,[pq - L(m,q)] \tag{2.8}$$

with

$$\begin{aligned} L(m,q) &= \sup_{p\in\mathbb{R}} \,[pq - H(m,p)] \\ &= \frac{q}{2}\log\left(\frac{q+\sqrt{q^2+4(1-m^2)}}{2(1+m)}\right) - \frac{1}{2}\sqrt{q^2+4(1-m^2)} + 1. \end{aligned} \tag{2.9}$$

Hence, using the theory developed in Feng and Kurtz [11], Chapter 1, Example 1.5, we indeed have the large deviation principle in (2.3) with $L(\gamma_t,\dot\gamma_t)$ given by (2.9) with $m=\gamma_t$ and $q=\dot\gamma_t$.

2.2 Optimal trajectories

We may think of the typical trajectories $(m_N(t))_{t\in[0,T]}$ as being exponentially close to *optimal trajectories* minimizing the *action functional* $\gamma = (\gamma_t)_{t\in[0,T]} \mapsto \int_0^T L(\gamma_t,\dot\gamma_t)\,dt$. The optimal trajectories satisfy the Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot\gamma_t} = \frac{\partial L}{\partial \gamma_t} \tag{2.10}$$

or, equivalently, the Hamilton-Jacobi equations corresponding to the Hamiltonian in (2.7),

$$\dot m = \frac{\partial H}{\partial p}, \qquad \dot p = -\frac{\partial H}{\partial m}, \tag{2.11}$$

which gives

$$\dot m = -m(e^{2p}+e^{-2p}) + (e^{2p}-e^{-2p}), \qquad \dot p = \tfrac{1}{2}\,(e^{2p}-e^{-2p}). \tag{2.12}$$

Putting $h = \tanh(p)$ and integrating the second equation in (2.12), we obtain

$$h(t) = C\, e^{2t}. \tag{2.13}$$

Using that $\operatorname{arctanh}(x) = \frac{1}{2}\log(\frac{1+x}{1-x})$, we get

$$\dot{m} = -m\,\frac{2+2h^2}{1-h^2} + \frac{4h}{1-h^2}, \tag{2.14}$$

which can be integrated to yield the solution

$$m(t) = C_1 e^{2t} + C_2 e^{-2t}, \tag{2.15}$$

where the constants C_1, C_2 are determined by the initial magnetization and the corresponding initial momentum. One example of an optimal trajectory corresponds to the dynamics *starting* from an initial magnetization m_0, giving $m(t) = m_0 e^{-2t}$, i.e., $C_1 = 0$ and $C_2 = m_0$. Another example of an optimal trajectory is the reversed dynamics *arriving* at magnetization m_T at time T, giving $m(t) = m_T e^{2(t-T)}$, i.e., $C_2 = 0$ and $C_1 = m_T e^{-2T}$.

Yet another example is the following. Suppose that we start the independent spin-flip dynamics from a measure under which the magnetization satisfies a large deviation principle with rate function, say, I, e.g. a Gibbs measure. If we want to arrive at a given magnetization m_T at time T, then the optimal trajectory is given by (2.15) with end condition $m(T) = m_T$ and satisfying the *open-end condition* relating the Lagrangian L at time $t = 0$ to the rate function I at magnetization $m = \gamma_0$ as follows:

$$\left[\frac{\partial L(\gamma_t, \dot{\gamma}_t)}{\partial \dot{\gamma}_t}\right]_{t=0} = -\left[\frac{\partial I(m)}{\partial m}\right]_{m=\gamma_0}. \tag{2.16}$$

This condition is obtained by minimizing $\gamma \mapsto I(\gamma_0) + \int_0^T L(\gamma_t, \dot{\gamma}_t)\, dt$ (see Ermolaev and Külske [10]).

3 Trajectory of the empirical measure for dependent spin-flips

We will generalize the computation in Section 2 in two directions. First, for independent spin-flips we are confronted with the problem that the rate at which the average of a local observable changes in general depends on the average of other observables. Second, for dependent spin-flips even the trajectory of the magnetization is not Markovian. Therefore, we are obliged to consider the time evolution of all spatial averages jointly, i.e., the *empirical measure*.

3.1 Setting and notation

For $N \in \mathbb{N}$, let $\mathbb{T}_N^d$ be the d-dimensional N-torus $(\mathbb{Z}/(2N+1)\mathbb{Z})^d$. For $i, j \in \mathbb{T}_N^d$, let $i + j$ denote coordinate-wise addition modulo $2N + 1$. We consider Glauber dynamics of Ising spins located at the sites of $\mathbb{T}_N^d$, i.e., on the configuration space $\Omega_N = \{-1, 1\}^{\mathbb{T}_N^d}$. We write $\Omega = \{-1, 1\}^{\mathbb{Z}^d}$ to denote the infinite-volume configuration space. Configurations are denoted by symbols like σ and η. For $\sigma \in \Omega_N$, σ_i denotes the value of the spin at site i. We write $\mathcal{M}_1(\Omega)$ to denote the set of probability measures on Ω, and similarly for $\mathcal{M}_1(\Omega_N)$.

The dynamics is defined via the generator L_N acting on functions $f\colon \Omega_N \to \mathbb{R}$ as

$$(L_N f)(\sigma) = \sum_{i\in\mathbb{T}_N^d} c_i(\sigma)\,[f(\sigma^i) - f(\sigma)], \tag{3.1}$$

where σ^i denotes the configuration obtained from σ by flipping the spin at site i. The rates $c_i(\sigma)$ are assumed to be strictly positive and translation invariant, i.e.,

$$c_i(\sigma) = c_0(\tau_i\sigma) = c(\tau_i\sigma) \quad \text{with} \quad (\tau_i\sigma)_j = \sigma_{i+j}. \tag{3.2}$$

We think of the dynamics with generator L_N as a finite-volume version with periodic boundary condition of the infinite-volume generator

$$(Lf)(\sigma) = \sum_{i\in\mathbb{Z}^d} c_i(\sigma)\,[f(\sigma^i) - f(\sigma)], \tag{3.3}$$

where now f is supposed to be a local function, i.e., a function depending on a finite number of σ_j, $j \in \mathbb{Z}^d$. We denote by $(S_t)_{t\geq 0}$ with $S_t = e^{tL}$ the semigroup acting on $C(\Omega)$ (the space of continuous functions on Ω)) associated with the generator in (3.3), and similarly $(S_t^N)_{t\geq 0}$ with $S_t^N = e^{tL_N}$. For $\mu \in \mathcal{M}_1(\Omega)$, we denote by $\mu S_t \in \mathcal{M}_1(\Omega)$ the distribution μ evolved over time t, and similarly for $\mu_N S_t^N$ and $\mu_N \in \mathcal{M}_1(\Omega_N)$.

We embed $\mathbb{T}_N^d$ in $\mathbb{Z}^d$ by identifying it with $\Lambda_N^d = ([-N,N]\cap\mathbb{Z})^d$. Through this identification, we give meaning to expressions like $\sum_{i\in\mathbb{T}_N^d} f(\tau_i\sigma)$ for $\sigma \in \Omega$ and $f\colon \Omega \to \mathbb{R}$. In this way we may also view local functions $f\colon \Omega \to \mathbb{R}$ as functions on Ω_N as soon as N is large enough for Λ_N^d to contain the dependence set of f. For a translation-invariant $\mu \in \mathcal{M}_1(\Omega)$, we denote by μ_N its natural restriction to Ω_N.

By the locality of the spin-flip rates, the infinite-volume dynamics is well-defined and is the uniform limit of the finite-volume dynamics, i.e., for every local function $f\colon \Omega \to \mathbb{R}$ and $t > 0$,

$$\lim_{N\to\infty} \|S_t^N f - S_t f\|_\infty = 0. \tag{3.4}$$

See Liggett [20], Chapters 1 and 3, for details on existence of the infinite-volume dynamics.

3.2 Empirical measure

For $N \in \mathbb{N}$ and $\sigma \in \Omega_N$, the empirical measure associated with σ is defined as

$$\mathcal{L}_N(\sigma) = \frac{1}{|\mathbb{T}_N^d|} \sum_{i\in\mathbb{T}_N^d} \delta_{\tau_i\sigma}. \tag{3.5}$$

This is an element of $\mathcal{M}_1(\Omega_N)$ which acts on functions $f\colon \Omega_n \to \mathbb{R}$ as

$$\langle f, \mathcal{L}_N\rangle = \int_{\Omega_N} f\, d\mathcal{L}_N = \frac{1}{|\mathbb{T}_N^d|} \sum_{i\in\mathbb{T}_N^d} f(\tau_i\sigma). \tag{3.6}$$

As already mentioned above, a local $f\colon \Omega \to \mathbb{R}$ may be considered as a function on Ω_N for N large enough. A sequence $(\mu_N)_{N\in\mathbb{N}}$ with $\mu_N \in \mathcal{M}_1(\Omega_N)$ converges weakly to some $\mu \in \mathcal{M}_1(\Omega)$ if

$$\lim_{N\to\infty} \int_{\Omega_N} f d\mu_N = \int_\Omega f d\mu \qquad \forall\, f \text{ local}. \tag{3.7}$$

For $\sigma \in \Omega$, we define its periodized version σ^N as $\sigma_i^N = \sigma_i$ for $i = (i_1, \dots, i_d)$ with $-N \le i_k < N$ for $k \in \{1, \dots, d\}$, and $\sigma_i^N = \sigma_{i \bmod (2N+1)}$ otherwise.

If μ is ergodic under translations, then by the locality and the translation invariance of the spin-flip rates also μS_t is ergodic under translations. Let μ^N be the distribution of σ^N when σ is drawn from μ. Since the semigroup $(S_t^N)_{t\ge0}$ uniformly approaches the semigroup $(S_t)_{t\ge0}$ as $N \to \infty$, the ergodic theorem implies that

$$\mathcal{L}_N(\sigma^N(t)) \to \mu S_t \quad \text{weakly as} \quad N \to \infty, \tag{3.8}$$

where $\sigma^N(t)$ denotes the random configuration that is obtained by evolving σ^N over time t in the process with generator L_N.

The deterministic trajectory $t \mapsto \mu S_t$ is the solution of the equation

$$\frac{d\mu_t}{dt} = L^* \mu_t, \tag{3.9}$$

where L^* denotes the adjoint of the generator acting on the space of finite signed measures $\mathcal{M}(\Omega)$. Thus, we can view the convergence in (3.8) as an infinite-dimensional law of large numbers, where the random measure-valued trajectory $(\mathcal{L}_N((\sigma^N(t)))_{t\in[0,T]}$ converges to the deterministic measure-valued trajectory $(\mu S_t)_{t\in[0,T]}$. It is therefore natural to ask for an associated large deviation principle, i.e., does there exist a rate function $\gamma \mapsto I(\gamma)$ such that

$$\mathbb{P}_N\Big(\big(\mathcal{L}_N((\sigma^N(t))\big)_{t\in[0,T]} \approx \gamma\Big) \approx \exp[-|\mathbb{T}_N^d| I(\gamma)]? \tag{3.10}$$

Inspired by the example of the magnetization described in Section 2, we expect the answer to be yes and the rate function to be of the form

$$I(\gamma) = \int_0^T L(\gamma_t, \dot\gamma_t)\, dt \tag{3.11}$$

for some appropriate Lagrangian L. In order to compute L, we must first find the generator of the non-linear semigroup.

3.3 The generator of the non-linear semigroup

In our setting the non-linear generator is defined as follows:

$$(H_N F)(\mathcal{L}_N(\sigma)) = \frac{1}{|\mathbb{T}_N^d|}\, e^{-|\mathbb{T}_N^d| F(\mathcal{L}_N(\sigma))} L_N \left(e^{|\mathbb{T}_N^d|(F\circ\mathcal{L}_N)}\right)(\sigma). \tag{3.12}$$

If the expression in (3.12) has a limit $(HF)(\mu)$ as $N \to \infty$ when $\mathcal{L}_N(\sigma) \to \mu$ weakly, then a candidate rate function can be constructed via Legendre transformation (see Section 5).

To compute the limit operator, we start with a simple function of the form

$$F(\mathcal{L}_N(\sigma)) = \langle f, \mathcal{L}_N(\sigma)\rangle, \tag{3.13}$$

where $f\colon\ \Omega \to \mathbb{R}$ is a local function. Such f's are linear combinations of the functions

$$H_A(\sigma) = \prod_{i\in A} \sigma_i, \qquad A \subseteq \mathbb{Z}^d \text{ finite}, \tag{3.14}$$

which live on Ω_N for N large enough.

Theorem 3.1. *For all local $f \in \Omega$ and N large enough,*

$$\frac{1}{|\mathbb{T}_N^d|} e^{-|\mathbb{T}_N^d|\langle f,\mathcal{L}_N(\sigma)\rangle} L_N \left(e^{|\mathbb{T}_N^d|\langle f,\mathcal{L}_N\rangle}\right)(\sigma) = \left\langle c\left(e^{\mathcal{D}_N f} - 1\right), \mathcal{L}_N(\sigma)\right\rangle, \tag{3.15}$$

where $\mathcal{D}_N$ is the linear operator, acting on functions on Ω_N, defined via

$$\mathcal{D}_N 1 = 0, \qquad \mathcal{D}_N H_A = \sum_{r\in -A} (-2) H_{A+r} \text{ for } A \subseteq \mathbb{T}_N^d, \tag{3.16}$$

where the N-dependence refers only to the fact that the addition $A+r$ is modulo $2N+1$.

Proof. Using the definition of the generator L_N in (3.1), we write (recall (3.2))

$$\begin{aligned} &e^{-|\mathbb{T}_N^d|\langle f,\mathcal{L}_N(\sigma)\rangle} L_N \left(e^{|\mathbb{T}_N^d|\langle f,\mathcal{L}_N\rangle}\right)(\sigma) \\ &\quad= \sum_{k\in\mathbb{T}_N^d} c(\tau_k\sigma) \left\{\exp\left[\sum_{j\in\mathbb{T}_N^d}\left[f\left(\tau_j(\sigma^k)\right) - f\left(\tau_j(\sigma)\right)\right]\right] - 1\right\}. \end{aligned} \tag{3.17}$$

Since

$$(\mathcal{D}_N^k f)(\sigma) = \sum_{j\in\mathbb{T}_N^d}\left[f\left(\tau_j(\sigma^k)\right) - f\left(\tau_j(\sigma)\right)\right] \tag{3.18}$$

is a linear operator, it suffices to prove that

$$(\mathcal{D}_N^k f)(\sigma) = (\mathcal{D}_N f)(\tau_k\sigma) \quad \text{for} \quad f = H_A, \tag{3.19}$$

where $\mathcal{D}_N f$ is given by (3.16) for $f = H_A$ (note that if $f = H_A$, then $f(\sigma^k) = -H_A(\sigma)$ for $k \in A$ and $f(\sigma^k) = f(\sigma)$ otherwise). Hence

$$\begin{aligned} (\mathcal{D}_N^k H_A)(\sigma) &= \sum_{j\in\mathbb{T}_N^d}\left(\prod_{i\in A}(\sigma^k)_{i+j} - \prod_{i\in A}\sigma_{i+j}\right) = \sum_{j\in\mathbb{T}_N^d} 1_{\{k-j\in A\}}\,(-2)\prod_{i\in A}\sigma_{i+j} \\ &= \sum_{j\in\mathbb{T}_N^d} 1_{\{j\in -A+k\}}\,(-2)\prod_{i\in A}\sigma_{i+j} = (-2)\sum_{r\in -A}\prod_{i\in A}\sigma_{i+r+k} \\ &= \left((-2)\sum_{r\in -A} H_{A+r}\right)(\tau_k\sigma). \end{aligned} \tag{3.20}$$

□

Remark: Note that, in the limit as $N \to \infty$, $\mathcal{D}_N$ becomes an unbounded operator, defined on local functions $f\colon \Omega \to \mathbb{R}$ via

$$\mathcal{D}1 = 0, \qquad \mathcal{D}\left(\sum_A \alpha_A H_A\right) = \sum_A H_A\left(\sum_{r\in -A}(-2)\alpha_{A-r}\right). \tag{3.21}$$

The domain of $\mathcal{D}$ can be extended to functions $f = \sum_A \alpha_A H_A$ for which

$$\sum_A \sum_{r\in -A} |\alpha_{A-r}| < \infty. \tag{3.22}$$

The dual operator $\mathcal{D}^*$ acts on $\mathcal{M}(\Omega)$, the space of finite signed measures on Ω, and since $\mathcal{D}1 = 0$, $\mathcal{D}^*$ has the measures of total mass zero as image set. The intuitive idea is that when the dynamics starts from the empirical measure μ, after an infinitesimal time t the empirical measure is $\mu + t\mathcal{D}^*\mu + o(t)$.

Remark: From Theorem 3.1 it follows that, for $f = \sum_{i=1}^N \lambda_i f_i$,

$$L_N \left(e^{|\mathbb{T}_N^d| \sum_{i=1}^N \lambda_i \langle f_i, \mathcal{L}_N \rangle} \right)(\sigma) = |\mathbb{T}_N^d| \, e^{|\mathbb{T}_N^d| \sum_{i=1}^n \lambda_i \langle f_i, \mathcal{L}_N(\sigma) \rangle} \left\langle c \left(e^{\sum_{i=1}^n \lambda_i \mathcal{D} f_i} - 1 \right), \mathcal{L}_N(\sigma) \right\rangle. \tag{3.23}$$

The right-hand side is a function of $\mathcal{L}_N$. By taking derivatives with respect to the variables λ_i, we see that the generator maps any function of $\mathcal{L}_N$ into a function of $\mathcal{L}_N$. This shows that $(\mathcal{L}_N(\sigma^N(t)))_{t \geq 0}$ is a *Markov process*. Roughly speaking, this Markov process can be viewed as a random walk that makes jumps of size $1/|\mathbb{T}_N^d|$ at rate $|\mathbb{T}_N^d|$. Of course, the problem is that this random walk is infinite-dimensional, and therefore we cannot directly apply standard random-walk theory.

Theorem 3.1 shows that the operator H defined by

$$(HF)(\mu) = \lim_{N \to \infty} (H_N F)(\mathcal{L}_N(\sigma)) \quad \text{when} \quad \lim_{N \to \infty} \mathcal{L}_N = \mu \text{ weakly} \tag{3.24}$$

is well-defined for $F(\mu) = \langle f, \mu \rangle$. We next extend Theorem 3.1 to F of the form

$$F(\mu) = \Psi\big(\langle f_1, \mu \rangle, \ldots, \langle f_n, \mu \rangle\big), \tag{3.25}$$

where $\Psi \colon \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{C}^\infty$ with uniformly bounded derivatives of all orders.

Theorem 3.2. *If* $\lim_{N \to \infty} \mathcal{L}_N = \mu$ *and* F *is of the form* (3.25)*, then (with the same notation as in* (3.12)*)*

$$(HF)(\mu) = \lim_{N \to \infty} (H_N F)(\mu) = \left\langle c \left(\exp\left[\sum_{i=1}^n \frac{\partial \Psi}{\partial x_i} (\langle f_1, \mu \rangle, \ldots, \langle f_n, \mu \rangle) \mathcal{D} f_i \right] - 1 \right), \mu \right\rangle. \tag{3.26}$$

Proof. Compute

$$\begin{aligned} &\frac{1}{|\mathbb{T}_N^d|} \, e^{-|\mathbb{T}_N^d| F(\mathcal{L}_N(\sigma))} \, L_N \left(e^{|\mathbb{T}_N^d| F(\mathcal{L}_N)} \right)(\sigma) \\ &\qquad = \sum_{k \in \mathbb{T}_N^d} c(\tau_k \sigma) \left(\exp\left[|\mathbb{T}_N^d| \left(F\big(\mathcal{L}_N(\sigma^k)\big) - F\big(\mathcal{L}_N(\sigma)\big) \right) \right] - 1 \right). \end{aligned} \tag{3.27}$$

Next, use the fact that

$$\langle f, \mathcal{L}_N(\sigma^k) \rangle - \langle f, \mathcal{L}_N(\sigma) \rangle = \frac{1}{|\mathbb{T}_N^d|} \, (\mathcal{D}_N f)(\tau_k(\sigma)) \tag{3.28}$$

to see that

$$\begin{aligned} &\Psi\big(\langle f_1, \mathcal{L}_N(\sigma^k) \rangle, \ldots, \langle f_n, \mathcal{L}_N(\sigma^k) \rangle\big) - \Psi\big(\langle f_1, \mathcal{L}_N(\sigma) \rangle, \ldots, \langle f_n, \mathcal{L}_N(\sigma) \rangle\big) \\ &\qquad = \sum_{i=1}^n \frac{\partial \Psi}{\partial x_i} \big(\langle f_1, \mathcal{L}_N(\sigma) \rangle, \ldots, \langle f_n, \mathcal{L}_N(\sigma) \rangle\big) \, (\mathcal{D} f_i)(\tau_k \sigma) + o\left(\frac{1}{|\mathbb{T}_N^d|} \right). \end{aligned} \tag{3.29}$$

Combine (3.27) and (3.29) and take the limit $N \to \infty$, to obtain (3.26). $\square$

Remark: For

$$F(\mu) = \Psi\big(\langle f_1, \mu\rangle, \ldots, \langle f_n, \mu\rangle\big), \tag{3.30}$$

the functional derivative of F with respect to μ is defined as

$$\frac{\delta F}{\delta \mu} = \sum_{i=1}^{n} \frac{\partial \Psi}{\partial x_i}\big(\langle f_1, \mu\rangle, \ldots, \langle f_n, \mu\rangle\big)\, f_i. \tag{3.31}$$

We may therefore rewrite (3.29) as

$$H(F)(\mu) = \left\langle c\, \left(\exp\left[\mathcal{D}\left(\frac{\delta F}{\delta \mu}\right)\right] - 1\right), \mu\right\rangle. \tag{3.32}$$

4 The rate function for independent spin-flips

4.1 Legendre transform

Having completed the computation of the non-linear generator in Section 3, we are ready to compute its Legendre transform. As a warm-up, we will first do this for independent spin-flips, i.e., when $c \equiv 1$ in (3.2). In Section 5 we will extend the calculation to general c, which will not represent a serious obstacle.

The non-linear generator in (3.26) is of the form

$$(HF)(\mu) = \mathcal{H}\left(\mu, \frac{\delta F}{\delta \mu}\right), \tag{4.1}$$

where, for $\mu \in \mathcal{M}_1(\Omega)$ and $f\colon\ \Omega \to \mathbb{R}$ continuous,

$$\mathcal{H}(\mu, f) = \left\langle c\, \left(e^{\mathcal{D}f} - 1\right), \mu\right\rangle. \tag{4.2}$$

By the convexity of $f \mapsto \mathcal{H}(\mu, f)$, we have

$$\mathcal{H}(\mu, f) = \sup_{\alpha \in \mathcal{M}(\Omega)} \left[\int_\Omega f\, d\alpha - L(\mu, \alpha)\right] \tag{4.3}$$

with

$$L(\mu, \alpha) = \sup_{f \in \mathcal{C}(\Omega)} \left[\int_\Omega f\, d\alpha - \mathcal{H}(\mu, f)\right] \tag{4.4}$$

the Lagrangian appearing in the large deviation rate function in (3.11). As explained in Feng and Kurtz [11], Section 8.6.1.2, the representation of the generator in (4.1), where $\mathcal{H}(\mu, f)$ is a Legendre transform as in (4.3), implies that the generator in (4.1) generates a non-linear semigroup, called the *Nisio control semigroup*, associated with the function L (see [11], Section 8.1).

Remark: The operator $\mathcal{D}$ has the property

$$\mathcal{D} f_0 = -2 f_0, \qquad f_0(\sigma) = \sigma_0, \tag{4.5}$$

i.e., f_0 is an eigenfunction of $\mathcal{D}$. We recover the Hamiltonian in (2.7) (associated with the large deviation principle of the magnetization) from the infinite-dimensional Hamiltonian in (4.2) via the relation

$$\mathcal{H}(\mu, p f_0) = H\big(\langle f_0, \mu\rangle, p\big). \tag{4.6}$$

Remark: The infinite-dimensional Hamiltonian in (4.2) can be thought of as a function of the "position" variable μ and the "momentum" variable f. The corresponding Hamilton-Jacobi equations read

$$\dot{\mu} = \frac{\delta H}{\delta f}, \qquad \dot{f} = -\frac{\delta H}{\delta \mu}. \tag{4.7}$$

These give a closed equation for f, because the Hamiltonian in (4.2) is linear in μ. If we can solve the latter equation to find f, then we can integrate the equation for μ and find the solution for μ. This is precisely the same situation – but now infinite-dimensional – as we encountered in (2.12), where the equation for p was closed and could be integrated to give the solution for m.

4.2 Computation of the Lagrangian

To find L, the function appearing in the rate function in (3.11), we have to compute the Legendre transform in (4.4). To do so, we first consider the finite-dimensional analogue. We start with rates $c \equiv 1$, for which (4.4) becomes

$$\begin{aligned} L(\mu,\alpha) &= \sup_{f\in\mathbb{R}^n} \left[\sum_{i=1}^n f_i\alpha_i - \sum_{i=1}^n \left(e^{\sum_{j=1}^n D_{ij}f_j} - 1 \right)\mu_i \right], \\ \mu &= (\mu_1,\ldots,\mu_n),\ \alpha = (\alpha_1,\ldots,\alpha_n),\ f = (f_1,\ldots,f_n), \end{aligned} \tag{4.8}$$

where $\mu_i \in (0,\infty)$, $\sum_{i=1}^n \mu_i = 1$, $\alpha_i \in \mathbb{R}$, $f_i \in \mathbb{R}$, and $D_{ij} \in \mathbb{R}$. The matrix D has the additional property that $D(\overline{1}) = \overline{0}$, where $\overline{1}$ is the vector with all components equal to 1. Hence $\sum_{i=1}^n (D^T\mu)_i = 0$, i.e., the transposed matrix D^T maps any vector to a vector with zero sum. For a vector α, we say that $(D^T)^{-1}\alpha$ is well-defined if there exists a unique vector $\nu = \nu(\alpha)$ with sum equal to 1 such that $D^T\nu = \alpha$. For two column vectors $\alpha,\beta \in \mathbb{R}^n$, let $\alpha\beta$ be the vector with components $\alpha_i\beta_i$, α/β the vector with components α_i/β_i. For $f\colon \mathbb{R}\to\mathbb{R}$, write $f(\alpha)$ to denote the vector with components $f(\alpha_i)$. Then the equation for the maximizer $f = f^*$ of (4.8) becomes

$$\alpha_k = \sum_{i=1}^n \mu_i\, e^{\sum_{j=1}^n D_{ij}f_j^*}\, D_{ik}, \quad k = 1,\ldots,n, \tag{4.9}$$

which in vector notation reads

$$\alpha = D^T(\mu e^{Df^*}). \tag{4.10}$$

If $(D^T)^{-1}\alpha$ is well-defined, then we can rewrite the latter equation as

$$Df^* = \log\left[\frac{(D^T)^{-1}\alpha}{\mu}\right], \tag{4.11}$$

and for this f^* we have

$$\sum_{i=1}^n f_i^*\alpha_i = \langle f^*,\alpha\rangle = \left\langle \log\left[\frac{(D^T)^{-1}\alpha}{\mu}\right], (D^T)^{-1}\alpha \right\rangle \tag{4.12}$$

and

$$\sum_{i=1}^n \left(e^{\sum_{j=1}^n D_{ij}f_j^*} - 1 \right)\mu_i = 0, \tag{4.13}$$

because the total mass of μ and $(D^T)^{-1}$ are both equal to 1. Hence, inserting (4.12) and (4.13) into (4.8), we obtain the expression

$$L(\mu, \alpha) = \left\langle \log \left[\frac{(D^T)^{-1}\alpha}{\mu} \right], (D^T)^{-1}\alpha \right\rangle, \tag{4.14}$$

which is the *relative entropy* of $(D^T)^{-1}\alpha$ with respect to μ. The intuition behind (4.14) is that $L(\mu, \alpha)$ is the cost under the Markovian evolution for the initial measure to have derivative α at time zero.

Let us next consider the infinite-dimensional version of the above computation. First, for $\alpha \in \mathcal{M}(\Omega)$ with total mass zero, we declare $(\mathcal{D}^*)^{-1}\alpha = \nu$ to be well-defined if there exists a probability measure ν such that, for all f in the domain of $\mathcal{D}$,

$$\langle \nu, \mathcal{D} f \rangle = \langle \alpha, f \rangle. \tag{4.15}$$

If α is translation-invariant, then also $(\mathcal{D}^*)^{-1}\alpha$ is translation-invariant. For translation-invariant $\mu, \nu \in \mathcal{M}(\Omega)$, we denote by $s(\nu|\mu)$ the relative entropy density of ν with respect to μ, i.e.,

$$s(\nu|\mu) = \lim_{N \to \infty} \frac{1}{|\mathbb{T}_N^d|} \sum_{\sigma_{\mathbb{T}_N^d}} \nu(\sigma_{\mathbb{T}_N^d}) \log \left[\frac{\nu(\sigma_{\mathbb{T}_N^d})}{\mu(\sigma_{\mathbb{T}_N^d})} \right]. \tag{4.16}$$

Note that this limit does not necessarily exist. But if μ is a Gibbs measure, then for all translation-invariant ν both $s(\nu|\mu)$ and $s(\nu|\mu_t)$ exist, where μ_t is μ evolved over time t (see van Enter, Fernández and Sokal [4], Le Ny and Redig [19]). The rate function which is the analogue of (4.14) is now given by

$$L(\mu, \alpha) = s\left((\mathcal{D}^*)^{-1}\alpha | \mu \right) \tag{4.17}$$

with the same interpretation as for (4.14): $(\mathcal{D}^*)^{-1}\alpha$ produces derivative α at time zero for the trajectory of the empirical measure, and its cost is the relative entropy density of this measure with respect to the initial measure μ.

4.3 Optimal trajectories

In order to gain some intuition about the rate function corresponding to the Lagrangian in (4.17), we identify two easy optimal trajectories.

First, we consider a trajectory that starts from a product measure ν_{x_0} and ends at a product measure ν_{x_t} with $x_t = x_0 e^{-2t}$. The typical trajectory is then simply the product-measure-valued trajectory $\gamma_t = \nu_{x_t}$ with $x_t = x_0 e^{-2t}$. We can easily verify that this trajectory has zero cost. Indeed, $\langle \gamma_t, H_A \rangle = x_t^{|A|}$, and hence $\langle \dot{\gamma}_t, H_A \rangle = |A| x_t^{|A|-1} \dot{x}_t$. On the other hand, $\langle \mathcal{D}^*(\gamma_t), H_A \rangle = -2|A| x_t^{|A|}$ and, since $\dot{x}_s = -2x_t$, we thus see that $\langle \dot{\gamma}_t, H_A \rangle = \langle \mathcal{D}^*(\gamma_t), H_A \rangle$. Therefore $(\mathcal{D}^*)^{-1}(\dot{\gamma}_t) = \gamma_t$, and (4.17) gives

$$L(\gamma_t, \dot{\gamma}_t) = s\left((\mathcal{D}^*)^{-1}(\dot{\gamma}_t) | \gamma_t \right) = s(\gamma_t|\gamma_t) = 0. \tag{4.18}$$

Note that this is the only product-measure-valued trajectory that has zero cost. Indeed, if $\gamma_t = \nu_{x_t}$ has zero cost, then from the requirement that $\langle \dot{\gamma}_t, H_A \rangle = \langle \mathcal{D}^*(\gamma_t), H_A \rangle = -2|A| x_t^{|A|}$ we find that $\dot{x}_t = -2x_t$. For a general starting measure μ, the trajectory that has zero cost satisfies $\langle \dot{\gamma}_t, H_A \rangle = -2|A| \langle \gamma_t, H_A \rangle$, which has as solution $\langle \gamma_t, H_A \rangle = \langle \mu, H_A \rangle e^{-2|A|t}$, corresponding to the Markovian independent spin-flip evolution started from μ. Note that, for a general trajectory γ, $\langle (\mathcal{D}^*)^{-1}(\dot{\gamma}_t), H_A \rangle = -2|A| \langle \dot{\gamma}_t, H_A \rangle$.

Second, we consider the case where $\mu = \mu_y$ is a product measure with

$$\langle \mu_y, H_A \rangle = y^{|A|}, \qquad -1 < y < 1, \tag{4.19}$$

and $\alpha = \alpha_x$ is the derivative at time zero of another product measure, i.e.,

$$\langle \alpha_x, H_A \rangle = -2|A| x^{|A|}, \qquad -1 < x < 1. \tag{4.20}$$

In that case $D^* \alpha = \nu_x$ with ν_x the translation-invariant product measure with $\langle \nu_x, H_A \rangle = x^{|A|}$. The latter follows from the identity

$$\left\langle \left[\sum_{i \in A} H_A(\sigma^i) - H_A(\sigma) \right], \nu_x \right\rangle = -2|A| x^{|A|}, \tag{4.21}$$

and the rate function becomes

$$L(\mu_y, \alpha_x) = \frac{1+x}{2} \log \left(\frac{1+x}{1+y} \right) + \frac{1-x}{2} \log \left(\frac{1-x}{1-y} \right). \tag{4.22}$$

5 The rate function for dependent spin-flips

5.1 Computation of the Lagrangian

For general spin-flip rates c in (3.2), let us return to the matrix calculation in (4.8) and (4.9). Equation (4.8) has to be replaced by

$$L(\mu, \alpha) = \sup_{f \in \mathbb{R}^n} \left[\sum_{i=1}^n f_i \alpha_i - \sum_{i=1}^n c_i \left(e^{\sum_{j=1}^n D_{ij} f_j} - 1 \right) \mu_i \right], \tag{5.1}$$

where $c_i > 0$, $i = 1, \ldots, n$. Put $C_\mu = \sum_{i=1}^n c_i \mu_i$. In the calculation with $c_i = 1$, $i = 1, \ldots, n$, this "total mass" does not depend on μ and is equal to 1. Now, however, it becomes a normalization that depends on μ. We say that $(D^T)^{-1}(\alpha, \mu)$ is well-defined if there exists a non-negative vector $\nu = \nu(\alpha, \mu) = (\nu_1 \ldots, \nu_n)$ with sum C_μ such that $D^T \nu = \alpha$. The analogue of (4.14) reads

$$L(\mu, \alpha) = \left\langle \log \left[\frac{(D^T)^{-1}(\alpha, \mu)}{\mu} \right], (D^T)^{-1}(\alpha, \mu) \right\rangle. \tag{5.2}$$

In order to find the analogue of this expression in the infinite-dimensional setting, we proceed as follows. For two finite positive measures μ, ν of equal total mass M, we define $S(\mu|\nu)$ to be the relative entropy density of the probability measures $\mu/M, \nu/M$, i.e., $S(\mu|\nu) = s(\nu/M|\mu/M)$. For $\mu \in \mathcal{M}_1(\Omega)$, we define the c-modification of μ as the positive measure defined via $\int_\Omega f(\sigma) d\mu_c(\sigma) = \int_\Omega c(\sigma) f(\sigma) d\mu(\sigma)$. For a signed measure of total mass zero and $\mu \in \mathcal{M}_1(\Omega)$, we say that $(\mathcal{D}^*)^{-1}(\alpha, \mu, c) = \nu$ is well-defined if there exists a positive measure ν of total mass equal to that of μ_c such that $\mathcal{D}^*(\nu) = \alpha$. Then the analogue of (5.2) becomes

$$L(\mu, \alpha) = s\left((\mathcal{D}^*)^{-1}(\alpha, \mu, c) | \mu_c \right). \tag{5.3}$$

5.2 The non-linear semigroup and its relation with relative entropy

The non-linear semigroup with generator (3.12) is defined as follows. Let $\mathcal{P}^{\text{inv}}(\Omega)$ be the set of translation-invariant probability measures on Ω. For local functions $f_1, \ldots, f_n \colon \Omega \to \mathbb{R}$ and a $\mathcal{C}^\infty$-function $\Psi \colon \mathbb{R}^n \to \mathbb{R}$, we define an associated function $F_\Psi^{f_1,\ldots,f_N} \colon \mathcal{P}^{\text{inv}}(\Omega) \to \mathbb{R}$ via

$$F_\Psi^{f_1,\ldots,f_n}(\mu) = \Psi\left(\int_\Omega f_1 d\mu, \ldots, \int_\Omega f_n d\mu\right). \tag{5.4}$$

Since $\langle f_i, \mathcal{L}_N\rangle$ is well-defined for N large enough, we can define $F_\Psi^{f_1,\ldots,f_n}(\mathcal{L}_N)$ for N large enough as well. This allows us to define a non-linear semigroup $(V(t))_{t\geq 0}$ via

$$\big(V(t)F_\Psi^{f_1,\ldots,f_n}\big)(\mu) = \lim_{N\to\infty} \frac{1}{|\mathbb{T}_N^d|} \log \mathbb{E}_{\sigma^N}\left(\exp\Big[|\mathbb{T}_N^d| F_\Psi^{f_1,\ldots,f_n}\big(\mathcal{L}_N(\sigma^N(t))\big)\Big]\right), \tag{5.5}$$

where $\mathbb{E}_{\sigma^N}$ denotes expectation with respect to the law of the process starting from σ^N, and the limit is taken along a sequence of configurations $(\sigma^N)_{N\in\mathbb{N}}$ with $\sigma^N \in \Omega_N$ such that the associated empirical measure $\mathcal{L}_N(\sigma^N)$ converges weakly to μ as $N \to \infty$. If $V(t)$ exists, then it defines a non-linear semigroup, and the generator of $V(t)$ is given by (3.32).

Conversely, if H in (3.32) generates a semigroup, then this must be $(V(t))_{t\geq 0}$. The fact that this semigroup is well-defined is sufficient to imply the large deviation principle for the trajectory of the empirical measure (Feng and Kurtz [11], Theorem 5.15). Technically, the difficulty consists in showing that the generator in (3.32) actually generates a semigroup.

We now make the link between the non-linear semigroup, its generator and some familiar objects of statistical mechanics, such as pressure and relative entropy density.

Definition 5.1. *The constrained pressure at time t associated with a function $f \colon \Omega \to \mathbb{R}$ and a Gibbs measure $\mu \in \mathcal{P}^{\text{inv}}(\Omega)$ is defined as*

$$p_t(f|\mu) = \lim_{N\to\infty} \frac{1}{|\mathbb{T}_N^d|} \log \mathbb{E}_{\sigma^N}\left(e^{\sum_{x\in\mathbb{T}_N} \tau_x f(\sigma_t)}\right), \tag{5.6}$$

where the limit is taken along a sequence of configurations $(\sigma^N)_{N\in\mathbb{N}}$ with $\sigma^N \in \Omega_N$ such that the associated empirical measure $\mathcal{L}_N(\sigma^N)$ converges weakly to μ as $N \to \infty$.

In particular, $p_0(f|\mu) = \int_\Omega f d\mu$. The relation between the non-linear semigroup in (5.5) and the constrained pressure at time t reads

$$\big(V(t)\langle f, \cdot\rangle\big)(\mu) = p_t(f|\mu). \tag{5.7}$$

The pressure at time t is defined as

$$p(f|\mu_t) = \lim_{N\to\infty} \frac{1}{|\mathbb{T}_N^d|} \log \mathbb{E}_\mu\left(e^{\sum_{x\in\mathbb{T}_N} \tau_x f(\sigma_t)}\right). \tag{5.8}$$

This is well-defined as soon as the dynamics starts from a Gibbs measure $\mu_0 = \mu$ (see Le Ny and Redig [19]). The relation between the pressure and the constrained pressure reads

$$p(f|\mu_t) = \sup_{\nu\in\mathcal{P}^{\text{inv}}(\Omega)} [p_t(f|\nu) - s(\nu|\mu)]. \tag{5.9}$$

On the other hand, the pressure at time t is the Legendre transform of the relative entropy density with respect to μ_t, i.e.,

$$p(f|\mu_t) = \sup_{\nu \in \mathcal{P}^{\mathrm{inv}}(\Omega)} \left[\int_\Omega f d\nu - s(\nu|\mu_t) \right], \tag{5.10}$$

where the relative entropy density $s(\nu|\mu_t)$ exists because μ_t is *asymptotically decoupled* (see Pfister [24]) as soon as $\mu_0 = \mu$ is a Gibbs measure (see Le Ny and Redig [19]).

The relation between the non-linear generator and the constrained pressure is now as follows. Define the Legendre transform of the constrained pressure as

$$p_t^*(\nu|\mu) = \sup_{f \in \mathcal{C}(\Omega)} \left[\int f d\nu - p_t(f|\mu) \right]. \tag{5.11}$$

Then the relation with the Hamiltonian in (4.2) and the Lagrangian in (5.3) is

$$H(\mu, f) = \left[\frac{d}{dt} p_t(f|\mu) \right]_{t=0} \tag{5.12}$$

and

$$L(\mu, \alpha) = \lim_{t \to 0} \frac{1}{t} p_t^*(\mu + t\alpha|\mu). \tag{5.13}$$

Remark: The operator $\mathcal{D}$, acting on the space $C(\Omega)$ of continuous functions on Ω, has a dual operator $\mathcal{D}^*$, acting on the space $\mathcal{M}(\Omega)$ of finite signed measures on Ω, defined via

$$\langle f, \mathcal{D}^* \mu \rangle = \langle \mathcal{D} f, \mu \rangle. \tag{5.14}$$

In order to gain some understanding for $\mathcal{D}^*$ (which will be useful later on), we first compute $\mathcal{D}^*$ for a Gibbs measure $\mu \in \mathcal{P}^{\mathrm{inv}}(\Omega)$. Without loss of generality we may assume that the interaction potential of μ is a sum of terms of the form $\Phi(A, \sigma) = J_A H(A, \sigma)$, $A \subseteq \mathbb{Z}^d$ finite, where J_A is translation invariant, i.e., $J_{A+k} = J_A$, $k \in \mathbb{Z}^d$. We also assume absolute summability, i.e.,

$$\sum_{A \ni 0} |J_A| < \infty. \tag{5.15}$$

Remember that

$$(\mathcal{D}f)(\sigma) = \sum_{j \in \mathbb{Z}^d} \left[f(\tau_j(\sigma^0)) - f(\tau_j(\sigma)) \right]. \tag{5.16}$$

Therefore, for the Gibbs measure μ under consideration, we have

$$\langle \mu, \mathcal{D} f \rangle = \int_\Omega \left(\sum_{j \in \mathbb{Z}^d} \frac{d\mu^0}{d\mu} \circ \tau_{-j} - 1 \right) f d\mu, \tag{5.17}$$

where μ^0 denotes the distribution of σ^0 when σ is distributed according to μ. Note that the sum in the right-hand side of (5.17) is formal, i.e., the integral is well-defined due to the multiplication with the local function f. In terms of J_A, $A \subseteq \mathbb{Z}^d$ finite, we have

$$\left(\sum_{j \in \mathbb{Z}^d} \frac{d\mu^0}{d\mu} \circ \tau_{-j} - 1 \right)(\sigma) = \sum_{j \in \mathbb{Z}^d} \left(e^{-\sum_{A \ni 0} -2J_A H(A-j, \sigma)} - 1 \right), \tag{5.18}$$

where, once again, this expression is well-defined only after multiplication with a local function and integrated over μ.

6 Bad empirical measures

In Section 7 we will see what consequences the large deviation principle for the trajectory of the empirical measure derived in Sections 3 and 5 has for the question of Gibbs versus non-Gibbs. This needs the notion of bad empirical measure, which we define next.

If we start our spin-flip dynamics from a Gibbs measure $\mu \in \mathcal{P}^{\mathrm{inv}}(\Omega)$, then a probability-measure-valued trajectory $\gamma = (\gamma_t)_{t\in[0,T]}$ has cost

$$\mathcal{I}_\mu(\gamma) = s(\gamma_0|\mu) + \int_0^T L(\gamma_t, \dot\gamma_t)\, dt, \tag{6.1}$$

where the term $s(\gamma_0|\mu)$ is the cost of the initial distribution γ_0. We are interested in the minimizers of $\mathcal{I}_\mu(\gamma)$ over the set of trajectories γ that *end* at a given measure ν. Let

$$\begin{aligned} K_T(\mu',\nu) &= \inf_{\gamma\colon\ \gamma_0=\mu',\gamma_T=\nu} \int_0^T L(\gamma_t,\dot\gamma_t)\, dt \\ &= -\lim_{N\to\infty} \frac{1}{|\mathbb{T}_N^d|} \log \mathbb{P}_\mu \left(\mathcal{L}_N(\sigma_T) = \nu | \mathcal{L}_N(\sigma_0) = \mu'\right). \end{aligned} \tag{6.2}$$

Then $e^{|\mathbb{T}_N^d| - K_T(\mu',\nu)}$ can be thought of as the transition probability for the empirical measure $\mathcal{L}_N$ to go from μ' to ν, up to factors of order $e^{o(|\mathbb{T}_N^d|)}$. Hence

$$\begin{aligned} &-\frac{1}{|\mathbb{T}_N^d|} \log \mathbb{P}_\mu \left(\mathcal{L}_N(\sigma_0) = \mu' | \mathcal{L}_N(\sigma_T) = \nu\right) \\ &\quad = [s(\mu'|\mu) + K_T(\mu',\nu)] - \inf_{\mu'\in\mathcal{P}^{\mathrm{inv}}(\Omega)} [s(\mu'|\mu) + K_T(\mu',\nu)]. \end{aligned} \tag{6.3}$$

Let $M^*(\mu,\nu)$ be the set of probability measure μ' for which the infimum in the right-hand side of (6.3) is attained. We can then think of each element in this set as a typical empirical measure at time $t = 0$ given that the empirical measure at time T is ν. When M^* is a singleton, we denote its unique element by $\mu^*(\mu,\nu)$.

Definition 6.1. *(a) A measure ν is called bad at time t if $M^*(\mu,\nu)$ contains at least two elements μ_1 and μ_2 and there exist two sequences $(\nu_n^1)_{n\in\mathbb{N}}$ and $(\nu_n^2)_{n\in\mathbb{N}}$, both converging to ν as $n\to\infty$, such that $\mu^*(\mu,\nu_n^1)$ converges to μ_1 and $\mu^*(\mu,\nu_n^2)$ converges to μ_2.*
(b) A measure ν that is bad at time t has at least two possible histories, stated as a two-layer property: seeing the measure ν at time t is compatible (in the sense of optimal trajectories) with two different measures at time $t = 0$.

Badness of a measure can be detected as follows.

Proposition 6.2. *A measure ν is bad at time t if there exists a local function $f\colon \Omega\to\mathbb{R}$, two sequences $(\nu_n^1)_{n\in\mathbb{N}}$ and $(\nu_n^2)_{n\in\mathbb{N}}$ both converging to ν, and an $\epsilon > 0$ such that*

$$\Big|\mathbb{E}\Big(f(\sigma(0)) \mid \mathcal{L}_N(\sigma(t)) = (\nu_n^1)_N\Big) - \mathbb{E}\Big(f(\sigma(0)) \mid \mathcal{L}_N(\sigma(t)) = (\nu_n^2)_N\Big)\Big| > \epsilon \quad \forall\, N,n\in\mathbb{N}, \tag{6.4}$$

where $(\nu_n)_N$ denotes the projection of ν_n on $\mathbb{T}_N^d$.

7 A large deviation view on dynamical Gibbs-non-Gibbs transitions

In van Enter, Fernández, den Hollander and Redig [5] we studied the evolution of a Gibbs measure μ under a high-temperature spin-flip dynamics. We showed that the Gibbsianness of the measure μ_t at time $t > 0$ is equivalent to the absence of a phase transition in the double-layer system. More precisely, conditioned on end configuration η at time t, the distribution at time $t = 0$ is a Gibbs measure μ^η with η-dependent formal Hamiltonian

$$H_t^\eta(\sigma, \eta) = H(\sigma) + h_t \sum_{i \in \mathbb{Z}^d} \sigma_i \eta_i, \tag{7.1}$$

where $t \mapsto h_t$ is a monotone function with $\lim_{t \downarrow 0} h_t = \infty$ and $\lim_{t \to \infty} h_t = 0$. If the double-layer system has a phase transition for an end configuration η, then η is called bad. In that case η is an essential point of discontinuity for any version of the conditional probability $\mu_t(\sigma_\Lambda = \cdot \,|\sigma_{\Lambda^c})$, $\Lambda \subseteq \mathbb{Z}^d$ finite.

The relation between the double-layer system and the trajectory of the empirical distribution is as follows. Suppose that the double-layer system has no phase transition for any end configuration η. If we condition the empirical measure at time $t > 0$ to be ν, then (by further conditioning on the configuration η at time $t > 0$) we conclude that at time $t = 0$ we have the measure $\int_\Omega \mu^\eta \nu(d\eta)$. Hence the optimizing trajectory is unique. Conversely, if there exist a bad configuration η, then (because of the translation invariance of the initial measure and of the dynamics) all translates of η are bad also. Hence we expect that a translation-invariant measure ν arising as any weak limit point of $|\mathbb{T}_N^d|^{-1} \sum_{x \in \mathbb{T}_N^d} \delta_{\tau_x \eta}$ is bad also.

As an example, let us consider a situation studied in [5]. The dynamics starts from μ_β^+, the low-temperature plus-phase of the Ising model with zero magnetic field, and evolves according to independent spin-flips. Then, from some time onwards, the alternating configuration $\eta_{\text{alt}}(x) = (-1)^{\sum_{i=1}^d |x_i|}$ becomes bad. The same is true for $-\eta_{\text{alt}}$, and so the translation-invariant measure

$$\nu = \tfrac{1}{2} \left(\delta_{\eta_{\text{alt}}} + \delta_{-\eta_{\text{alt}}} \right) \tag{7.2}$$

has the property that, for ν-a.e. configuration η, the double-layer system has a phase transition when the end configuration is η. Moreover, the Hamiltonian H_t^η has a plus-phase μ_η^+ and a minus-phase μ_η^-. Therefore, when we condition on the empirical measure in (7.2) we get two possible optimal trajectories, one starting at $\frac{1}{2}(\mu_\eta^+ + \mu_{-\eta}^+)$ and one starting at $\frac{1}{2}(\mu_\eta^- + \mu_{-\eta}^-)$. To realize the approximating measures of Proposition 6.2, we choose ν_n^1, ν_n^2 to be the randomized versions of ν where we first choose a configuration according to ν and then independently flip spins with probability $1/n$, to change either from minus to plus or stay plus if it was plus to begin with, respectively to change to minus or stay minus. Clearly, by the FKG-inequality, when conditioning on ν_n^1, respectively, ν_n^2 as empirical distribution, we get a measure at time $t = 0$ that is above $\mu_\eta^+ + \mu_{-\eta}^+$, respectively, below $\mu_\eta^- + \mu_{-\eta}^-$. Hence (6.4) holds with $f(\sigma) = \sigma_0$, and ν is bad.

A A simple example of the Feng-Kurtz formalism

A.1 Poisson walk with small increments

In order to introduce the general formalism developed in Feng and Kurtz [11], let us consider a simple example where computations are simple yet the fundamental objects of the general theory already appear naturally.

Let $X_N = (X_N(t))_{t\geq 0}$ be the continuous-time random walk on $\mathbb{R}$ that jumps N^{-1} forward at rate bN and $-N^{-1}$ backward at rate dN, with $b, d \in (0, \infty)$. This is the Markov process with generator

$$(L_N f)(x) = bN\left[f\left(x + N^{-1}\right) - f(x)\right] + dN\left[f\left(x - N^{-1}\right) - f(x)\right]. \tag{A.1}$$

Clearly, if $\lim_{N\to\infty} X_N(0) = x \in \mathbb{R}$, then

$$\lim_{N\to\infty} X_N(t) = x + (b-d)t, \qquad t > 0, \tag{A.2}$$

i.e., in the limit as $N \to \infty$ the random process X_N becomes a deterministic process $(x(t))_{t\geq 0}$ solving the limiting equation

$$\dot{x} = (b-d), \qquad x(0) = x. \tag{A.3}$$

For all $N \in \mathbb{N}$, we have

$$X_N(t) = N^{-1}\left[\mathcal{N}^+(Nbt) - \mathcal{N}^-(Ndt)\right] = \sum_{i=1}^N (X_i^{bt} - Y_i^{dt}) \tag{A.4}$$

with $\mathcal{N}^+ = (\mathcal{N}^+(t))_{t\geq 0}$ and $\mathcal{N}^- = (\mathcal{N}^-(t))_{t\geq 0}$ independent rate-1 Poisson processes, and X_i^t, Y_i^t, $i = 1, \dots, N$, independent Poisson random variables with mean bt, respectively, dt. Consequently, we can use Cramér's theorem for sums of i.i.d. random variables to compute

$$I(at) = \lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\big(X_N(t) = at \mid X_N(0) = 0\big) = \sup_{\lambda\in\mathbb{R}} \big[at\lambda - F(\lambda)\big], \tag{A.5}$$

where

$$F(\lambda) = \lim_{N\to\infty} \frac{1}{N} \log \mathbb{E}_N\left(e^{\lambda N X_N(t)}\right) = b\big(e^{\lambda} - 1\big) + d\big(e^{-\lambda} - 1\big). \tag{A.6}$$

Thus, we see that

$$I(at) = tL(a) \tag{A.7}$$

with

$$L(a) = \sup_{\lambda\in\mathbb{R}} \Big[a\lambda - b\big(e^{\lambda} - 1\big) - d\big(e^{-\lambda} - 1\big)\Big]. \tag{A.8}$$

Using the property that the increments of the Poisson process are independent over disjoint time intervals, we can now compute

$$\begin{aligned} &\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\Big((X_N(t))_{t\in[0,T]} \approx (\gamma_t)_{t\in[0,T]}\Big) \\ &= \lim_{n\to\infty} \sum_{i=1}^n \lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\Big(X_N(t_i) - X_N(t_{i-1}) \approx \dot{\gamma}_{t_{i-1}}(t_i - t_{i-1})\Big) \\ &= \lim_{n\to\infty} \sum_{i=1}^n (t_i - t_{i-1})\, L\big(\dot{\gamma}_{t_{i-1}}\big) = \int_0^T L\big(\dot{\gamma}_t\big)\, dt, \end{aligned} \tag{A.9}$$

where L is given by (A.8) and t_i, $i = 1, \ldots, n$, is a partition of the time interval $[0, T]$ that becomes dense in the limit as $n \to \infty$.

We see from the above elementary computation that, in the limit as $N \to \infty$,

$$\mathbb{P}_N\Big((X_N(t))_{t\in[0,T]} \approx (\gamma(t))_{t\in[0,T]}\Big) \approx \exp\left[-N \int_0^T L(\gamma_t, \dot\gamma_t)\, dt\right], \tag{A.10}$$

where the *Lagrangian* L only depends on the second variable, namely,

$$L(\gamma_t, \dot\gamma_t) = L(\dot\gamma_t) \tag{A.11}$$

with L given by (A.8). We interpret (A.10) as follows: if the trajectory is not differentiable at some time $t \in [0, T]$, then the probability in the left-hand side of (A.10) decays superexponentially fast with N, i.e.,

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N\Big((X_N(t))_{t\in[0,T]} \approx (\gamma_t)_{t\in[0,T]}\Big) = -\infty, \tag{A.12}$$

and otherwise it is given by the formula in (A.10) (read in the standard large-deviation language).

The *Lagrangian* in (A.8) is the Legendre transform of the *Hamiltonian*

$$H(\lambda) = b\big(e^{\lambda} - 1\big) - d\big(e^{-\lambda} - 1\big). \tag{A.13}$$

This Hamiltonian can be obtained from the generator in (A.1) as follows:

$$H(\lambda) = \lim_{N\to\infty} \frac{1}{N}\, e^{-Nf_\lambda(x)} \left(L_N e^{Nf_\lambda}\right)(x), \qquad f_\lambda(x) = \lambda x. \tag{A.14}$$

More generally, by considering the operator

$$(\mathcal{H}f)(x) = \lim_{N\to\infty} \frac{1}{N}\, e^{-Nf(x)} \left(L_N e^{Nf}\right)(x) = b\big(e^{f'(x)} - 1\big) - d\big(e^{-f'(x)} - 1\big), \tag{A.15}$$

we see that the Hamiltonian equals

$$H(\lambda) = (\mathcal{H}f_\lambda)(x), \tag{A.16}$$

and that, by the convexity of $\lambda \mapsto H(\lambda)$,

$$(\mathcal{H}f)(x) = H(f'(x)) = \sup_{a\in\mathbb{R}} [af'(x) - L(a)]. \tag{A.17}$$

The operator $\mathcal{H}$ is called the *generator of the non-linear semigroup.*

A.2 The scheme of Feng and Kurtz

The scheme that produces the Lagrangian in (A.8) from the operator in (A.15) actually works in much greater generality. Consider a sequence of Markov processes $X = (X_N)_{N\in\mathbb{N}}$ with $X_N = (X_N(t))_{t\geq 0}$, living on a common state space (like $\mathbb{R}$, $\mathbb{R}^d$ or a space of probability measures). Suppose that X_N has generator L_N and in the limit as $N \to \infty$ converges to a process $(x(t))_{t\geq 0}$, which can be either deterministic (as in the previous example) or stochastic. We want to identify the Lagrangian controlling the large deviations of the trajectories:

$$\mathbb{P}_N\Big((X_N(t))_{t\in[0,T]} \approx (\gamma_t)_{t\in[0,T]}\Big) \approx \exp\left[-N \int_0^T L(\gamma_t, \dot\gamma_t)\, dt\right]. \tag{A.18}$$

Omitting technical conditions, we see that this can be done in four steps:

1. Compute the generator of the non-linear semigroup

$$(\mathcal{H}f)(x) = \lim_{N\to\infty} \frac{1}{N}\, e^{-Nf(x)} \left(L_N e^{Nf} \right)(x). \tag{A.19}$$

2. Look for a function $H(x,p)$ of two variables such that

$$(\mathcal{H}f)(x) = H(x, \nabla f(x)). \tag{A.20}$$

   What ∇f means depends on the context: on $\mathbb{R}^d$ it simply is the gradient of f, while on an infinite-dimensional state space it is a functional derivative.

3. Express the function H as a Legendre transform:

$$H(x,p) = \sup_{p} \left[\langle p, \lambda \rangle - L(x,\lambda) \right]. \tag{A.21}$$

   What $\langle \cdot \rangle$ means also depends on the context: on $\mathbb{R}^d$ it simply is the inner product, while in general it is a natural pairing between a space and its dual, such as $\langle f, \mu \rangle = \int f d\mu$.

4. The Lagrangian in (A.18) is the function L with $x = \gamma_t$ and $\lambda = \dot{\gamma}_t$.

References

[1] D. Dereudre and S. Roelly, Propagation of Gibbsianness for infinite-dimensional gradient Brownian diffusions, J. Stat. Phys. 121 (2005) 511–551.

[2] R.L. Dobrushin and S.B. Shlosman, Gibbsian description on "non-Gibbs" fields (russian), Uspekhi Mat. Nauk 52 (1997) 45–58; translation in Russian Math. Surveys 52 (1997) 285–297.

[3] R.L. Dobrushin and S.B. Shlosman, "Non-Gibbsian" states and their Gibbs description, Comm. Math. Phys. 200 (1999) 125–179.

[4] A.C.D. van Enter, R. Fernández and A. Sokal, Regularity properties and pathologies of position-space renormalization-group transformations: scope and limitations of Gibbsian theory, J. Stat. Phys. 72 (1993) 879–1167.

[5] A.C.D. van Enter, R. Fernández, F. den Hollander and F. Redig, Possible loss and recovery of Gibbsianness during the stochastic evolution of Gibbs measures, Comm. Math. Phys. 226 (2002) 101–130.

[6] A.C.D. van Enter, C. Külske, A.A. Opoku and W.M. Ruszel, Gibbs-non-Gibbs properties for n-vector lattice and mean-field models, Braz. J. Prob. Stat., 24 (2010) 226–255.

[7] A.C.D. van Enter and W.M. Ruszel, Loss and recovery of Gibbsianness for XY spins in a small external field, J. Math. Phys. 49 (2008) 125208.

[8] A.C.D. van Enter and W.M. Ruszel, Gibbsianness versus non-Gibbsianness of time-evolved planar rotor models, Stoch. Proc. Appl. 119 (2010) 1866–1888.

[9] A.C.D. van Enter and E.A. Verbitskiy, On the variational principle for generalized Gibbs measures, Markov Proc. Relat. Fields 10 (2004) 411–434.

[10] V. Ermolaev and C. Külske, Low-temperature dynamics of the Curie-Weiss model: Periodic orbits, multiple histories, and loss of Gibbsianness, Bochum-Groningen (2010) preprint.

[11] J. Feng and T.G. Kurtz, *Large Deviations for Stochastic Processes*, American Mathematical Society, Providence RI, 2006.

[12] R. Fernández, Gibbsianness and non-Gibbsianness in lattice random fields, in: Proceedings of Les Houches Summer School LXXXIII "Mathematical Statistical Physics" (eds. A. Bovier, J. Dalibard, F. Dunlop, A. van Enter and F. den Hollander), Elsevier, 2006, pp. 731–798.

[13] R. Fernández and C.-E. Pfister, Global specifications and non-quasilocality of projections of Gibbs measures, Ann. Prob. 25 (1997) 1284–1315.

[14] C. Külske and A. Le Ny, Spin-flip dynamics of the Curie-Weiss model: loss of Gibbsianness with possibly broken symmetry, Comm. Math. Phys. 271 (2007) 431–454.

[15] C. Külske and A.A. Opoku, The posterior metric and the goodness of Gibbsianness for transforms of Gibbs measures, Elect. J. Prob. 13 (2008) 1307–1344.

[16] C. Külske and A.A. Opoku, Continuous mean-field models: limiting kernels and Gibbs properties of local transforms, J. Math. Phys. 49 (2008) 125215.

[17] C. Külske and F. Redig, Loss without recovery of Gibbsianness during diffusion of continuous spins, Probab. Theory Relat. Fields 135 (2006) 428–456.

[18] A. Le Ny and F. Redig, Short time conservation of Gibbsianness under local stochastic evolutions, J. Stat. Phys. 109 (2002) 1073–1090.

[19] A. Le Ny and F. Redig, Large deviation principle at fixed time in Glauber evolutions, Markov Proc. Relat. Fields 10 (2004) 65–74.

[20] T. Liggett, *Interacting Particle Systems*, Springer, Berlin, 1985.

[21] C. Maes, F. Redig and A. Van Moffaert, Almost Gibbsian versus weakly Gibbsian measures, Stoch. Proc. Appl. 79 (1999) 1–15.

[22] C. Maes and K. Netocný, Spacetime expansions for weakly coupled interacting systems, J. Phys. A: Math. Gen. 35 (2002) 3053–3077.

[23] A.A. Opoku, *On Gibbs Properties of Transforms of Lattice and Mean-Field Systems*, PhD thesis, Groningen University, 2009.

[24] C.-E. Pfister, Thermodynamical aspects of classical lattice systems, in " In and Out of Equilibrium" (Mambucaba 2000), vol 51 of Prog. Prob. (ed. V. Sidoravicius), Birkhäuser, 2002, pp. 393–472.

[25] F. Redig, S. Roelly and W.M. Ruszel, Short-time Gibbsianness for infinite-dimensional diffusions with space-time interaction, J. Stat. Phys. 138 (2010) 1124–1144.

[26] Proceedings of the conference "Gibbs versus non-Gibbs in Statistical Mechanics and Related Fields", December 2003, EURANDOM, Eindhoven, The Netherlands (eds. A.C.D. van Enter. A. Le Ny and F. Redig), Markov Proc. Relat. Fields 10 (2004).