SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
I STUDI AVANZATI
TRIESTE
Strada Costiera 11

ISAS - INTERNATIONAL SCHOOL
FOR ADVANCED STUDIES

Long–time behaviour of time–inhomogeneous evolutions with applications to Langevin diffusions

Thesis submitted for the degree of
"Doctor Philosophiae"

Functional Analysis Sector

Candidate:
Gabriele Grillo

Supervisor:
Prof. Alberto Frigerio

Academic Year 1991/92

TRIESTE

Long–time behaviour of time–inhomogeneous evolutions with applications to Langevin diffusions

Thesis submitted for the degree of

"Doctor Philosophiae"

Functional Analysis Sector

Candidate:

Gabriele Grillo

Supervisor:

Prof. Alberto Frigerio

Academic Year 1991/92

CONTENTS

INTRODUCTION

The main goal of this thesis is to discuss the long–time behaviour of a class of time–inhomogeneous evolutions on von Neumann algebras (commutative or not). To explain the origin of what has been done, we briefly discuss what is now known as the "*simulated annealing*" algorithm [1–9].

Let X be a finite set with $|X|$ points, whose elements describe the states of some (fictitious) physical system, and $\beta_n \to \infty$ be a monotonically nondecreasing sequence of nonnegative numbers. Let there be given a function $U : X \to [0,\infty)$, to be interpreted as the energy of the above mentioned system, and consider a Markov chain X_n on X with $X_0 = x \in X$ and with transition probabilities given, for $x \neq y$, by

$$P\left[X_n = y \mid X_{n-1} = x\right] = q_0(x,y)\exp\left[-\beta_n(U(y) - U(x))_+\right], \tag{1}$$

q_0 being a symmetric matrix.

Intuitively, starting from a point x at "time" $t = 0$, one chooses a new point y with probability $q_0(x,y)$. Then the jump from x to y is actually performed, with probability $\exp\left[-\beta_n(U(y) - U(x))_+\right]$ (conditional upon the choice of y), when the energy difference $U(y) - U(x)$ is positive, and with probability one otherwise.

The system is thus allowed to make energy increasing transition, therefore it can climb out of local minima of U; infact, what has been proved is that, under suitable assumptions on the sequence β_n, the distribution of X_n tends to be concentrated, as $n \to \infty$, on the set of absolute minima of U.

To give an idea of one possible strategy for proving that, note that, if β_n were constant and equal to β, one can show that iteration of the above procedure drives the system to the Gibbs distribution $\mu_\beta = Z_\beta^{-1}\exp[-\beta U]$, Z_β being the normalization constant. Besides, it is known (cf. [10] for the more general case of Gibbs measures on $\mathbb{R}^n$) that the weak limit of μ_β as $\beta \to \infty$ exists, and that it is a measure concentrated on the set of absolute minima of U. Therefore one can try to compare the distribution of $X_{\bar n}$ at step $\bar n$ with what would be the asymptotic equilibrium distribution $\mu_{\beta_{\bar n}}$ if β_n was constant and equal to $\beta_{\bar n}$ for all n.

In turn, if β_n tends to ∞ sufficiently slowly (typically, $\beta_n = c^{-1}\log n$, c not smaller than a critical constant $c_0 > 0$), it can be proved that, at "time" $t = n$, the distribution of X_n is close (in the weak topology) to the Gibbs measure μ_{β_n}, and asymptotically the same objects become indistinguishable; this can be used to prove the above assertion.

A very similar idea underlines the introduction of the following class of diffusion processes in $\mathbb{R}^d$ (cf. [11–15]), which are the continuous space, continuous time version of simulated annealing, and which are often called *Langevin algorithms.* Consider a function $U : \mathbb{R}^d \to [0,\infty)$ of class C^2, and a piecewise continuous function $\beta : [0,\infty) \to (0,\infty)$. Then we can study the following ordinary stochastic differential equation in $\mathbb{R}^d$:

$$dx_t = -\nabla U(x_t)dt + \sqrt{\frac{2}{\beta(t)}}dw_t, \quad x_0 = x \in \mathbb{R}^d, \tag{2}$$

w_t being a standard d–dimensional Brownian motion. One can think of $\sqrt{2/\beta(t)}dw_t$ as a "thermal noise" which allows x_t to climb out of the local minima of U. Once more, if $\beta(t) = c^{-1}\log(t_0 + t)$, $c \geq c_0 > 0$, $t_0 > 1$ one can prove that, at least under suitable assumptions on U, the distribution of (2) tends weakly to a measure concentrated on the set of absolute minima of U.

The critical constant c_0 is in many cases determined, informally speaking, as follows: let $\gamma_{x,y} : [0,1] \to \mathbb{R}^d$ be a C^1 path joining x to y, and let its elevation be defined as $\mathrm{Elev}(\gamma_{x,y}) = \{\max[U(\gamma_{x,y}(t)] \,|\, t \in [0,1]\}$. Define the energy gain between x and y as $H(x,y) = \min\{\mathrm{Elev}(\gamma), \gamma \text{ joining } x \text{ and } y\}$. Consider a global minimum x_0 and, for each $x \in \mathbb{R}^d$, consider a path γ^*_{x,x_0} between x and y which requires the least energy gain. Finally, make this least energy gain as large as possible taking the supremum as (x, x_0) vary over $\mathbb{R}^{2d}$, provided that it is finite, as is usually assumed. The resulting quantity turns out to be the critical constant c_0 mentioned before; if $c < c_0$ and in general in the case in which $\beta(t)$ diverges faster than logarithmically, there are explicit examples in which the distribution of x_t satisfying (2) does not converge weakly. Similar consideration also hold for the Markov chain (1).

The constant c_0 has another interesting characterization (cf. [6, 12, 15–17]); let $L_\beta = \beta\triangle - \nabla U \cdot \nabla$ be the generator of (2). Under suitable assumptions on U, L_β is, for all fixed β, a nonpositive essentially self–operator on C_0^∞ as an operator acting in the Hilbert space $\mathrm{L}^2(d\mu_\beta)$, with $\lambda = 0$ as (simple) highest eigenvalue and a gap $-\gamma(\beta) < 0$ between $\lambda = 0$ and the rest of its L^2–spectrum. It can be proved that

$$\lim_{\beta\to\infty} -\frac{1}{\beta}\log(\gamma(\beta)) = c_0. \tag{3}$$

Therefore, the problem of finding the critical constant can be rephrased in term of the spectral analysis for the operator L_β in what is usually called the *semiclassical limit* (cf. [18–27]). In more general situations than the commutative examples that we have just sketched and in which the intuitive description of c_0 in term of paths does not make sense, it is this latter concept which plays a fundamental role.

In the first part of this thesis, we concentrate on stochastic differential equation of the form (2). We prove some detailed convergence result of the previous type in the form of bounds on the Radon–Nikodym derivative ν_t of the distribution p_t of x_t with respect to the "instantaneous" equilibrium distribution $d\mu_\beta = Z_\beta^{-1}\exp[-\beta U]dx$. Specifically, we bound from above the $\mathrm{L}^q(t)(d\mu_{\beta(t)})$–norm of p_t, $q(t)$ being a suitable function with $q(t) \to \infty$ as $t \to \infty$. This choice of q is motivated by the fact that the larger q is, the faster the convergence of p_t takes place. This bounds give also information about the probability of being away from global minima at time t, information potentially more important for applications than the statement of weak convergence in itself.

While L^2–bounds turn out to be technically easy to establish, L^q–bounds are more tricky. Indeed, it is necessary to make use of a family of logarithmic Sobolev inequalities for the Dirichlet forms of L_β, as was originally pointed out in [6] at least for the Markov chain (1). These inequalities are proven here for the case in which U grows at infinity sufficiently fast, the quadratic potential being the borderline case. These inequality are closely related to what is called the intrinsic hypercontractivity (or ultracontractivity) of the semigroup generated by the operators L_β, which amounts to its boundedness between certain weighted L^p–spaces; see the original paper by Gross [28] and [29] and reference quoted therein for a recent discussion and a thorough bibliography. We just refer the interested reader to [30–32] for applications of the theory of logarithmic Sobolev inequalities to the spectral theory of Schrödinger operators and to lattice statistical mechanics.

Moreover, the non–compactness of state space yields some further complications, which are solved by modifying some well–known large deviation result contained for example in [6, 33, 34].

The discussion of chapter 1 raises one more technical problem. Indeed, all the estimates contained there rely on the existence of L^2–solutions to a time–dependent heat equation with unbounded and time–dependent potentials. This problem is tackled in chapter 2, which is somewhat detached from the rest of the paper, in which we show that at least three different approaches can be used to solve it, each of them requiring some further assumption on U. In particular, we first give a probabilistic proof which makes use of a (suitably established) time–dependent Feynman–Kac formula for the potential considered [35] and of ultracontractive estimates [36], and then show how general theorems due respectively to Lions [37] and to Acquistapace and Terreni [38, 39] can be used in the case at hand.

Chapter 3 is devoted to a generalization of the above results, at least of those involving estimates in L^2–norm. In fact, we study eq. (2) supposing in addition that the energy function U itself depend on time, and which we hope could be of some interest in connection with adaptive algorithms (cf. [40]). We have included a section in which the problem is discussed in the (computationally simpler) setting of compact riemannian manifolds, since the discussion seems to be new even in that case. We give conditions on $U(t)$ and on its time derivative under which the distribution of the process and the instantaneous equilibrium distribution become indistinguishable in the course of time, and L^2–estimates on the corresponding Radon–Nikodym derivative. The conditions on $U(t)$ do not necessarily imply that it converge pointwise as $t \to \infty$.

The following chapter concerns non–commutative generalizations of the above results. To understand what has been done, note that both processes defined by (1) and (2) give rise to a time–inhomogeneous evolution on an appropriate commutative von Neumann algebra. Indeed, in case (1), we can define a family of maps from $L^\infty(X) = \mathbb{C}^{|X|}$ into itself by

$$[T_n^{(1)}(f)] = \sum_{y \in X} P[X_n = y \mid Y_{n-1} = x] f(y), \quad f \in \mathbb{C}^{|X|}. \tag{4}$$

From now on, we restrict to the discrete time setting. In case (2), this amounts to taking β_n piecewise constant, so that $\beta(t) = \beta_n$ for $t \in [t_{n-1}, t_n)$. Let $\mathbb{E}$ denote expectation with respect to Wiener measure and define a family of maps from $L^\infty(\mathbb{R}^d)$ into itself as follows:

$$[T_n^{(2)}(f)](x) = \mathbb{E}\left(f[x(t_n)] \mid x_{t_{n-1}} = x\right). \tag{5}$$

Under very general conditions on U, β_n, the stochastic differential equation (2) does not blow up to infinity in finite time. Therefore it is easily seen that both maps $T_n^{(1)}$, $T_n^{(2)}$ are positivity and identity preserving weakly*–continuous normal linear maps on the corresponding commutative von Neumann algebras $\mathbb{C}^{|X|}$ and $L^\infty(\mathbb{R}^d)$, *dynamical maps* in the sequel (for terminology, we refer to [41]). As concerns the theory of dynamical maps and more generally of *dynamical semigroups*, which is outside the aim of this thesis, we refer to [42, 43] and references quoted therein.

Each $T_n^{(1)}, T_n^{(2)}$ has a unique faithful normal invariant state (positive linear functional μ on $\mathcal{M}$ with $\mu(\mathbf{1}) = 1$), given by the associated Gibbs measure. It may also be noted that in both cases T_n is μ_n–symmetric, in the sense that

$$\mu_n(fT_n(g)) = \mu_n(T_n(f)g); \tag{6}$$

here f, g belongs to the appropriate von Neumann algebra: this condition is often called of *detailed balance*. It can be seen that proofs of the kind we have sketched previously depend essentially only on:
i) estimates on the spectral gap of T_n extended to a contraction on the GNS space of $(\mathcal{M}, \omega_n)$;
ii) estimates on the difference (in the norm topology) between μ_n and μ_{n-1}.

In order to have a non–commutative generalization of the previous time–inhomogeneous evolutions, we replace the commutative algebra L^∞ with a general von Neumann algebra $\mathcal{M}$ with a representation acting on a separable Hilbert space $\mathcal{H}$, and having a cyclic and separating vector Ω. Then we consider a family of dynamical maps T_n, each of them having a unique faithful normal invariant state μ_n; for example, one could take

$$\mu_n = \langle \Omega_n, A\,\Omega_n \rangle, \; A \in \mathcal{M}, \tag{7}$$

where

$$\Omega_n = Z_{\beta_n}^{-1/2} J_\Omega \exp[-\beta_n H/2]\Omega, \tag{8}$$

H is a non–negative self–adjoint element of $\mathcal{M}$, J_Ω denotes the modular involution associated with the pair $(\mathcal{M}, \Omega)$ and Z_β is the normalization constant. If $\lambda = 0$ is in the point spectrum of H, then the weak limit of μ_β exists and equals

$$\mu_\infty(a) = \langle P_0\Omega, A\,P_0\Omega \rangle, \tag{9}$$

P_0 being the orthogonal projection onto the eigenspace corresponding to the eigenvalue $\lambda = 0$.

For any initial normal state φ_0, consider its time evolved defined by

$$\varphi_n = \varphi_0 \circ \tau_1 \circ \ldots \circ \tau_n. \tag{10}$$

In the commutative case, this definition gives nothing else than the probability distribution at time n starting from an initial probability distribution φ_0.

These definitions allow to state the problem in the following more general form: under which conditions φ_n and μ_n become indistinguishable in the course of time? An affirmative answer to this question depends exactly on estimates of the type *i)*–*ii)* above. We do not address ourself to question *i*), only reminding the reader that some results have been obtained for finite quantum system (cf. the discussion of [44, sec. 5], which we briefly review in section A.1), and for a class of infinite quantum systems in [45]. Concerning the problem *ii*) we give conditions in terms of the cyclic and separating vectors which represent μ_n, μ_{n-1}. These conditions are them rephrased in terms of the relative hamiltonians between the above couple of states. We want to remark that no detailed balance condition is necessary in our discussion, provided that one can prove a spectral gap condition without it.

The previous setting can be used successfully mainly for finite quantum systems. This can be seen from the fact that, when specializing the calculations to commutative algebras, the conditions involved are not satisfied if the corresponding energy function U is unbounded. We have tried to overcome this difficulty in section 4.4, but our results there are only partial. Indeed, it is not possible to describe the case in which μ_n and μ_m, $m \neq n$ are KMS (i.e. thermal) states at different temperatures on the C^*–algebra $\mathcal{A}$ of observables for an infinite quantum system, since in this case one expects in general that the two states are disjoint, in the sense that no subrepresentation of the GNS representation for the first state is unitarily equivalent to a subrepresentation of the GNS representation for the second state, and this can not be described in our setting due to technical complications which we were not able to overcome. We think that in such connection one should work in continuous time and have at one's disposal some sort of logarithmic Sobolev inequality in order to be able to proceed as in chapter 1; however a general theory of non-commutative hypercontractivity is far from being at hand.

Nevertheless, we give a generalization of the previous arguments which allows to deal, when working on a *fixed* von Neumann algebra, with some cases which correspond in the commutative situation to an unbounded energy function.

Chapter 5 is devoted to some applications of the above results, obtained by particular choices of the von Neumann algebra $\mathcal{M}$. The commutative examples are the usual ones, but with the additional bonus of being able to deal, without further technical difficulties, with a class of time–dependent potentials. In particular we study simulated annealing on finite sets and Langevin algorithms on compact manifolds with piecewise constant (as a function of time) energy function and temperature. Similar results also hold for a class of jump processes on compact manifolds which resemble simulated annealing (see Remark 5.9), provided that the spectral gap for their generators have the usual asymptotic behaviour. Section 5.3 describe a class of time–inhomogeneous evolution on the von Neumann algebra of all bounded operators on a separable Hilbert space.

Finally, the Appendix has a review nature, and concerns the main lines of the proofs of theorems concerning the spectral gap behaviour for the generators of the annealing algorithm on finite sets and of Langevin diffusions on $\mathbb{R}^d$.
Part of the arguments of this thesis have been discussed in [46–50].

Chapter 1:

LANGEVIN DIFFUSIONS

Let $U: \mathbb{R}^n \to \mathbb{R}$ be a function of class C^2 which is bounded below, and $T: [0,\infty) \to [0,\infty)$ be a function of class C^1; in this section we study ordinary stochastic differential equations in $\mathbb{R}^n$ of the form

$$dx_t = -\nabla U(x_t)dt + \sqrt{2T(t)}dw_t, \tag{1.0.1}$$

w_t being a standard n–dimensional Brownian motion. We shall prove in sections 1.1, 1.2, under suitable assumptions on U and T, the weak convergence of the distribution of the process to a measure concentrated on the set of global minima of U. First, we rephrase the problem in terms of a heat equation with time dependent sources, and then prove L^2–estimates on the rate of convergence of the distribution of (1.0.1). Then, we prove in section 1.3 a family of (weighted) logarithmic Sobolev inequalities for the generators L_T of the diffusions corresponding to fixing $T(t) = T$ =const. in (1.0.1), and we use such inequalities to prove, by means of large deviation techniques, some quantitative bounds on the above mentioned rate of convergence in the form of $L^{q}(t)$–estimates (with $q(t) \to +\infty$ as $t \to +\infty$) on the Radon–Nikodym derivative of the distribution of (1.0.1) with respect to the "instantaneous" equilibrium distribution.
In particular, this gives quantitative estimates about the efficiency of the Langevin algorithm as a tool for global minimization in $\mathbb{R}^n$. Our general references for the theory of stochastic differential equations will be [51–52].

In the sequel, U will be assumed to be a function of class C^2 defined on $\mathbb{R}^n$ with values in $\mathbb{R}$, which is bounded below. For the sake of notational simplicity, we shall assume in this chapter that $\min U = 0$ unless otherwise stated.

1.1 Statement of the problem

Let there be given a nonincreasing function of class C^1 $t \mapsto T(t)$ of $[0,\infty)$ to $[0,\infty)$. It will be technically convenient in the sequel to re-express t as a function $t = t(s)$ of a new parameter s such that $t(0) = 0$ and

$$\frac{dt}{ds} = \beta(s) = \frac{1}{T(t(s))}; \tag{1.1.1}$$

then $s \mapsto \beta(s)$ is a nondecreasing function of $[0,\infty)$ into $[0,\infty)$, with

$$\beta'(s) = \frac{d}{ds}\beta(s) = \frac{1}{T(t)}\frac{d}{dt}\frac{1}{T(t)}\Big|_{t=t(s)}. \tag{1.1.2}$$

$\beta(s)$ will have to increase to ∞ as $s \to \infty$, but slowly enough. For the time being, we assume that $\beta'(s)$ is bounded by a constant b for all s, so that $\beta(s) \le \beta(0) + bs$ and

$t(s) \leq \beta(0)s + bs^2/2$. Possibly by changing again the scale of time and the energy units, we assume that $\beta(s) \geq 2$, $\beta'(s) \leq 2$ for all $s \geq 0$.
We define a family $\{T_s : 0 \leq s < \infty\}$ of positivity- preserving linear maps of $\mathrm{L}^\infty = \mathrm{L}^\infty(\mathbb{R}^n, dx)$ into itself by

$$[T_s g](x) = \mathbf{E}_x^{(W)} \left[g(x_{t(s)}) \right] \quad : g \in \mathrm{L}^\infty, \tag{1.1.3}$$

where $\mathbf{E}_x^{(W)}$ stands for expectation with respect to Wiener measure starting at $x \in \mathbb{R}^n$, x_t is the solution to the stochastic differential equation (1.0.1), which is in principle defined up to an explosion time τ, and $t(s)$ is defined in (1.1.1). An application of Has'minskii's stochastic Liapunov theorem [53] shows that, under very general conditions on U, x_t does not blow up to ∞ in finite time with probability one, so that $T_s(\mathbf{1}) = \mathbf{1}$ for all s. In fact, the simplest choices of the Liapunov function V are either $V_1(x) = |x|^a + b$, $a > 0, b \geq 0$ or $V_2(x) = U(x) + c$, $c \geq 0$. The corresponding Liapunov inequalities amount respectively to:

i) there exists a constant $k > 0$ and a compact set $K \subset \mathbb{R}^n$ such that

$$(x \cdot \nabla U(x))_- \leq k|x|^2, \qquad x \in K^c. \tag{1.1.4}$$

ii) Let β_0 be the (strictly positive) lower bound of $1/T(t)$; there exist constants $l, m > 0$ such that

$$\triangle U(x) \leq lU(x) + \beta_0 |\nabla U(x)|^2 + m. \tag{1.1.5}$$

Here a subscript "$_-$" (resp. "$_+$" denotes the negative (resp. positive) part. We shall see in section 1.3 (see in particular Lemma 1.7) and in chapter 2 that (1.1.5) is implied by the hypercontractivity assumptions which will be crucial there, and we shall therefore *assume* in the sequel that the stochastic differential equation (1.0.1) is non–explosive.
In this section we also assume without further comments that

$$Z_\beta = \int e^{-\beta U} dx \tag{1.1.6}$$

and

$$\langle U \rangle_\beta = \int U e^{-\beta U} dx \tag{1.1.7}$$

are finite. In fact, in the following sections we shall be forced to impose stricter conditions on U; hereafter, all integrals will be performed on $\mathbb{R}^n$ unless otherwise stated. In particular, we shall often deal with the family of probability measures $\{\mu_\beta\}_{\beta \in (0,\infty)}$, defined by

$$d\mu_\beta(x) = Z_\beta^{-1} e^{-\beta U(x)} dx. \tag{1.1.8}$$

For g in a suitable domain (containing at least the space C_0^∞ of C^∞ functions of compact support), we have by Ito's formula

$$\frac{d}{ds} T_s g = -T_s(L_{\beta(s)} g), \tag{1.1.9}$$

where

$$L_\beta g = -\Delta g + \beta (\nabla U) \cdot (\nabla g). \tag{1.1.10}$$

L_β extends to a positive self-adjoint operator in $\mathrm{L}^2(\mathbb{R}^n, \Phi_\beta^2 dx)$ which we still denote by L_β, where Φ_β is the function $x \mapsto \Phi_\beta(x)$ defined by

$$\Phi_\beta(x) = Z(\beta)^{-1/2} e^{-\beta U(x)/2}. \tag{1.1.11}$$

Note that L_β can be also defined as the self-adjoint operator associated with the closure of the positive quadratic form

$$Q_\beta(g) = \int |\nabla g|^2 \Phi_\beta^2 dx \quad : \quad g \in C_0^\infty. \tag{1.1.12}$$

The Hilbert space $\mathrm{L}^2(\mathbb{R}^n, \Phi_\beta^2 dx)$ is mapped unitarily onto $\mathrm{L}^2 = \mathrm{L}^2(\mathbb{R}^n, dx)$ by $g \mapsto g\Phi_\beta$; under this mapping L_β is unitarily equivalent to the operator H_β given on C_0^∞ by

$$H_\beta f = \Phi_\beta L_\beta(\Phi_\beta^{-1} f) = (-\Delta + V_\beta) f, \tag{1.1.13}$$

where V_β is the operator of multiplication by the function

$$V_\beta(x) = \frac{\beta^2}{4} |\nabla U(x)|^2 - \frac{\beta}{2} \Delta U(x). \tag{1.1.14}$$

Indeed, H_β is the self-adjoint operator associated with the closure of the positive quadratic form

$$\mathcal{H}_\beta(g) = \int \left(|\nabla g|^2 + V_\beta |g|^2 \right) dx \quad : \quad g \in C_0^\infty. \tag{1.1.15}$$

Schrödinger operators of the form (1.1.13) are well understood. It is known that 0 is a nondegenerate eigenvalue of H_β, with the unique positive eigenvector Φ_β, and that the spectrum of H_β lies in $[0, \infty)$. Moreover, if $V_\beta(x) \to \infty$ as $|x| \to \infty$, as will be true in our assumptions (see in particular section 1.3 and chapter 2), then H_β has compact resolvent, with eigenvalues

$$0 = \gamma_0(\beta) < \gamma(\beta) \le \gamma_2(\beta) \le \cdots$$

The asymptotic behaviour of the *spectral gap* $\gamma(\beta)$ has been studied in several contexts (cf. for example [6, 7, 12, 15–18, 25, 26]. It is related with the problem of the semiclassical limit (as Planck's constant $\hbar \to 0$) in quantum mechanics where, however, one usually considers operators of the form $H = -\Delta + \hbar^{-2} V$ and not of the form $H = -\Delta + \hbar^{-2} V_2 + \hbar^{-1} V_1$. In several circumstances, there exists a number m such that

$$\lim_{\beta \to \infty} -\frac{1}{\beta} \log [\gamma(\beta)] = m \tag{1.1.16}$$

(see section A.2), and the number m can be described as follows (see section A.2). Let U have, for the sake of simplicity, finitely many, isolated, nondegenerate critical points

$x_1, \ldots, x_k$ (for more general conditions, see section A2). For each continuous path $\gamma : [0,1] \mapsto \mathbb{R}^n$, let

$$m(\gamma) = \max_{s\in[0,1]} \left[U(\gamma(s)) - U(\gamma(0)) - U(\gamma(1))\right].$$

For any two points $x, y \in \mathbb{R}^n$, let $m(x,y) = \min\{m(\gamma) \mid \gamma(0) = x,\ \gamma(1) = y\}$. Finally, let $m = \max\{m(x,y) \mid x, y \in \mathbb{R}^n\}$. It can be proved that if m is attained by $m(x,y)$, then either x or y (or both) is an absolute minimum for U, in which $U = 0$; so that if U has a unique global minimum m is the maximum height a point must gain in order to reach the global minimum on a path which takes the lowest passes. The natural setting for these results is that of differentiable manifolds.

Suppose that we are given an initial probability density function p_0, so that $p_0 \in \mathrm{L}^1 = \mathrm{L}^1(\mathbb{R}^n, dx)$ and, for any Borel set B in $\mathbb{R}^n$,

$$\mathrm{Prob}\,[x_0 \in B] = \int_B p_0(x)dx = \int p_0(x) I_B(x)dx, \tag{1.1.17}$$

where I_B denotes the indicator function of the set B. Let $\{p_s\}_{s\geq 0}$ be the probability distribution of (1.0.1), so that

$$\mathrm{Prob}\left[x_{t(s)} \in B\right] = \int p_0(x)\left(T_s I_B\right)(x)dx = \int p_s(y) I_B(y)dy. \tag{1.1.18}$$

Suppose that, for all $s \geq 0$, there exists $f_s \in \mathrm{L}^2$ such that

$$p_s(x) = f_s(x)\Phi_{\beta(s)}(x) \ :\ f_s \in \mathrm{L}^2, \tag{1.1.19}$$

Note that if (1.1.19) holds one has $\mathrm{Prob}\,[x_{t(s)} \in B] = (f_s, I_B\Phi_{\beta(s)})$. Then we have the following

Lemma 1.1. *If there exists a family $\{f_s\}_{s\geq 0} \subset \mathrm{L}^2$ satisfying (1.1.19), then f_s must satisfy the following time-dependent heat equation*

$$\frac{d}{ds} f_s = -H_{\beta(s)} f_s + \frac{1}{2}\beta'(s)\left(U - \langle U\rangle_{\beta(s)}\right) f_s = -\left(-\triangle + V(s)\right) f_s, \tag{1.1.20}$$

where $V(s)$ is the operator of multiplication by $V(s,x)$ given by

$$V(s,x) = V_{\beta(s)}(x) - \frac{1}{2}\beta'(s)\left(U(x) - \langle U\rangle_{\beta(s)}\right), \tag{1.1.21}$$

and $\beta'(s) = (d/ds)\beta(s)$.

Proof. Let $g \in C_0^\infty$; then we have on the one hand

$$\frac{d}{ds}\left(f_s, g\Phi_{\beta(s)}\right) = \frac{d}{ds}\int p_0 T_s g dx = -\int p_0 T_s \left(L_{\beta(s)} g\right) dx =$$

$$= -\left(f_s, \left[L_{\beta(s)} g\right]\Phi_{\beta(s)}\right) = -\left(f_s, H_{\beta(s)}\left[g\Phi_{\beta(s)}\right]\right), \tag{1.1.22}$$

and on the other hand we have, whenever the derivative exists,

$$\frac{d}{ds}\left(f_s, g\Phi_{\beta(s)}\right) = \left(\frac{d}{ds}f_s, g\Phi_{\beta(s)}\right) + \left(f_s, g\frac{d}{ds}\Phi_{\beta(s)}\right) =$$

$$= \left(\frac{d}{ds}f_s, g\Phi_{\beta(s)}\right) - \left(f_s, \frac{1}{2}g\beta'(s)\left[U - \langle U\rangle_{\beta(s)}\right]\Phi_{\beta(s)}\right), \tag{1.1.23}$$

□

Note that Eq. (1.1.20) involves a family of time-dependent Schrödinger operators. In chapter 2 we shall prove that, under suitable conditions on U, it has a unique solution for almost all $s > 0$ which is in L^2, and that this remain true also when the initial condition f_0 is a multiple of the Dirac δ delta measure concentrated at $x \in \mathbb{R}^n$, at least when $\beta(s)$ is constant in a neighbourhood of zero. In the remainder of this chapter, we shall *assume* that (1.1.20) has a solution with the above properties. In particular, we assume throughout, possibly by running the process at constant temperature for time $t \in (-\varepsilon, 0)$, that the Radon-Nikodym derivative ν_0 of the distribution of the process w.r.t. $P_{\beta(0)}$ exists and belongs to L^∞.

1.2. The L^2–estimates

In this section we shall prove, under suitable hypotheses, that simulated annealing works as a tool for global minimization. We shall restrict here to the computationally simplest even if quantitatively worse estimates, that is to those in L^2–norm, for the sake of clarity of exposition.
To explain what will be done in the sequel, note that

$$\mathrm{Prob}[x_{t(s)} \in B] - \mu_{\beta(s)}[B] = \left(f_s - \Phi_{\beta(s)}, I_B \Phi_{\beta(s)}\right). \tag{1.2.1}$$

By the Schwarz inequality, the absolute value of (1.2.1) is majorized by

$$N[t(s)]^{1/2} \left(I_B \Phi_{\beta(s)}, I_B \Phi_{\beta(s)}\right)^{1/2} = N[t(s)]^{1/2} \mu_{\beta(s)}[B]^{1/2}, \tag{1.2.2}$$

where

$$N(t(s)) = ||f_s - \Phi_{\beta(s)}||_2^2. \tag{1.2.3}$$

Moreover, it is well–known that, under suitable assumptions on U, one has

$$\lim_{\beta \to \infty} \mu_\beta[U \geq \min U + \delta] = 0 \tag{1.2.4}$$

for all $\delta > 0$. In particular, the weak limit μ of μ_β as $\beta \to \infty$ exists and is a measure concentrated on the set K of absolute minima of U under very general conditions: in fact, it is known that if $\mu_{\mathrm{Leb}}(K) > 0$ then the limiting measure exists and it is the uniform measure on K. If $\mu_{\mathrm{Leb}}(K) = 0$ the family $\{\mu_\beta\}_{\beta \geq 0}$ is tight, and in the case (cf. [10]) that K is the union of finitely many disjoint smooth manifolds (possibly of different dimension), and that U is of class C^3 with Hess $(U) :=$ det $(\partial^2 U / \partial t^2) \neq 0$ on K (the determinant being taken with respect to those coordinates in a tubular neighbourhood of K which are orthogonal to K), it follows that the limiting measure exists and is concentrated on the highest dimensional among these manifolds. Moreover, its density with respect to the sum m of the intrinsic measures on such manifolds exists and reads:

$$\frac{d\mu}{dm}(p) = \frac{[\mathrm{Hess}(U)(p)]^{-\frac{1}{2}}}{\int_N (\mathrm{Hess}(U))^{-\frac{1}{2}} \, dm}. \tag{1.2.5}$$

Therefore it follows that, under the above assumptions on U, one has

$$\mathrm{Prob}\left[U(x_{t(s)}) \geq \min U + \delta\right] \leq N(t(s))^{1/2} \left(\mu_{\beta(s)}[U \geq \min U + \delta]\right)^{1/2} \to 0 \text{ as } \beta \to \infty \tag{1.2.6}$$

for all $\delta > 0$, whenever $N(t)$ stays bounded. We shall prove in the remainder of this section that this is the case. In fact, we have the following

Proposition 1.2. *Assume that U is relatively form bounded with respect to H_β with relative bound strictly smaller than one, uniformly in β, so that*

$$U \leq aH_\beta + b \tag{1.2.7}$$

in the sense of quadratic forms, for some $a \in (0,1)$, $b \in \mathbb{R}$. Assume also that (1.1.16) *holds for some $m > 0$, and let $T(t) = c/\log(t + t_0)$ for some $t_0 > 1$, with $c > m$. Then $N(t) \to 0$ as $t \to \infty$. More precisely, let $p \in (1 - m/c, 1)$; then there exist constants k_1, k_2, depending on p but independent of $N(0)$, such that*

$$N(t) \leq k_1 N(0) e^{-(t+t_0)^{1-p}} + \frac{k_2}{(t + t_0)^{1-p}}. \tag{1.2.8}$$

Proof. Note that

$$N(s) = (f_s - \Phi_{\beta(s)}, f_s - \Phi_{\beta(s)}) = ||f_s||_2^2 - (f_s, \Phi_{\beta(s)}) - (\Phi_{\beta(s)}, f_s) + ||\Phi_{\beta(s)}||_2^2 =$$

$$= ||f_s||_2^2 - 1, \tag{1.2.9}$$

since $||\Phi_{\beta(s)}||_2 = 1$ and $(f_s, \Phi_{\beta(s)}) = \int p_s(x)dx = 1$. Next we compute the time derivative of (1.2.9), and we derive a differential inequality which will give us the required uniform bound.

$$\frac{d}{ds}||f_s - \Phi_{\beta(s)}||_2^2 = \frac{d}{ds}||f_s||_2^2 =$$

$$= -2(f_s, H_{\beta(s)} f_s) + \beta'(f_s, (U - \langle U\rangle_\beta) f_s). \tag{1.2.10}$$

We recall that H_β is positive and self-adjoint, and that Φ_β is an eigenvector of H_β with eigenvalue zero. Therefore one can rewrite (1.2.10) as

$$\frac{d}{ds}||f_s - \Phi_{\beta(s)}||_2^2 =$$

$$= -2(f_s - \Phi_{\beta(s)}, H_{\beta(s)}(f_s - \Phi_{\beta(s)})) + \beta'(f_s, (U - \langle U\rangle_\beta) f_s). \tag{1.2.11}$$

By assumption, $2 - \beta'(s)a > 0$ for all s; therefore, one can work out equation (1.2.10) as follows;

$$\frac{d}{ds}||f_s - \Phi_{\beta(s)}||_2^2 \leq$$

$$\leq -(2 - \beta' a)(f_s - \Phi_{\beta(s)}, H_{\beta(s)}(f_s - \Phi_{\beta(s)})) + \beta' b||f_s - \Phi_{\beta(s)}||^2 + \beta' b, \tag{1.2.12}$$

by the relative boundedness assumption. At this point, the spectral gap condition implies that

$$\frac{d}{ds}||f_s - \Phi_{\beta(s)}||_2^2 \leq$$

$$\leq -\left((2 - \beta' a)\gamma(\beta(s)) - b\beta'\right) ||f_s - \Phi_{\beta(s)}||_2^2 + b\beta'. \tag{1.2.13}$$

Restoring the original time variable t and integrating the differential inequality (1.2.13), we obtain an ordinary inequality for $N(t(s)) = ||f_s - \Phi_{\beta(s)}||_2^2$. Set

$$c(t(s)) = \frac{1}{\beta(s)} \left((2 - \beta'(s)a)\gamma(\beta(s)) - b\beta'(s) \right), \tag{1.2.14}$$

and observe that (1.2.11) can be rewritten as

$$\frac{d}{dt} N(t) = -c(t)N(t) + b\frac{d}{dt}\frac{1}{T(t)}. \tag{1.2.15}$$

It follows that

$$N(t) \leq e^{-\int_{t_0}^{t} c(u)du} N(t_0) + b \int_{t_0}^{t} e^{-\int_{u}^{t} c(v)dv} \frac{d}{du} \left(\frac{1}{T(u)} \right) du. \tag{1.2.16}$$

As is usual in the literature, we take a cooling schedule of the form

$$T(t) = \frac{c}{\log(t+d)}. \tag{1.2.17}$$

From the assumption (1.1.16) on the limiting behaviour of the spectral gap, we have that for any positive number $m^* > m$, one has

$$\gamma(\beta) \geq \Gamma e^{-\beta m^*}, \qquad \text{for some } \Gamma > 0. \tag{1.2.18}$$

Then we have

$$\begin{aligned} c(t) &= \left(2T(t) - a\frac{d}{dt}\frac{1}{T(t)} \right) \gamma \left(\frac{1}{T(t)} \right) - b\frac{d}{dt}\frac{1}{T(t)} \geq \\ &\geq \left(\frac{2c}{\log(t+d)} - \frac{a}{c(t+d)} \right) \Gamma(t+d)^{-m^*/c} - \frac{b}{c(t+d)}. \end{aligned} \tag{1.2.19}$$

For $c > m$ we may take $m < m^* < c$ so that $c(t) \geq k(t+d)^{-p}$ for some positive constant k, for some $p \in (m/c, 1)$, and for t large enough. Hence, the first term in the r.h.s. of (1.2.16) vanishes in the limit as $t \to \infty$. With suitable changes of variables, the second term is reduced to an expression of the form

$$e^{-b} \int_a^b \frac{e^x}{x} dx, \tag{1.2.20}$$

where $b = \text{const.}\ (t+t_0)^{1-p} \to \infty$; in this limit, (1.2.20) is of the order of $1/b$ (cf. [54]). This shows that also the second integral vanishes for large t, and that (1.2.8) holds, thus completing our proof.

□

Remark 1.3. Our estimates show the convergence of the annealing algorithm for $c > m$, while they do not give any information concerning the case $c \leq m$; however, in analogy with the results obtained by various authors in different contests such as compact manifolds (see [12]) or Markov chains on finite state space (see [8]), one would suspect that the process does not converge in probability when $c < m$ and converges for $c = m$.

Corollary 1.4. *Under the above assumptions,*

$$\mathrm{Prob}\,\{x_{t(s)} \geq \min U + \delta\} \leq N(t(s))^{1/2}\,(\mu_{\beta(s)}[U \geq \min U + \delta]^{1/2} \to 0 \quad \text{as } s \to \infty \tag{1.2.21}$$

for any $\delta > 0$, since N satisfies the bound (1.2.8).

Remark 1.5. The quadratic form inequality of the previous Proposition can be seen as a consequence of stricter hypercontractivity properties of the family H_β which we shall study in the following section and in chapter 2.

1.3. Logarithmic Sobolev inequalities

This section is devoted to proving a family of logarithmic Sobolev inequalities for the operators L_β of the previous section. This is done under assumptions which amount to a sufficiently fast growth of U at infinity. In fact, for potentials which are sufficiently regular in a neighbourhood of infinity our assumptions amount to $U(x) \geq c|x|^2 - d$ for some $c, d > 0$.

Assumptions of the form we shall use below will also be sufficient to prove that the time–dependent heat equation (1.1.20) has a solution in L^2 for almost all $s > 0$ (see chapter 2) and, as a simple corollary of the stochastic Liapunov theorem, to prove that the stochastic differential equation (1.0.1) is non–explosive (see Lemma 1.6 below).

The inequalities we prove here will be used in section 1.4 to prove quantitative bounds on the rate of convergence of the distribution of (1.0.1) towards a measure concentrated on the set of absolute minima of U.

Let Φ_β be defined by (1.1.11), and let $||\cdot||_{2,\beta}$ denote the norm in the Hilbert space $L^2(\mathbb{R}^n, d\mu_\beta)$. The aim of this section is to prove, under Assumption 1.6 below, a family of logarithmic Sobolev inequalities of the following form:

$$\int f^2 \log \frac{f}{||f||_{2,\beta}} d\mu_\beta \leq \varepsilon Q_\beta(f) + k_\beta ||f||^2_{2,\beta} \tag{1.3.1}$$

for some $\varepsilon \in (0,1)$, $k_\beta > 0$. Here Q_β stands for the quadratic form associated with the operator L_β acting in $L^2(\mathbb{R}^n, d\mu_\beta)$ and defined in (1.1.10). Our treatment follows [36], and aims at controlling the dependence on β of the Sobolev constants k_β (see also [55] for a discussion of the ground state representation). In order to do this, the main technical tool is the so–called Rosen's lemma, and in the sequel we estimate the dependence on β of the constants involved. As mentioned, our analysis will rely on the following

Assumption 1.6. U and $|\nabla U|$ diverge to infinity as $|x| \to \infty$ and, moreover, there exists constants $\alpha > 0$, $\varepsilon \in (0,1)$ such that the following inequality between quadratic forms holds:

$$0 \leq U(x) \leq \varepsilon(|\nabla U(x)|^2 - \Delta U(x)) + \alpha \quad \forall x \in \mathbb{R}^n. \tag{1.3.2}$$

Lemma 1.7. *The stochastic differential equation* (1.0.1) *is non–explosive.*
Proof. (1.3.2) implies that

$$\Delta U \leq |\nabla U| - \frac{1}{\varepsilon} U + \frac{\alpha}{\varepsilon} \leq \beta_0 |\nabla U|^2 + qU + \frac{\alpha}{\varepsilon}, \tag{1.3.3}$$

for any $q, \varepsilon > 0$, and hence there exist $l, m > 0$ such that (1.1.5) holds.

□

Lemma 1.8. *There exists a constant $a > 0$ such that the following inequality between quadratic forms holds:*

$$-\log \Phi_\beta \leq \varepsilon H_\beta + \beta a, \tag{1.3.4}$$

ε *being as in* (1.3.2).
Proof. It is clear from the definitions and from (1.3.2) that the following inequalities hold;

$$V_\beta = \left(\frac{1}{4}\beta^2 - \frac{1}{2}\beta\right) |\nabla U|^2 + \frac{1}{2}\beta V_2 \geq \frac{\beta}{2} V_2, \tag{1.3.5}$$

$$\varepsilon V_\beta \geq \varepsilon \frac{\beta}{2} V_2 \geq \frac{\beta}{2} U - \frac{\beta}{2}\alpha, \tag{1.3.6}$$

where α is as in (1.3.2). Hence

$$\frac{\beta}{2} U \leq \varepsilon V_\beta + \frac{\beta}{2}\alpha; \tag{1.3.7}$$

this implies that, for $a = \alpha/2 > 0$,

$$-\log \Phi_\beta \leq \varepsilon V_\beta + \beta a, \tag{1.3.8}$$

since $Z(\beta)$ is monotonically decreasing in β, and *a fortiori* (1.3.4) holds.

□

The second inequality which is necessary in order to prove Rosen's lemma is provided, at least when $n > 2$, by the following

Lemma 1.9. *Let $n > 2$. Then there exist constants $c = c(n), d > 0$ (independent of β) such that the following quadratic form inequality holds:*

$$g \leq c||g||_{\frac{n}{2}} \left[H_\beta + \beta d\right], \quad \forall g \in \mathrm{L}^{\frac{n}{2}}(dx). \tag{1.3.9}$$

Proof. Note that H_β is a positive operator, so that the Trotter product formula and (1.3.2) imply that, for some positive constant r

$$K_\beta = H_\beta + r\beta \tag{1.3.10}$$

satisfies the Beurling–Deny conditions; hence K_β is the generator of a symmetric Markov semigroup. At this point, one can proceed as follows; first, the use of the Trotter product formula shows that the integral kernel of $\exp[-tK_\beta]$ is pointwise dominated by the free heat kernel, and hence one finds that

$$||\exp[-tK_\beta]||_{1,\infty} \leq (4\pi t)^{-\frac{n}{2}}, \tag{1.3.11}$$

where $||\cdot||_{p,q}$ denotes the operator norm between $\mathrm{L}^p(dx)$ and $\mathrm{L}^q(dx)$. Interpolating between this bound and $||\exp[-tK_\beta]||_{\infty,\infty} \leq 1$ yields

$$||\exp[-tK_\beta]||_{2,\infty} \leq (4\pi t)^{-\frac{n}{4}} \tag{1.3.12}$$

At this point, we can make use of [36, Theorem 2.4.5], which asserts that a bound of the form $||\exp[-tH]||_{2,\infty} \leq c_1 t^{-\mu/4}$ for a $\mu > 2$ is equivalent to the quadratic form inequality

$g \leq c_2||g||_{\frac{\mu}{2}}H$, for all $g \in L^{\mu/2}$. The proof of this theorem shows also that the constant c_2 depends only on c_1 and on μ, thus proving the Lemma.

$\square$

Lemma 1.10. *Let ε be as in Assumption* 1.6. *There exist a positive constant c_1 (independent of β) such that the following logarithmic Sobolev inequalities hold:*

$$\int f^2 \log \frac{f}{||f||_{2,\beta}} d\mu_\beta \leq \varepsilon Q_\beta(f) + c_1\beta||f||_{2,\beta}^2, \tag{1.3.13}$$

for any $0 \leq f \in L^\infty(d\mu_\beta) \cap \mathrm{Dom}[Q_\beta]$, ε being as in (1.3.2).
Proof. If $N > 2$, this is a simple application of [36, Lemma 4.4.1, Corollary 4.4.2], and of the previous Lemmas. It is well known that inequalities of the form (1.3.13) with a Sobolev constant g_β which does not diverge too fast as $\varepsilon \to 0$ are equivalent to ultracontractive bounds for the corresponding symmetric Markov semigroups.

Even more directly, they can be used in order to estimate the corresponding integral kernels. This allows to prove the claim also in the case $N < 3$. In fact, the cases $N = 1,2$ are described by first estimating the integral kernels of the corresponding semigroups by means of tensor product operators on $L^2(\mathbb{R}^{3N})$ like $M_\beta = H_\beta \otimes 1 \otimes 1 + 1 \otimes H_\beta \otimes 1 + 1 \otimes 1 \otimes H_\beta$ as in [36, Lemma 4.5.4]; in fact, the integral kernel of $\exp[-tL_\beta]$ satisfies

$$K_{\Phi_\beta}(t,x,y) \leq a_1 \exp[\beta a_2(1+t+t^{-\gamma})], \quad t > 0,\ x,y \in \mathbb{R}^n, \tag{1.3.14}$$

for suitable positive constants $a_{1,2}, \gamma$, since the heat kernel of M_β factorizes. The application of [36, Corollary 2.2.8, Example 2.3.4] shows that the corresponding logarithmic Sobolev inequalities still hold.

$\square$

Next, the following Lemma due to J.D.Deuschel will be useful.

Lemma 1.11. *Let μ be a probability measure, and let f be in $L^2(d\mu)$. Then*

$$\int f^2 \log \left[\left(\frac{f}{||f||_{2,\mu}} \right)^2 \right] d\mu \leq \int (f - \langle f \rangle)^2 \log \left[\left(\frac{(f - \langle f \rangle)}{||(f - \langle f \rangle)||_{2,\mu}} \right)^2 \right] d\mu +$$

$$+2||f - \langle f \rangle||_{2,\mu}^2 \tag{1.3.15}$$

where $\langle f \rangle = \int f d\mu$.
Proof. See [6].

$\square$

Lemma 1.12. *There exists $m > 0$ such that*

$$\lim_{\beta \to \infty} \frac{1}{\beta} \log(\gamma(\beta)) = -m. \tag{1.3.16}$$

Proof. The thesis follows from [16] by observing that both U and $|\nabla U|$ tend to infinity as $|x| \to \infty$ by assumption, and that V_β is bounded below.

□

More precisely, it has been proved in [16] that, at least for β large enough,

$$\gamma(\beta) \geq b^{-1}\beta^{-k} \exp[-\beta m], \tag{1.3.17}$$

for suitable constants $b > 0$, $k \in \mathbb{R}$, depending on d. Once more we assume, possibly by changing the constants involved, that (1.3.17) holds for any $\beta \geq 2$. The number m will turn out to be the critical value for cooling schedules of the form $T(t) = c/\log(t_0 + t)$. Indeed, we shall prove that if $c > m$, the Langevin algorithm works.

The spectral gap estimate (1.3.17) can be rewritten in terms of an ordinary Sobolev inequality as follows:

$$||f - \langle f\rangle_\beta||^2_{2,\beta} \leq b\beta^k \exp[\beta m] Q_\beta(f), \tag{1.3.18}$$

for each f in the form domain of L_β. This fact, together with the previous Lemmas, allows us to prove the following

Theorem 1.13. *There exists $A > 0$ such that*

$$\int f^2 \log\left[\left(\frac{f}{||f||_{2,\beta}}\right)^2\right] d\mu_\beta \leq A\beta^{k+1} \exp[\beta m] Q_\beta(f), \tag{1.3.19}$$

for any $0 \leq f \in \mathrm{L}^\infty(d\mu_\beta) \cap \mathrm{Dom}\ [Q_\beta]$, $\beta \geq 2$.

Proof. The previous Lemmas imply that

$$\int f^2 \log\left[\left(\frac{f}{||f||_{2,\mu_\beta}}\right)^2\right] d\mu_\beta \leq$$

$$\begin{aligned} &\leq \int (f - \langle f\rangle_\beta)^2 \log\left[\left(\frac{(f - \langle f\rangle_\beta)}{||(f - \langle f\rangle_\beta)||_{2,\mu_\beta}}\right)^2\right] d\mu_\beta + 2||f - \langle f\rangle_\beta||^2_{2,\mu_\beta} \leq \\ &\leq 2\varepsilon Q_\beta[(f - \langle f\rangle)] + 2(c_1\beta + 1)||(f - \langle f\rangle_\beta)||^2_{2,\beta} \leq \\ &\leq 2\varepsilon Q_\beta(f) + 2(c_1\beta + 1) b\beta^k \exp[\beta m] Q_\beta(f). \end{aligned} \tag{1.3.20}$$

□

1.4. The L^q–estimates

The conclusion of Theorem 1.13 is precisely of the form required in [6, Theorem 3.9]. Let p_t denote the probability distribution of (1.0.1); the aim of this section is to use the large deviation methods originally introduced in [28] and developed for example in [6, 33] to estimate the $L^{q(t)}(d\mu_{1/T(t)})$–norm of the Radon–Nikodym derivative $\nu_t = dp_t/d\mu_{1/T(t)}$ for a suitable function $q(t)$ with $q(t) \to \infty$ as $t \to \infty$, and where $T(t)$ has the usual form $T(t) = c/\log(t_0 + t)$, $c > m$, m as in Lemma 1.12. However, in [6], inequalities of the form (1.3.19) have been used to prove estimates which depend crucially on the existence of an upper bound for U. The aim of this section is to show how to replace such estimates with bounds which do not depend on the L^∞–norm of U but only on its $L^2(d\mu_\beta)$–norm; see in particular the proof of Lemma 1.16 below. These considerations will allow us to prove L^p estimates similar to those proved in [6] also in the non–compact case, despite of the fact that U is not bounded.

We find more convenient to write the differential inequalities in the original time variable t. So, we denote by $||\cdot||_{q,t}$ the L^q–norm with respect to $\mu_{1/T(t)}$, and define $\hat{Q}_\beta = Q_\beta/\beta$; in this section, a dash will denote derivative with respect to t. We postpone to section 2.1 (see in particular Corollary 2.7) the proof the all the norms in the remainder of this section are finite, and *assume* in the sequel that this is the case. We also suppose, for the sake of simplicity, that ν_t exists even for $t = 0$, possibly (see section 2.1) by running (1.0.1) at constant "temperature" for times $(-\varepsilon, 0)$ for some $\varepsilon > 0$.

We start recalling the following

Lemma 1.14. *Let $t \mapsto q(t)$, $t \mapsto T(t)$ be C^1 positive functions such that $q(t) \geq 2$ for all $t \geq 0$. Then the following estimate holds, with $q'(t) = dq/dt(t)$:*

$$\frac{d}{dt}||\nu_t||_{q(t),t} \leq -4\frac{q(t)-1}{q^2(t)}||\nu_t||_{q(t),t}^{1-q(t)}\hat{Q}_{1/T(t)}(\nu_t^{q(t)/2}) +$$

$$+\frac{q'(t)}{q^2(t)}||\nu_t||_{q(t),t}^{1-q(t)}\int \nu_t^{q(t)}\log\left(\frac{\nu_t^{q(t)}}{||\nu_t||_{q(t),t}^{q(t)}}\right)d\mu_{1/T(t)} +$$

$$+\frac{d}{dt}\frac{1}{T(t)}\left(1-\frac{1}{q(t)}\right)||\nu_t||_{q(t),t}\int\frac{\nu_t^{q(t)}}{||\nu_t||_{q(t),t}^{q(t)}}(U - \langle U\rangle_{1/T(t)})d\mu_{1/T(t)}. \tag{1.4.1}$$

Proof. Let $P_\beta(t,\sigma,\Gamma)$ be the transition probability function of the diffusion associated with $\hat{L}_\beta = L_\beta/\beta$. We recall that μ_β is $P_\beta(t,\sigma,\Gamma)$–reversing (see [33, pg. 251]). This fact implies that, considering a process corresponding to a fixed $T = 1/\beta$, one would have

$$\frac{d}{ds}||\nu_t||_{q(t),\beta} \leq -4\frac{q(t)-1}{q^2(t)}||\nu_t||_{q(t),\beta}^{1-q(t)}\hat{Q}_\beta(\nu_t^{q(t)/2}) +$$

$$+\frac{q'(t)}{q^2(t)}||\nu_t||_{q(t),\beta}^{1-q(t)}\int \nu_t^{q(t)}\log\left(\frac{\nu_t^{q(t)}}{||\nu_t||_{q(t),\beta}^{q(t)}}\right)d\mu_\beta$$

(see also [36, pg. 67]). Taking into account the time–dependence of β and the form (1.1.20), (1.1.21) of the generator yields the claim, since

$$\frac{d\nu_t}{dt} = -\left[\hat{L}_{1/T(t)} - \frac{d}{dt}\left(\frac{1}{T(t)}\right)(U - \langle U\rangle_{1/T(t)})\right]\nu_t. \tag{1.4.2}$$

□

In order to estimate the last integral on the right hand side of (1.4.1) we shall recall the following well–known lemma (cf. for example [33]).

Lemma 1.15. *Let μ, ν be probability measures defined on the same measurable space, with $\mu << \nu$ and $d\mu/d\nu = \phi$; let $H(\mu,\nu)$ denote the relative entropy of μ with respect to ν. Then, for any positive function $\psi \in \mathrm{L}^1(d\nu)$ the following estimate holds:*

$$H(\mu,\nu) \geq \int \log\psi \; d\mu - \log \int \psi d\nu. \tag{1.4.3}$$

Proof. Recall that in the above hypothesis the relative entropy is given by

$$H(\mu,\nu) = \int \log\phi \; d\mu = \int \phi \log\phi \; d\nu, \tag{1.4.4}$$

Hence, defining the probability measure α, absolutely continuous with respect to ν, by $d\alpha/d\nu = \psi(\int \psi d\nu)^{-1}$, it follows that

$$H(\mu,\nu) = H(\mu,\alpha) + \int \log\psi \; d\mu - \log \int \psi \; d\nu. \tag{1.4.5}$$

Since, by Jensen's inequality, the relative entropy is always non–negative, (1.4.3) holds.

□

It is also well known and very simple to prove that the relative entropy is given by the supremum over the positive L^1–functions of the right–hand side of (1.4.3). The supremum might be taken only on L^∞–functions, but this is not convenient for our purposes, since we shall have to apply Lemma 1.15 to the case $\mu = \mu_\beta$, $\psi = \exp(U - \langle U\rangle_\beta)$, U being unbounded in our situation. In fact, Lemma 1.16 below is a generalization of Lemma 3.3 of [6] in the case μ and ψ are as above: however it could be stated in somewhat more general situations provided for example that $\psi = \exp\lambda$ is such that there exist $A > 0$ such that $\exp(a\lambda) \in \mathrm{L}^1(d\mu)$ for any $a \in [-A, A]$, and that $\int \lambda d\mu = 0$.

Lemma 1.16. *Let $f \in \mathrm{L}^1(d\mu_\beta)$ be positive, normalized and such that*

$$\left|\int (U - \langle U\rangle_\beta) f d\mu_\beta\right| < \infty. \tag{1.4.6}$$

Then there exists $k > 0$ such that the following estimate holds:

$$\left|\int (U - \langle U\rangle_\beta) f d\mu_\beta\right| \leq k \left(\int f \log f d\mu_\beta\right)^{\frac{1}{2}}. \tag{1.4.7}$$

Proof. Let λ be such that $\exp(\lambda) \in \mathrm{L}^1(d\mu_\beta)$. Lemma 1.15 considered in the case $\nu = \mu_\beta$, $d\mu/d\mu_\beta = f$, $\psi = \exp(\lambda)$ shows that

$$\int f \log f d\mu_\beta \geq \int f\lambda d\mu_\beta - \log \int \exp(\lambda) d\mu_\beta. \tag{1.4.8}$$

Following [6], set now $\lambda = a(U - \langle U\rangle_\beta)$, for example for a belonging to $[-1,1]$. Let $\varepsilon = |\int (U - \langle U\rangle_\beta) f d\mu_\beta|$. Let $F(a) = \log \int \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta$. Since $\int (U - \langle U\rangle_\beta) d\mu_\beta = 0$, Jensen's inequality implies that $F(a) \geq 0$. Note that by Lemma 1.16

$$\int f \log f d\mu_\beta \geq \max_{a\in[0,1]} (a\varepsilon - F(a)) := K(\varepsilon). \tag{1.4.9}$$

Since $F(0) = (dF/da)(0) = 0$, it follows that $K(\varepsilon) \geq 0$ for all $\varepsilon \geq 0$, the equality holding if and only if $\varepsilon = 0$. Finally, for any $a \in [0,1]$,

$$\begin{aligned} &\frac{d^2F}{da^2}(a) \times \int \left[\exp[a(U - \langle U\rangle_\beta)] d\mu_\beta\right]^2 = \\ &= \left[\int (U - \langle U\rangle_\beta)^2 \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta\right] \left[\int \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta\right] + \\ &\quad - \left[\int (U - \langle U\rangle_\beta) \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta\right]^2 . \end{aligned} \tag{1.4.10}$$

Jensen's inequality implies that

$$\int \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta \geq 1, \quad \forall a \in [0,1], \tag{1.4.11}$$

and hence

$$\begin{aligned} &\frac{d^2F}{da^2}(a) \leq \\ &\leq \left[\int (U - \langle U\rangle_\beta)^2 \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta\right] \left[\int \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta\right] \\ &\quad - \left[\int (U - \langle U\rangle_\beta) \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta\right]^2 . \end{aligned} \tag{1.4.12}$$

We have to estimate quantities of the form

$$\int (U - \langle U\rangle_\beta)^k \exp[a(U - \langle U\rangle_\beta)] d\mu_\beta, \quad k = 0,1,2, \quad a \in [0,1], \quad \beta \in [2,\infty). \tag{1.4.13}$$

In order to do this, the fact that U is unbounded can be circumvented as follows; first of all, since

$$\frac{d}{d\beta} \langle U\rangle_\beta = -\left(\langle U^2\rangle_\beta - \langle U\rangle_\beta^2\right) := -\mathrm{Var}_\beta U \leq 0, \tag{1.4.14}$$

$\langle U\rangle_\beta > 0$ is a bounded function of β. Furthermore, from the equation $d/d\beta[\log Z(\beta)] = -\langle U\rangle_\beta$ it follows that

$$\frac{Z(\beta_1)}{Z(\beta_2)} = \exp\left[\int_{\beta_1}^{\beta_2} \langle U\rangle_\beta d\beta\right] \leq \exp\left[(\beta_1-\beta_2)\langle U\rangle_2\right]. \tag{1.4.15}$$

Next we note that since $\langle U\rangle_\beta$ is bounded, it is sufficient to consider the case $a = 1$ in (1.4.12). In particular

$$\frac{1}{Z(\beta)}\int \exp[-(\beta-1)U]dx = \frac{Z(\beta-1)}{Z(\beta)} \tag{1.4.16}$$

is bounded as a function of β, and so is

$$\frac{1}{Z(\beta)}\int U\exp[-(\beta-1)U]dx = \langle U\rangle_{\beta-1}\frac{Z(\beta-1)}{Z(\beta)}. \tag{1.4.17}$$

Finally, consider the quantity $[1/Z(\beta)]\int U^2\exp[-(\beta-1)U]dx$. It is clear that we have to estimate the quantity $B(\beta) = \langle U^2\rangle_\beta$. Since by Jensen's inequality we have

$$\frac{dB}{d\beta}(\beta) = -\left(\langle U^3\rangle_\beta - \langle U^2\rangle_\beta\langle U\rangle_\beta\right) \leq -\langle U\rangle_\beta \mathrm{Var}_\beta U \leq 0, \tag{1.4.18}$$

$B(\beta)$ is bounded.

It follows that the right–hand side of (1.4.12) is uniformly bounded in a by a constant c_β for all $\beta \geq 2$, with c_β uniformly bounded in β. Hence there exists k_1 such that $(d^2F/da^2)(a) \leq 2k_1$ for all $a \in [0,1]$, for all $\beta \in [2,\infty]$; therefore $F(a) \leq k_1a^2$. Possibly by substituting k_1 by a sufficiently large $k_2 > k_1$, it follows that there exists $k > 0$ such that $K(\varepsilon) \geq \varepsilon/k^2$. This concludes the proof.

□

Remark 1.17. The proof of the previous Lemma shows that the constant k in (1.4.7) depends only on the $\mathrm{L}^2(d\mu_\beta)$–norm of U, which in turn is a monotonically nonincreasing function of β as shown by (1.4.18).

The following argument will be used repeatedly in the sequel

Lemma 1.18. *Consider the ordinary differential equation*

$$\frac{d}{dt}x(t) = -a(t)x(t) + b(t), \quad t \geq 0, \quad x(0) = x_0 \tag{1.4.19}$$

where for suitable positive constants $t_0, a, b, \varepsilon, \delta$ ($\delta < \varepsilon$),

$$a(t) \geq a(t_0+t)^{\varepsilon-1} \tag{1.4.20}$$

$$0 \leq b(t) \leq b(t_0 + t)^{\delta-1}. \tag{1.4.21}$$

Then the solution $x(t)$ satisfies both a bound of the form

$$x(t) \leq c_1 \exp\left[-\frac{a}{\varepsilon}(t_0+t)^{\varepsilon}\right] + c_2(t_0+t)^{\delta-\varepsilon}, \tag{1.4.22}$$

and a bound of the form

$$x(t) \leq c_3 \exp\left[-\frac{a}{\varepsilon}(t_0+t)^{\varepsilon}\right] + c_4(t_0+t)^{(\delta-1)\varepsilon} \tag{1.4.23}$$

In particular, $x(t) \to 0$ as $t \to \infty$ for all initial conditions x_0.
Proof. The solution to (1.4.19) is given by

$$x(t) = \exp\left[-\int_0^t a(u)du\right] x_0 + \int_0^t \exp\left[-\int_s^t a(u)du\right] b(s)ds. \tag{1.4.22}$$

Inserting (1.4.20) and (1.4.21) into (1.4.22) we obtain the bound

$$x(t) \leq \exp\left[-\frac{a}{\varepsilon}(t_0+t)^{\varepsilon}\right] \times$$

$$\times \left\{\exp\left(\frac{a}{\varepsilon}t_0^{\varepsilon}\right) x_0 + b\int_0^t \exp\left[\frac{a}{\varepsilon}(t_0+s)^{\varepsilon}\right](t_0+s)^{\delta-1}ds\right\}. \tag{1.4.25}$$

In order to estimate the integral in the right-hand side of (1.4.25), we must study a quantity of the form

$$I(t) = \int_0^t \exp\left[\frac{a}{\varepsilon}(t_0+s)^{\varepsilon}\right](t_0+s)^{\delta-1}ds.$$

Setting

$$z = \frac{a}{\varepsilon}(t_0+s)^{\varepsilon}$$

we find:

$$I(t) = \frac{1}{\varepsilon}\left(\frac{\varepsilon}{a}\right)^{\delta/\varepsilon} \int_{at_0^{\varepsilon}/\varepsilon}^{a(t_0+t)^{\varepsilon}/\varepsilon} e^z\, z^{\frac{\delta}{\varepsilon}-1}dz =$$

$$= \frac{1}{\varepsilon}\left(\frac{\varepsilon}{a}\right)^{\delta/\varepsilon} J\left(\frac{a}{\varepsilon}t_0^{\varepsilon}, \frac{a}{\varepsilon}(t_0+t)^{\varepsilon}, \frac{\delta}{\varepsilon}\right), \tag{1.4.26}$$

where

$$J(x,y,\alpha) = \int_x^y e^z\, z^{\alpha-1}, \quad (0 < x < y,\ 0 < \alpha < 1). \tag{1.4.27}$$

Integrating by parts one finds

$$J(x,y,\alpha) = e^y\, y^{\alpha-1} - e^x\, x^{\alpha-1} + (1-\alpha)\int_x^y e^z\, z^{\alpha-2}dz. \tag{1.4.28}$$

If $x \geq 1$ we have $z^{\alpha-2} \leq z^{\alpha-1}$ for $z \geq x$, and hence

$$J(x,y,\alpha) \leq \frac{1}{\alpha}\left(e^y\ y^{\alpha-1} - e_x\ x^{\alpha-1}\right) < \frac{1}{\alpha} e^y\ y^{\alpha-1}. \quad (x \geq 1) \tag{1.4.29}$$

In the general case we have instead

$$J(x,y,\alpha) \leq \frac{1}{\alpha} e^y\ y^{\alpha-1} + I_{(0,1)}(x) J(x,1,\alpha) \tag{1.4.30}$$

Inserting the above result into (1.4.24), we obtain

$$x(t) \leq \exp\left[-\frac{a}{\varepsilon}(t_0+t)^\varepsilon\right]\left\{\exp\left(\frac{a}{\varepsilon}t_0^\varepsilon\right)x_0 + bI_{(0,1)}J\left(\frac{a}{\varepsilon}t_0^\varepsilon, 1, \frac{\delta}{\varepsilon}\right)\right\} +$$

$$+\frac{b\varepsilon}{a\delta}(t_0+t)^{\delta-\varepsilon}, \tag{1.4.31}$$

which is indeed of the form (1.4.22). An inequality of the form (1.4.23) is proved similarly by noting that

$$e^z\ z^{\alpha-2} \leq e^y\ y^{\alpha-2}, \quad \forall z \in [x,y], \tag{1.4.32}$$

at least for sufficiently large y.

$\square$

The use of the previous Lemmas allows us to give the required uniform bound on $||\nu_t||_{q,t}$ for a suitable function $q(t) : [0,\infty) \to [2,\infty)$. In fact one finds:

Theorem 1.19. *Let $T(t) = c/\log(t_0+t)$, $t_0 > 1$, $c > m$, $\delta > 0$ and $\varepsilon \in (0, 1-m/c)$. Assume that $p_0 << \mu_{1/T(0)}$. Then there exists a constant h, depending on c, m, A, k, ε and δ such that*

$$||\nu_t||_{q(t),t} \leq ||\nu_0||_{q_0,0} \exp\left[\frac{k^2}{8}T(0)\right] \quad \forall t \geq 0 \tag{1.4.33}$$

and

$$q(t) \geq h(t_0+t)^{\varepsilon(1-\delta)}. \tag{1.4.34}$$

Proof. The following inequality holds:

$$\frac{d}{dt}||\nu_t||_{q(t),t} \leq \frac{1}{q^2(t)}||\nu_t||^{1-q(t)}_{q(t),t}\hat{Q}_{1/T(t)}(\nu_t^{q(t)/2})\left[q'(t)\hat{\alpha}\left(1/T(t)\right) - 2q(t)\right] +$$

$$+k\frac{d}{dt}\left(\frac{1}{T(t)}\right)||\nu_t||^{1-q(t)/2}_{q(t),t}\left[\hat{\alpha}\left(1/T(t)\right)\hat{Q}_{1/T(t)}(\nu_t^{q(t)/2})\right]^{\frac{1}{2}}, \tag{1.4.35}$$

since $q \geq 2$, where

$$\hat{\alpha}\left(1/T(t)\right) = \frac{1}{T(t)}\tilde{\alpha}\left(1/T(t)\right) = A\left(\frac{1}{T(t)^{k+2}}\right)\exp\left[\frac{m}{T(t)}\right]. \tag{1.4.36}$$

In addition, noting with [6] that, for all $t \in (0,\infty)$,

$$f(t) = \frac{1}{2}\left[\frac{4t}{k}\hat{\alpha}\left(1/T(t)\right)\hat{Q}_{1/T(t)}(\nu_t^{q(t)/2})||\nu_t||_{q(t),t}^{-q(t)/2} + \frac{k}{4t}||\nu_t||_{q(t),t}^{q(t)/2}\right] \geq$$

$$\geq \left[\hat{\alpha}\left(1/T(t)\right)\hat{Q}_{1/T(t)}(\nu_t^{q(t)/2})\right]^{\frac{1}{2}}, \tag{1.4.37}$$

one finds that for any function $\varrho : (0,\infty) \mapsto (0,\infty)$ the following inequality holds:

$$\frac{d}{dt}||\nu_t||_{q(t),t} \leq \frac{1}{q^2(t)}||\nu_t||_{q(t),t}^{1-q(t)}\hat{Q}_{1/T(t)}(\nu_t^{q(t)/2})\times$$

$$\times\left[\hat{\alpha}\left(1/T(t)\right)\left(q'(t) + \frac{d}{dt}\left(\frac{1}{T(t)}\right)\varrho\left(1/T(t)\right)q^2(t)\right) - 2q(t)\right] +$$

$$+\frac{d}{dt}\left(\frac{1}{T(t)}\right)\frac{k^2}{8\varrho\left(1/T(t)\right)}||\nu_t||_{q(t),t}. \tag{1.4.38}$$

Let $M(t) = ||\nu_t||_{q(t),t}$: it follows that

$$\frac{d}{dt}M(t) \leq$$

$$\leq \frac{M(t)^{1-q(t)}}{q^2(t)}\left[\left(q'(t) + q^2(t)\varrho\left(1/T(t)\right)\frac{d}{dt}\left(\frac{1}{T(t)}\right)\right)\hat{\alpha}\left(1/T(t)\right) - 2q(t)\right]\hat{Q}_{1/T(t)}(\nu_t^{q(t)/2}) +$$

$$+\frac{k^2}{8\varrho\left(1/T(t)\right)}\frac{d}{dt}\left(\frac{1}{T(t)}\right)N(t). \tag{1.4.39}$$

For any fixed function ϱ and for the usual cooling schedule, one can equate to zero the term in square brackets in the right–hand side of (1.4.39), thus finding $q = q_\varrho(t)$. The corresponding inequality for $N(t)$ is easily integrable and implies that

$$M(t) \leq M(0)\exp\left[\frac{k^2}{8}\int_{1/T(0)}^{\infty}\frac{dx}{\varrho(x)}\right]. \tag{1.4.40}$$

The bound (1.4.33) for $M(t)$ follows immediately from (1.4.40) choosing $\varrho(x) = x^2$. In addition $q(t)$ satisfies the following ordinary differential equation:

$$\frac{d}{dt}\left(\frac{1}{q(t)}\right) + \frac{2c^{k+2}}{A[\log(t_0+t)]^{2+k}(t_0+t)^{m/c}}\frac{1}{q(t)} - \frac{[\log(t_0+t)]^2}{c^3(t_0+t)} = 0. \tag{1.4.41}$$

The thesis now follows as in Lemma 1.17, since for all $\delta > 0$, $t \geq 0$ a bound of the form $\log(t_0+t) \leq c(\delta)(t_0+t)^\delta$ holds. For example, we can take $c(\delta) = 1/(e\delta)$.

□

We remark that the constant h in (1.4.34) may be very small. Indeed, the constants a, b in the Appendix may be chosen in the situation at hand as:

$$a = \frac{2}{A}\left[\left(1 - \frac{m}{c} - \varepsilon\right)\frac{ec}{k+2}\right]^{k+2}, \tag{1.4.42}$$

and

$$b = \frac{4}{(\delta e)^2}, \tag{1.4.43}$$

at least when $k > -2$. In fact, one could replace the previous bound by another one of the form

$$q(t) \geq h_1(t_0 + t)^{\varepsilon-\delta}, \ \varepsilon > \delta, \tag{1.4.44}$$

as indicated in Lemma 1.17, where the constant h_1 may in some situation be larger than the constant h in (1.4.34), or by the argument of [6], who find a bound of the form $q(t) \geq [1 + \beta(t)]/4$. These latter bounds may be better than the previous one for large but finite times, even if they are weaker in the limit as $t \to \infty$.
As a consequence of the previous Theorem and of the Schwarz inequality, we state the final result of this section.

Corollary 1.20. *Under the assumptions and with the notation of Theorem* 1.18, *one has*

$$\text{Prob}\,\{x_{t(s)} \geq \min U + \delta\} \leq$$

$$\leq ||\nu_0||_{q_0,0} \exp\left[\frac{k^2}{8}T(0)\right] \mu_{\beta(s)}\{x \mid U(x) \geq \min U + \delta\}^{1-1/\overline{q}(t)} \to 0 \quad \text{as } s \to \infty \tag{1.4.45}$$

for any $\delta > 0$, where

$$\overline{q}(t) = h(t_0 + t)^{\varepsilon(1-\delta)}. \tag{1.4.46}$$

Chapter 2:

A TIME–DEPENDENT HEAT EQUATION

The estimates in chapter 1 were obtained by exploiting the time evolution of the Radon–Nikodym derivative of the distribution of (1.0.1) with respect to the instantaneous equilibrium distribution $\mu_{\beta(t)}$ at time t. In particular, we have considered the time–dependent heat equation (1.1.20), and all our conclusions depend on the existence of L^2–solutions to such equation.

In this chapter we outline some arguments which allow to prove existence and uniqueness results for (1.1.20). The arguments are of three different kinds: namely, we first give in section 2.1 a probabilistic proof in term of ultracontractive bounds by proving a time–dependent Feynman–Kac formula for the heat equation at hand; in section 2.2 we show that, at least as concerns the existence of a weak solution, a straightforward application of Lions's theorem for parabolic equations works, and finally we use in section 2.3 the general theory of Acquistapace and Terreni for time–dependent evolution equations. This latter result is probably the better one, since it involves rather weak conditions on U; for example most functions U which increase monotonically in a neighbourhood of infinity would work.

Assumption 1.6 *is assumed throughout*, and some additional assumptions will be necessary in sections 2.1, 2.3.

2.1. Solution via a time–dependent Feynman–Kac formula

As in the previous chapter, we assume that $s \mapsto \beta(s)$ is a function of class C^1 in $[0, \infty)$, and moreover that $\beta(0) > 2$, $0 \le \beta'(s) \le 2$ for all s. We use the notation that a subscript "2" denotes that the corresponding quantity is evaluated at $\beta = 2$.
Concerning U, we make the following

Assumption 2.1. are positive constants $k, h, l, \gamma, \alpha_1, \alpha_2,$, with $\gamma < k$, such that

$$U(x) \ge k|x|^h - l, \quad |\nabla U|^2 + |\triangle U| \le \alpha_1 e^{\gamma x^h} - \alpha_2 \tag{2.1.1}$$

for all x in $\mathbb{R}^n$. Moreover, there exist positive constants c, d, δ such that the following strengthened form of Assumption 1.6 holds *for all* $\varepsilon > 0$ and for all $x \in \mathbb{R}^n$:

$$0 \le U(x) \le \varepsilon(|\nabla U(x)|^2 - \triangle U(x)) + c + \frac{d}{\varepsilon^\delta}. \tag{2.1.2}$$

Lemma 2.1. *The potential*

$$V(s) = V_{\beta(s)} - \frac{1}{2}\beta'(s)\left(U - \langle U\rangle_{\beta(s)}\right) \tag{2.1.3}$$

satisfies the bounds

$$V(s) \geq \frac{1}{2}\left[\beta(s) - \beta'(s)\right] V_2 - \frac{1}{2}\beta'(s)(c+d), \tag{2.1.4}$$

$$V(s) \geq -\frac{1}{2}\beta(s)(c+d). \tag{2.1.5}$$

Proof. For all $s > 0$, we have

$$V_{\beta(s)} = \left[\frac{1}{4}\beta(s)^2 - \frac{1}{2}\beta(s)\right] |\nabla U|^2 + \frac{1}{2}\beta(s)V_2 \geq \frac{1}{2}\beta(s)V_2.$$

In particular, it follows from Assumption 1.6 that $U \leq V_2+(c+d)$. Since $\langle U\rangle_\beta$ is positive for all β, the estimate (2.1.4) follows. From Assumption 1.6 it follows also that $V_2 \geq -(c+d)$; recalling that $\beta(s) \geq \beta'(s)$, we obtain (2.1.5).

□

Using the theory of ultracontractive semigroups, as in [36], we can prove the following

Lemma 2.2. *The semigroup* $\exp[-tH_2]$ *has an integral kernel* $K_2(t,x,y)$*, and there exist positive constants* a, b, α *such that, for all* x, y *in* $\mathbb{R}^n$ *and* t *in* $[0,\infty)$*, one has*

$$0 \leq K_2(t,x,y) \leq a \exp\left[b(t^{-\alpha} \vee 1)\right] \Phi_2(x)\Phi_2(y). \tag{2.1.6}$$

Proof. Let L_2 be the positive self-adjoint operator in $\mathrm{L}^2(\mathbb{R}^n, \Phi_2^2 dx)$ which is unitarily equivalent to H_2 according to (1.1.13). Under the above assumptions we can apply [36, Theorem 4.7.1], whence it follows that $\exp[-tA_2]$ has a heat kernel $K_{\Phi,2}(t,x,y)$ satisfying

$$0 \leq K_{\Phi,2}(t,x,y) \leq a \exp\left[b(t^{-\alpha} \vee 1)\right] \tag{2.1.7}$$

for some positive a, b, α. Then (2.1.6) follows as in [36, Lemma (4.2.2)], since $K_2(t,x,y) = K_{\Phi,2}(t,x,y)\Phi_2(x)\Phi_2(y)$.

□

Lemma 2.3. . *For all* f *in* L^2 *and* t *in* $(0,\infty)$, $\exp[-tH_2]f$ *is in* $\mathrm{Dom}\, V(s)$ *for all* s.
Proof. It suffices to prove the lemma for $f \geq 0$ in L^2. Then (2.1.6) implies that

$$0 \leq \left(\exp[-tH_2]f\right)(x) \leq a \exp\left[b(t^{-\alpha} \vee 1)\right] \Phi_2(x)\left(\Phi_2, f\right), \tag{2.1.8}$$

and the claim follows upon taking into account (2.1.1).

□

Next we recall some basic facts about the Feynman-Kac formula, having Reed–Simon [56] and Simon [35] as general references. Let Ω be the set of Brownian paths $w : [0,\infty) \to \mathbb{R}^n$, with $w(s) = x + \sqrt{2}w_s$, and let $d\mu$ be the product measure on Ω of Lebesgue measure

dx on the starting point $x \in \mathbb{R}^n$ with the probability measure dP which makes $\{w_s : s \in [0,\infty)\}$ an n-dimensional standard Brownian motion starting at x.

Let V be a real-valued function defined on $\mathbb{R}^n$ (continuous and bounded, for simplicity), and let H be the self-adjoint closure in L^2 of the operator $H = -\Delta + V$. Then, for all f,g in L^2 and for all t in $(0,\infty)$, we have

$$\begin{aligned}(\exp[-tH]f,g) &= (f,\exp[-tH]g) = \\ &= \int_\Omega \overline{f(w(0))} \exp\left[-\int_0^t V(w(u))du\right] g(w(t))d\mu(w),\end{aligned} \tag{2.1.9}$$

which may be also written in terms of a heat kernel $K(t,x,y) \geq 0$ as

$$(\exp[-tH]f,g) = (f,\exp[-tH]g) = \int dxdy\overline{f(x)}K(t,x,y)g(y).$$

Note that the factor of $\sqrt{2}$ in front of w_s in the definition of $w(s)$ serves to obtain $H = -\Delta + V$ instead of $H = -\Delta/2 + V$. In particular, if $V \equiv 0$, so that $H = H_0 = -\Delta$, we have

$$(\exp[-tH_0]f,g) = (f,\exp[-tH_0]g) = \int dxdy\overline{f(x)}K_0(t,x,y)g(y),$$

$$K_0(t,x,y) = (4\pi t)^{-n/2}\exp[-\frac{1}{4t}|x-y|^2]. \tag{2.1.10}$$

In fact, (2.1.9) holds for a very general class of potentials (see [35]), and for example it suffices that $V_+ \in L^1_{\mathrm{loc}}(\mathbb{R}^n \setminus G)$ for a measure-zero closed set G and that V_- is relatively $-\Delta$-form bounded with relative bound smaller than one.

Lemma 2.4. *There exists a (not necessarily symmetric) heat kernel $\tilde{K}(t,x,y)$ such that, for all f,g in L^2,*

$$\int_\Omega \overline{f(w(0))} \exp\left[-\int_0^t V(u,w(u))du\right] g(w(t))d\mu(w) = \int \overline{f(x)}\tilde{K}(t,x,y)g(y)dxdy, \tag{2.1.11}$$

and such that, for all x,y in $\mathbb{R}^n$ and s in $(0,\infty)$, one has

$$0 \leq \tilde{K}(s,x,y) \leq \exp[\tau_1(s)]K_0(s,x,y), \tag{2.1.12}$$

$$0 \leq \tilde{K}(s,x,y) \leq \exp[\tau_2(s)]K_2(\tau_3(s),x,y), \tag{2.1.13}$$

where K_0 is given by (2.1.10), K_2 is as in Lemma 2.2, and where

$$\tau_1(s) = \frac{1}{2}(c+d)\int_0^s \beta(u)du = \frac{1}{2}(c+d)t(s), \tag{2.1.14}$$

$$\tau_2(s) = \frac{1}{2}(c+d)\int_0^s \beta'(u)du = \frac{1}{2}(c+d)\beta(s), \tag{2.1.15}$$

$$\tau_3(s) = \frac{1}{2}\int_0^s [\beta(u) - \beta'(u)]du = \frac{1}{2}[t(s) - \beta(s)]. \tag{2.1.16}$$

Proof. For positive real r, let

$$V_r(s,x) = V(s,x) \wedge r. \tag{2.1.17}$$

Then $x \mapsto V_r(s,x)$ is a continuous bounded function, and the same is true for $s \mapsto V_r(s,w_s)$; so there is no problem in generalizing (2.1.9) to define

$$\int_\Omega \overline{f(w(0))} \exp\left[-\int_0^t V_r(u,w(u))du\right] g(w(t))d\mu(w). \tag{2.1.19}$$

Since also $V_r(s)$ satisfies the bound (2.1.5), it follows from consideration of positive f,g in L^2 converging in the sense of distributions to the Dirac delta at x and at y respectively that there exists a (not necessarily symmetric) heat kernel $\tilde{K}_r(t,x,y)$ such that

$$(2.1.19) = \int dxdy\overline{f(x)}\tilde{K}_r(t,x,y)g(y), \tag{2.1.20}$$

$$0 \leq \tilde{K}_r(s,x,y) \leq \exp[\tau_1(s)]K_0(s,x,y). \tag{2.1.21}$$

Now let $r \to \infty$. Then $\exp\left[-\int_0^t V_r(u,w(u))du\right]$ is a nonincreasing function of r for each w. For positive f,g in L^2, (1.2.18) is a monotonic nonincreasing function of r, and the same is true for $\tilde{K}_r(t,x,y)$, pointwise in t,x,y. Then (2.1.11) and (2.1.12) follow, with $\tilde{K}(t,x,y) = \lim_{r\to\infty} \tilde{K}_r(t,x,y)$. In order to obtain (2.1.13), we use the bound (2.1.4) on $V(s)$.

□

In the following, for the sake of simplicity, we shall assume that the initial condition f at time $t=0$ is obtained by starting from an arbitrary $f_0 \in L^2$ at time $t = -\varepsilon\beta(0)$, and letting the process x_t evolve from $t = -\varepsilon\beta(0)$ to $t=0$ with constant "temperature" $T(0) = 1/\beta(0)$. Then $f = \exp\left[-\varepsilon H_{\beta(0)}\right] f_0$. Since $\beta(0) \geq 2$, we have $V_{\beta(0)} \geq V_2$. It follows that for positive f_0 we have

$$0 \leq f(y) = \int f_0(x)K_{\beta(0)}(\varepsilon,x,y)dx \leq$$

$$\leq \int f_0(x)K_2(\varepsilon,x,y)dx \leq a\exp\left[b(\varepsilon^{-\alpha} \vee 1)\right](f_0,\Phi_2)\Phi_2(y), \tag{2.1.22}$$

which is in Dom $V(s)$ for all s, where by Dom $V(s)$ we mean the L^2–domain of the associated multiplication operator in L^2. Since Φ_2 is a continuous function, a similar bound holds also when f_0 is replaced by a multiple of the Dirac delta.

Theorem 2.5. *For $f = \exp\left[-\varepsilon H_{\beta(0)}\right] f_0$, f_0 in L^2 and t in $[0,\infty)$, let*

$$(P(t)f)(y) = \int f(x)\tilde{K}(t,x,y)dx. \tag{2.1.23}$$

Then $P(t)f$ is in $\mathrm{Dom}\, V(s) \subseteq L^2$ *for all s,t in $[0,\infty)$, and the following du Hamel formula holds:*

$$P(t)f = \exp\left[-tH_0\right] f - \int_0^t \exp\left[-(t-s)H_0\right] V(s)P(s)f ds. \tag{2.1.24}$$

The conclusion remains true if $f_0 \in L^2$ is replaced by a multiple of the Dirac delta measure at some point x.

Proof. We adapt the proof in [35]. First we work with $V_r(s)$. We have

$$\frac{d}{ds}\exp\left[-\int_0^s V_r(u,w(u))du\right] = \exp\left[-\int_0^s V_r(u,w(u))du\right]\left[-V_r(s,w(s))\right],$$

and integrating between 0 and t we obtain

$$\exp\left[-\int_0^t V_r(u,w(u))du\right] - 1 = \int_0^t \exp\left[-\int_0^s V_r(u,w(u))du\right]\left[-V_r(s,w(s))\right]ds.$$

Take $f,g \in \mathrm{L}^2$, multiply both sides of the above equality by $\overline{f(w(0))}g(w(t))$, and integrate on Ω with respect to $d\mu(w)$. By independence of the increments of Brownian motion in disjoint time intervals we obtain, using the definition (2.1.20) of $\tilde{K}$,

$$\int dxdy\overline{f(x)}\tilde{K}_r(t,x,y)g(y) - (f,\exp[-tH_0]g) =$$

$$= \int_0^t ds\left[\int dxdy\overline{f(x)}\tilde{K}_r(t,x,y)V_r(s,y)\left(\exp[-tH_0]g\right)(y)\right]. \tag{2.1.25}$$

Let

$$\left[P_r(t)f\right](y) = \int f(x)\tilde{K}_r(t,x,y)dx; \tag{2.1.26}$$

(2.1.26) tells us that

$$P_r(t)f = \exp[-tH_0]f - \int_0^t \left[-(t-s)H_0\right] V_r(s)P_r(s)f ds \tag{2.1.27}$$

(the passage from weak to strong form being allowed as all operators involved are bounded). There remains to prove that one may let $r \to \infty$ in (2.1.27) and obtain (2.1.24). Note that for f of the form $\exp\left[-\varepsilon H_{\beta(0)}\right] f_0$, we have

$$\left[P(t)f\right](y) = \int f(x)\tilde{K}(t,x,y)dx = \int f_0(x)\hat{K}(t+\varepsilon,x,y)dx, \tag{2.1.28}$$

where $\hat{K}(t,x,y)$ is obtained by replacing $V(s,w(s))$ in (2.1.11) by $V(s-\varepsilon,w(s))$, with $V(t,x) = V(0,x)$ for $t<0$. An application of Lemmas 2.2 and 2.4 shows that

$$\left[P(t)f\right](y) \le a\exp\left[b(\varepsilon^{-\alpha}\vee 1)\right](f_0,\Phi_2)\Phi_2(y), \quad y \in \mathbb{R}^n,\ t \in [0,\infty), \tag{2.1.29}$$

implying that $P(t)f$ and $V(s)P(t)f$ are in L^2 with L^2–norm bounded uniformly in s, t. By dominated convergence, we can take the limit as $r \to \infty$ in (2.1.25), thus obtaining (2.1.24) in weak form. Upon noting that $V(s)P(s)$ is bounded by the closed graph theorem, the passage from weak to strong form is allowed.

□

Finally, in order to show that the propagator (2.1.23) is a solution to the original differential equation (1.1.20) it suffices to recall the following classical result (cf. Lion and Magenes [37]); suppose that A is the generator of an analytic semigroup of some angle θ (which is of course the case for H_0); then the convolution

$$g(t) = \int_0^t e^{(t-s)A} f(s)ds, \qquad t \in (0,T), \tag{2.1.30}$$

with $f \in L^2((0,T),\mathcal{H})$, where $\mathcal{H}$ is a Hilbert space, defines an element g of $L^2((0,T), D(A))$, where $D(A)$ is the domain of A endowed with the graph norm. This implies in particular that the right-hand side of (2.1.30) is in $D(A)$ for almost all t, and this is exactly what we need to be allowed to take derivatives in du Hamel's formula corresponding to the potential $V(s)$ at least for almost all t. Since this derivative is used to derive a differential inequality which is afterwards integrated, this is all that we need to complete the prove of the existence of a solution to the evolution equation (1.1.20) for a class of time–dependent potentials arising from Langevin diffusions.

Corollary 2.6. *The solution $f_t = P(t)f$ belongs to L^p for all $t \in [0,\infty)$ and for all $p \in [1,\infty]$*

The corollary follows from the choice of the initial condition, which corresponds to starting the process at time $-\varepsilon\beta(0)$ and running the process with constant temperature until $t = 0$, and from the fact that, in analogy with (2.1.13), (2.1.6), there exists a positive continuous function $c(s)$ such that

$$K(s,x,y) \leq c(s)\Phi_{\beta(s)}(x)\Phi_{\beta(s)}(y). \tag{2.1.31}$$

A remark is in order: we used in section 1.4 the Radon–Nikodym derivatives $\nu_t = d\mu_t/d\mu_{1/T(t)}$, μ_t being the distribution of the process at time t. It is clear that $\nu_{t(s)} = f_s\Phi^{-1}_{\beta(s)}$. An immediate consequence of Theorem 2.4 is the following

Corollary 2.7. *The Radon–Nikodym derivatives ν_t belong to $L^\infty(d\mu_{1/T(t)})$ for all $t \geq 0$.*

This proves that the calculations of section 1.4 make sense, since all the quantities involved are finite.

2.2. Solution via Lions's theorem

In this section we prove that, under Assumption 1.6, Lions's theorem for parabolic equations [37] can be used in order to prove existence and unicity of a weak solution to the equation at hand. Without loss of generality we also assume that $\beta(0) > 2$ strictly. We start with some pointwise estimates on the potential V_β.

Lemma 2.8. *There exists $A(\beta,\beta_0) \in \mathbb{R}$ such that*

$$V_\beta - V_{\beta_0} \le K(\beta,\beta_0)V_{\beta_0} + A(\beta,\beta_0), \tag{2.2.1}$$

where $\beta \in [\beta_0, 1+[1+2\beta_0(\beta_0-2)]^{1/2})$ and $K(\beta,\beta_0) \in (\beta^2/\beta_0^2 - 1, 1)$ is given by

$$K(\beta,\beta_0) = \frac{1}{\beta_0^2 - 2\beta_0}(\beta^2 - \beta_0^2 - 2\beta + 2\beta_0). \tag{2.2.2}$$

Proof. Note that Assumption 1.6 implies that

$$\Delta U \le |\nabla U|^2 + b, \tag{2.2.3}$$

holds for a suitable constant b. Let $\beta \in [\beta_0, \sqrt{2}\beta_0)$. Proving (2.2.1) is equivalent to proving an inequality of the form

$$\Delta U \le \frac{1}{2}\frac{\beta_0^2 K(\beta,\beta_0) - \beta^2 + \beta_0^2}{\beta_0 K(\beta,\beta_0) - \beta + \beta_0}|\nabla U|^2 + \frac{2A(\beta,\beta_0)}{\beta_0 K(\beta,\beta_0) - \beta + \beta_0}. \tag{2.2.4}$$

But we know that (2.2.4) holds with the coefficient of $|\nabla U|^2$ equal to one; this implies that (2.2.1) holds with $K(\beta,\beta_0)$ given by (2.2.2). In order that the relative bound be strictly smaller then one, it is necessary that $\beta < 1 + [1+2\beta_0(\beta_0-2)]^{1/2}$.

$\square$

Fix now $T > 0$; and choose a constant $c = c(T)$ such that $H(s)+c > k > 0$ for any $s \in [0,T]$. For the sake of notational simplicity, we shall still denote $H(s)+c$ by $H(s)$; this simple trick makes $H(s)$ boundedly invertible for all $s \in [0,T]$. Observe that a pointwise bound for functions is equivalent to a quadratic form bound for the corresponding multiplication operators, and let Dom $[H(t)]^{1/2}$ denote the form domain of $H(t)$, that is the L^2–domain of the positive square root of $H(t)$.

Lemma 2.9. *The form domain of $H(s)$ does not depend on $s \in [0,T]$.*
Proof. Let $\beta_{n+1} = 1 + [1+2\beta_n(\beta_n - 2)]^{1/2}$, $\beta_0 = \beta(0) > 2$. It follows that β_n is strictly increasing, with $\lim_{n\to\infty}\beta_n = \infty$. By using a finite number of times a standard relative boundedness argument and recalling that $-\Delta$ is positive, it follows that $Q(H_\beta) = Q(H_{\beta_0})$ for all $\beta > \beta_0$. Noting that $U \le \varepsilon V_2 + b$, $V_\beta \ge \beta V_2/2$, $\beta \ge \beta'$, we find that

$$\frac{1}{2}\beta' U \le \delta V_\beta + k, \tag{2.2.5}$$

for a suitable constant k, where $\delta = \varepsilon\beta'/\beta$ is strictly smaller than one.

□

We define V as the Hilbert space $\mathrm{Dom}[H(t)]^{1/2}$ endowed with the graph norm

$$|f| = \left[||f||^2 + ||H(0)^{1/2} f||^2\right]^{\frac{1}{2}}, \quad f \in V. \tag{2.2.6}$$

The immersion of V in $\mathcal{H} = \mathrm{L}^2$ is continuous since $||f|| \leq |f|$ by definition for any $f \in V$, and is dense since $C_0^\infty \subset V$. Let $a(t)$, $t \in [0, T]$ be the family of sesquilinear forms defined on V by

$$a(t, u, v) = (H(t)^{1/2} u, H(t)^{1/2} v), \quad u, v \in V \tag{2.2.7}$$

Lemma 2.10. *The following assertions hold:*
i) the function $t \mapsto a(t, u, v)$ defined for $t \in [0, T]$ is measurable for all $u, v \in V$.
ii) There exists $M > 0$ such that

$$|a(t, u, v)| \leq M|u||v| \tag{2.2.8}$$

for any $u, v \in V, t \in [0.T]$.
iii) There exists $K > 0$ such that

$$a(t, u, u) \geq K|u| - ||u|| \tag{2.2.9}$$

for any $u \in V, t \in [0, T]$.
Proof. The dependence on t of the function under consideration is continuous for any u, v in a suitable form core such as C_0^∞. Since the pointwise limit of continuous functions is measurable, *i)* follows. Observe now that

$$d_{t,s} := ||H(t)^{\frac{1}{2}} H(s)^{-\frac{1}{2}}|| \tag{2.2.10}$$

is finite for any $t, s \in [0, T]$ by the closed graph theorem.
Since $d_{t,0}^2 = \sup_{||f||=1} a(t, H(0)^{-1/2} f, H(0)^{-1/2} f)$, applying the relative boundedness argument of the previous Lemma a finite number of times, and using the fact that $H(0)^{-1/2}$ is bounded, one sees that there exists $d > 0$ such that $d_{t,0} < d$ for any $t \in [0, T]$. Hence there exists $M > 0$ such that for any $t \in [0, T]$

$$|a(t, u, v)| \leq ||H(t)^{\frac{1}{2}} u|| \; ||H(t)^{\frac{1}{2}} v|| \leq$$

$$\leq d_{t,0} ||H(0)^{\frac{1}{2}} u|| \; ||H(0)^{\frac{1}{2}} v|| \leq (d_{t,0}^2 \vee 1)|u||v| \leq M|u||v|, \tag{2.2.11}$$

thus proving *ii)*.

In addition, reversing the roles of 0 and t shows that there exists $K > 0$ such that $d_{0,t} < K$ for any $t \in [0, T]$, and that

$$a(t, u, u) = \left(||H(t)^{\frac{1}{2}} u||^2 + ||u||^2\right) - ||u||^2 \geq \left(d_{0,t}^{-1} ||H(0)^{\frac{1}{2}} u||^2 + ||u||^2\right) - ||u||^2 \geq$$

$$\geq (1 \wedge d_{0,t}^{-1})|u|^2 - ||u||^2 \geq (1 \wedge K^{-1})|u|^2 - ||u||^2. \tag{2.2.12}$$

□

Theorem 2.11. *Denote by $\langle \cdot, \cdot \rangle$ the pairing between V' and V. Then, for every $u_0 \in \mathcal{H}$ there exists one and only one $u \in L^2([0,T]; V) \cap C([0,T]; \mathcal{H})$ with $du/dt \in L^2([0,T]; V')$ and such that*

$$\langle \frac{du}{dt}, v \rangle = -(H(t)^{\frac{1}{2}} u, H(t)^{\frac{1}{2}} v), \quad u(0) = u_0 \tag{2.1.13}$$

for almost all $t \in [0,T]$ and for all $v \in V$.
Proof. A straightforward application of Lions's theorem [37] and of Lemma 2.10.

□

2.3 Solution via Acquistapace–Terreni theory.

Here we use the general theory of abstract linear evolution equation developed by Acquistapace and Terreni (cf. [38, 39] and references quoted therein) to prove, under suitable assumptions, existence and uniqueness results of the above kind. Here, however, the results obtained hold for general initial condition and, moreover, the solution $u(t)$ solves (1.1.20) for all t and not only for almost all t. Moreover, the class of functions U for which this approach is efficient is rather large, since it includes for example all functions which grow like a power of $|x|$ in a neighbourhood of infinity. As in the previous chapter, we fix $T > 0$ and assume that $H(t) > \varepsilon > 0$ for all $t \in [0, T]$.
With the notation of [38] we give the following

Definition 2.12. A function $u \in C([0,T], \mathrm{L}^2)$ is said to be a strict (respectively classical) solution of (1.1.10) if $u \in C^1([0,T], \mathrm{L}^2) \cap C([0,T], D[H(t)])$ (respectively $u \in C^1((0,T], \mathrm{L}^2)$ $\cap C((0,T], D[H(t)])$).

The additional assumption of this section is as follows:

Assumption 2.13 (continuity). For each $t > s \geq 0$ the operator $(V(t) - V(s))H(t)^{-1}$ is bounded, and moreover there exists $C > 0$, $\gamma \in (0, 2)$ such that

$$||(V(t) - V(s))H(s)^{-1}|| \leq c(t-s)^{\gamma}. \tag{2.3.1}$$

Using some of the results of [38], in particular the theorems contained in section 6 of that paper (see also [39] and references quoted therein), we have the following.

Proposition 2.14. *Let $f_0 \in D(H(0))$ (respectively $f_0 \in \mathrm{L}^2$). Then there exist a unique strict (respectively classical) solution to* (1.1.20) *with initial condition f_0.*
Proof. We shall prove that, for λ in a sector of any angle $\theta \in (0, \pi)$, the quantity

$$||H(t)(H(t) - \lambda)^{-1}[H(t)^{-1} - H(s)^{-1}]|| \tag{2.3.2}$$

can be estimated in terms of the left–hand side of (2.3.1), provided that it makes sense. Let $g \in \mathrm{L}^2$ be given; we have

$$H(t)(H(t) - \lambda)^{-1}[H(t)^{-1} - H(s)^{-1}]g =$$

$$= (H(t) - \lambda)^{-1}g - H(s)^{-1}g - \lambda(H(t) - \lambda)^{-1}H(s)^{-1}g =$$
$$= [(H(t) - \lambda)^{-1}(H(s) - \lambda) - 1]H(s)^{-1}g =$$
$$= (H(t) - \lambda)^{-1}(H(s) - \lambda)H(s)^{-1}g - H(s)^{-1}g. \tag{2.3.3}$$

Define the following quantities:

$$u = (H(t) - \lambda)^{-1}(H(s) - \lambda)H(s)^{-1}g, \tag{2.3.4}$$

$$v = H(s)^{-1} g. \tag{2.3.5}$$

Then $u \in D(H(t))$, $v \in D(H(s))$ and

$$(H(t) - \lambda)u = (H(s) - \lambda)v = g - \lambda v, \tag{2.3.6}$$

$$H(s)v = g \tag{2.3.7}$$

Writing $H(t)$ as $H(t) = \triangle + V(t)$, and subtracting (2.3.7) from (2.3.6), we have (the Laplacian being meant in distributional sense),

$$\triangle(u - v) - \lambda(u - v) + V(t)(u - v) = [(V(s) - V(t))v]; \tag{2.3.8}$$

note that each single summand in the left–hand side of (2.3.8) need not be in L^2. Eq. (2.3.8) can be rewritten as

$$u - v = (H(t) - \lambda)^{-1}(V(s) - V(t))v, \tag{2.3.9}$$

where the right–hand side makes sense by Assumption 2.13. Since for all $t \in [0, T]$, for all λ in a sector of any angle $\theta \in (0, \pi)$, and for a suitable constant $c = c(T)$ an estimate of the form

$$||(H(t) - \lambda)^{-1}|| \leq \frac{c}{|\lambda|} \tag{2.3.10}$$

holds, we have by (2.3.1), (2.3.3), (2.3.9) and (2.3.10) that

$$||H(t)(H(t) - \lambda)^{-1}[H(t)^{-1} - H(s)^{-1}]|| \leq \frac{c}{|\lambda|}|t - s|^{\gamma}. \tag{2.3.11}$$

The inequality (2.3.11) is precisely of the form ([38], Hypothesis 2), whence the thesis follows as in ([38], Section 6).

□

Corollary 2.15. *Suppose that*

$$\mathrm{Dom}\,(H(t)) = H^2(\mathbb{R}^n) \cap D(V(t)), \tag{2.3.12}$$

where $\mathrm{Dom}\,(V(t))$ *denotes the domain of* $V(t)$ *as a multiplication operator in* L^2. *Then Assumption* 2.13 *is satisfied with* $\gamma = 1$, *so that Theorem* 2.14 *holds.*
Proof. We have shown in the previous section that

$$V_t - V_s \leq K(t, s)V_s + A(t, s), \tag{2.3.13}$$

where

$$K(t, s) = \frac{1}{\beta(s)^2 - 2\beta(s)}(\beta(t)^2 - \beta(s)^2 - 2\beta(t) + 2\beta(s)), \tag{2.3.14}$$

$$A(t, s) = \text{const.} \times (\beta(s)K(t, s) - \beta(t) + \beta(s)). \tag{2.3.15}$$

If $\beta(t) = \text{const.} \times \log(t_0 + t)$, it follows that, for $t > s$ $|K(t,s)| \leq \text{const.} \times (t-s)$, $|A(t,s)| \leq \text{const.} \times (t-s)$. Moreover, it is simple to check that

$$|\beta'(t) - \beta'(s)|U \leq \text{const.} \times \beta'(s)(t-s)U$$

and that

$$|\beta'(s)\langle U\rangle_s - \beta'(t)\langle U\rangle_t| \leq \text{const.}(t-s).$$

This implies that

$$|V(t) - V(s)| \leq c_1(t-s)(V(s)+1). \tag{2.3.16}$$

Since $H(s)^{-1}g$ belongs to the domain of $H(s)$, and hence, by assumption, to the domain of $V(s)$, the claim follows from the closed graph theorem.

□

Remark 2.16. Assumption (2.3.12) holds for example if the so–called condition V of Kato (see [57, 58]) is the satisfied, that is if $(V(t)+c)^{-1/2}$ (for sufficiently large c) is Lipschitz in a neighbourhood of infinity with Lipschitz constant strictly smaller than one. If $U(x) = p(x)$, $U(x) = \exp[p(x)]$, $U(x) = \exp[\exp[p(x)]], \ldots$ in $B(0,r)^c$ for some $r > 0$ and for some polynomial $p(x)$ with $\lim_{|x|\to\infty} p(x) = +\infty$, the latter condition holds.

Chapter 3:

TIME–DEPENDENT ENERGY FUNCTIONS

In this section we discuss the case in which the "energy function" U is allowed to depend on time. Specifically, we consider the following stochastic differential equation in $\mathbb{R}^n$:

$$dx_t = -\nabla U(x_t, t)dt + \sqrt{2T(t)}dw_t, \tag{3.0.1}$$

where $T : [0, \infty) \to [0, \infty)$ is a monotonically non–increasing function of class C^1, and $U: [0, \infty) \times \mathbb{R}^n \to [0, \infty)$ is a function of class $C^{1,2}([0, \infty) \times \mathbb{R}^n)$.

The techniques used here are essentially similar to those of the sections 1.1, 1.2. In fact for the sake of simplicity we have not pursued the goal of proving L^p–estimates via logarithmic Sobolev inequalities as in section 1.3, 1.4. However, since the problem seems to be new also in the (computationally simpler) case of diffusions on compact manifolds, we include a proof of our result also in this setting. This is done is section 3.1, whereas sections 3.2 and 3.3 are devoted to the case of $\mathbb{R}^n$, and specifically to the reformulation of the problem in L^2 and to the L^2-estimates.

3.1 Compact manifolds

Let $\mathbb{M}$ be a compact connected finite dimensional Riemannian manifold, with Riemannian metric g. Denote by ∇ the covariant gradient, and by ν the Riemannian measure associated with g.

To introduce the diffusion process which we shall be concerned with, we need some more notation. Let $\Omega = C([0, \infty), \mathbb{M})$ be the *path space*, that is space of continuous functions from $[0, \infty)$ with values in $\mathbb{M}$. Then Ω is a Polish space when endowed with the topology of uniform convergence on finite intervals.

Define the coordinate process x_t on Ω by $x_t(\omega) = \omega(t)$, $\omega \in \Omega$. Let also $\mathcal{F}_t$ be the filtration generated by X_t. Then $\mathcal{F} = \cup_{t \geq 0} \mathcal{F}_t$ generates the Borel field of Ω.

Let $\beta : [0, \infty) \to [0, \infty)$ be of class C^∞ and monotonically non–decreasing, with $\beta(0) = 0$, and $U: \mathbb{M} \times \mathbb{R}^n \to [0, \infty)$ be a function of class $C^\infty(\mathbb{M})$.

Finally, define a family of operators $L_t : C^\infty(\mathbb{M}) \to C^\infty(\mathbb{M})$ by

$$L_t f = e^{\beta(t)U(t)} \nabla \cdot \left(e^{-\beta(t)U(t)} \nabla f \right); \tag{3.1.1}$$

here we have denoted by $U(t)$ the function $U(t, \cdot)$ on $\mathbb{R}^n$. It is known that, for each fixed t, L_t is essentially self–adjoint in $\mathcal{L}^2(\mathbb{M}, \mu_t)$ where

$$\mu_t(dx) = Z_t^{-1} e^{-\beta(t)U(t,x)} \nu(dx), \tag{3.1.2}$$

$$Z_t = \int e^{-\beta(t)U(t,x)} \nu(dx). \tag{3.1.3}$$

Furthermore, L_t is a non-positive operator, which has $\lambda = 0$ as simple, non–degenerate, highest eigenvector, and there is a gap $-\gamma_t < 0$ between $\lambda = 0$ and the rest of its spectrum. By general theorems about the so–called martingale approach to diffusions [59, 60], we have the following

Proposition 3.1. *For each $(s,x) \in [0,\infty) \times \mathbb{M}$, there exists a unique probability measure $P_{s,x}$ on $(\Omega, \mathcal{F})$ such that*

$$\left(f(X_t) - f(x) - \int_s^t L_u f(X_u) du,\ \mathcal{F}_t\ ,\ P_{s,x} \right) \tag{3.1.4}$$

is a mean zero martingale for all $f \in C^\infty(\mathbb{M})$. Moreover, the family $P_{s,x}$ is Feller continuous and strongly Markov.

Similar properties hold also if the process starts from a measurable initial probability distribution.

Without entering into details, we only remark that, with the terminology of the previous chapters, the process described in the above Proposition corresponds to a Langevin algorithm of the kind considered in chapter 1 with starting point x at time s, to a time–dependent "energy function" $U(t)$ and to a "cooling schedule" $\beta(t)$. It should be noted the the time scaling of section 1.1 has already been performed.

Let $P(x,t,dy)$ be the distribution of X_t under $P_{s,x}$. Then $P(x,t,\cdot)$ is absolutely continuous with respect to μ_t for all $t > 0$.

In the sequel, following the lines of chapter 1, our aim is to control the "size" of the Radon–Nikodym derivative $\nu_t(\cdot)$ $(\in (0,\infty))$, in the form of estimates on its $\mathrm{L}^2(\mu_t)$–norm. In fact, let

$$T_t f(x) = \mathbf{E}_x[f(X_t)], \quad f \in C^\infty(\mathbb{M}), \tag{3.1.5}$$

where $\mathbf{E}_x$ denotes expectation with respect to $P_{0,x}$. It follows that

$$\frac{d}{dt}(T_t f) = T_t(L_t f). \tag{3.1.6}$$

Consider now the unitary operator between $\mathrm{L}^2(\mu_t)$ and $\mathrm{L}^2(\nu)$ defined by the correspondence $f \mapsto f\Phi_t$ for all $f \in \mathrm{L}^2(\mu_t)$, Φ_t being defined as

$$\Phi_t = \frac{\exp[-\beta(t)U(t)/2]}{Z_t^{1/2}}. \tag{3.1.7}$$

Under this mapping L_t is unitarily equivalent to H_t, where

$$H_t = \Delta - V_t = \Delta - \left(\frac{\beta^2}{4} |\nabla U(t)|^2 - \frac{\beta}{2} \Delta U(t) \right), \tag{3.1.8}$$

$|\cdot|$ denotes the Riemannian norm of a vector field, and Δ is the Laplace–Beltrami operator on $\mathbb{M}$. Here, and often in the sequel, we avoid to write explicitly the time dependence of β (and sometimes of U) for the sake of notational simplicity.

Fix $x \in \mathbb{M}$, and let $p_t(y) = p(x,t,y)$, where $P(x,t,dy) = p(x,t,y)dy$ for $t > 0$, so that

$$\mathrm{Prob}[x_t \in B] = \int p_t(y) I_B(y) \nu(dy) \tag{3.1.9}$$

for any Borel set B, I_B being the indicator function of B. Finally, define also a family $\{f_t\}_{t>0}$ of continuous (and hence square–integrable) functions by

$$p_t(x) = f_t(x)\Phi_t(x) \tag{3.1.10}$$

Define, for all $g \in \mathrm{L}^1(\mu_t)$, $\langle g \rangle_t$ as

$$\langle g \rangle_t = \int g(x)\mu_t(dx). \tag{3.1.11}$$

Assume that f_t is defined and continuous also for $t = 0$. Then we have the following

Lemma 3.2. *f_t satisfies the following time–dependent heat equation:*

$$\frac{d}{dt} f_t = -H(t) f_t = -[-\triangle + V(t)] f_t, \tag{3.1.12}$$

$V(t)$ being defined as the operator of multiplication by the function

$$V(t,x) = V_t(x) - \frac{1}{2}\beta'(t)[U(t,x) - \langle U(t,\cdot)\rangle_t] - \frac{1}{2}\beta(t)[U'(t,x) - \langle U'(t,\cdot)\rangle_t]; \tag{3.1.13}$$

here a prime stands for the derivative with respect to t.
Proof. The thesis follows as in Lemma 1.1, by noting that

$$\frac{d}{dt}\log \Phi_t = -\frac{1}{2}\beta'(t)\,[U(t,x) - \langle U(t,\cdot)\rangle_t] - \frac{\beta}{2}\,[U'(t,x) - \langle U'(t,\cdot)\rangle_t]\,.$$

□

In the sequel we suppose for the sake of simplicity that the initial condition f_0 is obtained by letting an arbitrary $g \in \mathrm{L}^2(\nu)$ (or the Dirac measure concentrated at $x_0 \in \mathbb{M}$) evolve for times $t \in (-\varepsilon, 0)$, $\varepsilon > 0$, under the equation $(d/dt)g_t = -H_{t=0}g_t$. This corresponds to consider the time-evolved of g with temperature and energy function which do not depend on time in $t \in (-\varepsilon, 0)$. In the sequel f_t will denote the unique solution to (3.1.13) with initial condition of the above mentioned form; see the next section for a more general discussion.

The Radon–Nikodym derivative ν_t is given by $\nu_t = f_t\Phi_t^{-1}$, so that

$$\frac{d}{dt}\nu_t = [L_t + \beta'(U - \langle U\rangle_t) + \beta(U' - \langle U'\rangle_t)]\nu_t. \tag{3.1.14}$$

Set now

$$N(t) = ||f_t - \Phi_t||_2^2; \tag{3.1.15}$$

then we have

$$|\text{Prob}[X_t \in A] - \mu_t(A)| = |(f_t - \Phi_t, I_A \Phi_t)| \leq$$
$$\leq \mu_t(A)^{1/2} N(t)^{1/2}; \tag{3.1.16}$$

here $||\cdot||_2$ and $(\cdot\ ,\ \cdot)$ denote the norm and the scalar product in $L^2(\nu)$. Eq. (3.1.16) shows that asymptotic indistinguishability of the distribution of the process and of the instantaneous equilibrium distribution is implied by the existence of a uniform bound on $N(t)$. We want to find sufficient conditions for this to hold.

Lemma 3.3. *Let*

$$Q_t(f) = -(f, H_t f) \tag{3.1.17}$$

denote the Dirichlet form associated with the non–negative operator $-H_t$. Then

$$\frac{d}{dt}N(t) = -2Q_t(f_t) + \beta'(f_t, [U - \langle U\rangle_t]f_t) + \beta(f_t, [U' - \langle U'\rangle_t]f_t) \tag{3.1.18}$$

Proof. It suffices to observe that

$$||f_t - \Phi_t||_2^2 = ||f_t||^2 - 1,$$

and to recall (3.1.13).

Let now f_+ (resp. f_-) denote the positive (resp. negative) part of a real function f. Set

$$||U(t)||_\infty = a(t), \quad ||U'_+(t)||_\infty = b(t), \quad \langle U'_-\rangle_t = c(t). \tag{3.1.19}$$

Observe that the non–negative self–adjoint operator $-H_t$ has, for all $t \in \mathbb{R}$, a gap γ_t between $\lambda = 0$ and the rest of its spectrum. It is also known (cf. [12]) that there exists a positive $m = m(t)$ such that

$$\gamma_t \geq \gamma \exp[-\beta(t)m(t)], \tag{3.1.20}$$

for a suitable positive constant γ which depends only on $\mathbb{M}$. With these notations at hand, we have the following

Lemma 3.4. *$N(t)$ satisfies the following differential inequality:*

$$\frac{d}{dt}N(t) \leq [-2\gamma_t + \beta' a(t) + \beta(b(t) + c(t))]N(t) + \beta' a(t) + \beta(b(t) + c(t)). \tag{3.1.21}$$

Proof. We note that

$$-Q_t(f_t) = (f_t, H_t f_t) = (f_t - \Phi_t, H_t(f_t - \Phi_t)) = -Q_t(f_t - \Phi_t).$$

Since $(f_t - \Phi_t, \Phi_t) = 0$, the spectral gap condition implies that

$$Q_t(f_t - \Phi_t) \geq \gamma_t N(t),$$

and hence we have the claim, upon recalling that $||f_t||^2 = 1 + N(t)$, and writing

$$U = U_+ - U_-, \qquad U' = U'_+ - U'_-.$$

□

Lemmas 3.4 and 1.18 show that in order to prove that $N(t)$ tends to zero for $t \to \infty$, it suffices to give bounds on the coefficients of (3.1.21) in the form (1.4.20), (1.4.21). To this end, let us make the following

Assumption 3.5.
i) There exist $m > 0$ such that $m(t) < m$, $m(t)$ being as in (3.1.20).
ii) The "cooling schedule" $\beta(t)$ is given by

$$\beta(t) = \frac{1}{c}\log(t_0 + t), \tag{3.1.22}$$

with $c > m$, for some $t_0 > 1$.
iii) Set $\varepsilon = 1 - m/c$, and note that $\varepsilon \in (0,1)$. We assume that

$$a(t) \le a(t_0 + t)^\delta, \tag{3.1.23}$$

$$b(t) + c(t) \le \frac{b}{\log(t_0 + t)t^{1-\delta}}, \tag{3.1.24}$$

for some positive constant a, b, δ with $\delta < \varepsilon$.

Theorem 3.6. *Under the previous assumptions, we have*

$$N(t) \le (t_0 + t)^{(\delta-1)\varepsilon}/C^2, \tag{3.1.25}$$

for a suitable $C > 0$, so that

$$|\mathrm{Prob}[X_t \in A] - \mu_t(A)| \le C\frac{\mu_t(A)^{1/2}}{(t_0 + t)^{(\delta-1)\varepsilon/2}}. \tag{3.1.26}$$

Proof. By the assumptions collected above, eq. (3.1.20) and Lemma 3.4, we have that $N(t)$ satisfies a differential inequality of the form (1.4.19), with coefficients satisfying (1.4.20), (1.4.21). The thesis therefore follows from Lemma 1.18 and from (3.1.16).

□

Remark 3.7. The hypotheses of Theorem 3.6 do not necessarily imply that U converge pointwise. In particular, it may happen that $||U(t)||_\infty$ diverges as $t \to \infty$. However, if $U(t) = U$ independent of t sufficient conditions are known in order that the weak limit μ_∞ of μ_t as $t \to \infty$ exist. In such a case μ_∞ is concentrated on the set K of absolute minima of U. The same conclusions hold if $U(t)$ tends to U pointwise sufficiently fast so that

$$\beta(t)||U(t) - U||_\infty \to 0 \text{ as } t \to \infty. \tag{3.1.27}$$

Remark 3.8. We have restricted ourselves to the simplest estimates (i.e. in L^2–norm) on the Radon–Nikodym derivative ν_t. Better bounds could be obtained generalizing the use of families of ordinary Sobolev inequalities for the instantaneous generators of the diffusion considered, which was explained in [6, section 2] for the case of a time–independent energy function. However such estimates are not easily generalizable to the case of a non–compact state space such as $\mathbb{R}^n$. In fact, the aim of the next section is to prove results of the form of Theorem 3.6, with methods modelled on the analysis of this section.

Remark 3.9. The conclusion of Theorem 3.7 imply in particular that the distribution of the process and μ_t become indistinguishable also in the total variation norm in the limit as $t \longrightarrow \infty$.

3.2 The L^2–approach in $\mathbb{R}^n$

The analysis of the previous section has provided estimates which depend on the existence of an upper bound for $U(t)$ for each $t > 0$. Once more, this is not convenient if one wants to analyse the case in which the state space is non–compact.
We show in this section how certain relative boundedness assumptions on U and on its time derivative U' can be used to prove bounds of the previous form. The additional problem of proving the existence of L^2–solutions to (the analogous of) the time–dependent heat equation (3.1.10) will be solved under certain assumptions on the time dependence of $U(t)$.

The notation here will be the same of section 3.1, unless explicitly indicated, with $\mathbb{R}^n$ replacing $\mathbb{M}$ and Lebesgue measure replacing ν. L^2 will denote the Hilbert space of (equivalence classes of) square integrable functions w.r.t. Lebesgue measure. We suppose that $U : [0,\infty)\times\mathbb{R}^n \to [0,\infty)$ is a non–negative function of class $C^{1,2}$, that is continuously differentiable in $t \in [0,\infty)$ and twice continuously differentiable in $x \in \mathbb{R}^n$; we also assume that $\lim_{|x|\to\infty} U(x) = \infty$.

In addition, let $L_t : C_0^\infty(\mathbb{R}^n) \to C_0^\infty(\mathbb{R}^n)$ be defined as in (3.1.1), and let H_t be the Schrödinger operator (densely defined in L^2) which is unitarily equivalent to L_t under the mapping $f \mapsto f\Phi_t$, $f \in L^2(\mu_t)$. Here β is a monotonically non–decreasing bijection of $[0,\infty)$ onto its range, this latter one being contained in $[0,\infty)$.
We now list the assumptions which will be used in the sequel.

Assumption 3.10 (relative boundedness);
$U'(t)$ belongs to $L^1(\mu_t)$ for all $t \geq 0$, and in addition there exist constants $a, b > 0$, $d \in (0,1)$ such that

$$0 \leq U(t) \leq aV_t + b; \tag{3.2.1}$$

$$\frac{1}{2}\beta(t)[U'_+(t) + \langle U'_-\rangle_t] \leq dV_t + h(t), \tag{3.2.2}$$

pointwise, and hence in the sense of quadratic forms for the corresponding multiplication operators in L^2, for a suitable function function $h(t)$ which tends to zero as $t \to \infty$ at a rate to be determined below (see Theorem 3.20).

Remark 3.11. Assumption 3.10 implies that

$$V(t) \geq (1 - d - \beta' a/2)V_t + c(t), \tag{3.2.3}$$

for a suitable continuous function $c(t)$. Possibly by changing the scale of time we may suppose that $1 - d - \beta' a/2 > 0$ for all t. Hence $V(t)$ is bounded below. This in addition implies that $-H(t)$ is essentially self–adjoint, with closure bounded below.

Moreover, (3.2.1) implies in particular that $\lim_{|x|\to\infty} V_t(x) = \infty$ for all t. Hence $-H_t$ has compact resolvent and hence a purely discrete spectrum. Finally, $\lambda = 0$ is its lowest eigenvalue, whose unique eigenvector is Φ_t. Let γ_t be the lowest strictly positive eigenvalue of $-H_t$. It is known (cf. [16, 17]) that

$$\gamma_t \geq \gamma \exp[-\beta(t)m(t)] \tag{3.2.4}$$

for a suitable $\gamma > 0$ independent of V_t, where $m(t)$ has an informal description as the maximum height which must be crossed in order to reach a global mimimum of $U(t)$ from a local one.

Remark 3.12. As explained in section 1.3, (3.2.1) is related, for each $t > 0$, to hypercontractivity properties of (the closure of) $-H_t$ and, at least for sufficiently regular potentials, implies that each $U(t)$ must diverge as $|x| \to \infty$ at least as const.$\times |x|^2$. Moreover, if $U(t) = U$ independent of time, (3.2.2) is a consequence of (3.2.1), and the same holds if $U'(t,x)$ tends to zero as $t \to \infty$ uniformly in x sufficiently fast.

Lemma 3.13. *Under Assumption* 3.10, *the martingale problem for L_t is well-posed in the sense of* [60].

Proof. Since the coefficients of the generator L_t are locally bounded, we have only to check that explosion does not occur. We shall prove that $U(t)$ is a Liapunov function. In fact, for all $T > 0$,

$$\lim_{|x|\to\infty} \min_{t\in[0,T]} U(t,x) = \infty$$

by assumption, whereby (3.2.2) implies that

$$U' \le U'_+ \le d\left(\frac{\beta}{2}|\nabla U|^2 - \Delta U\right) + \frac{2h}{\beta} - \langle U'_- \rangle_t. \tag{3.2.5}$$

In turn, the Liapunov inequality for $U(t) + c$, $c > 0$, amounts to

$$U' \le \beta|\nabla U|^2 - \Delta U + l(U + c), \quad l > 0, \tag{3.2.6}$$

which is implied by (3.2.5).

□

We now turn to the second assumption of this section, which is necessary in order to prove existence and uniqueness for the L^2–evolution equation given by the analogous in $\mathbb{R}^n$ of (3.1.12). To this end, fix as usual $T > 0$, and let $c = c(T)$ be such that $-H(t) + c > \varepsilon > 0$ for all $t \in [0,T]$. With a slight abuse of notation we shall denote with $-H(t)$ also the closure of $-H(t) + c$, so that $-H(t)$ is boundedly invertible for such t and generates an analytic semigroup of any angle $\theta \in (0, \pi)$. We can now state the following

Assumption 3.14 (continuity);
for each $t > s \ge 0$ the operator $(V(t) - V(s))H(t)^{-1}$ is bounded, and there exists $C > 0$, $\gamma \in (0,2)$ such that

$$||(V(t) - V(s))H(s)^{-1}|| \le c(t-s)^\gamma. \tag{3.2.7}$$

Consider the time dependent heat equation

$$\frac{d}{dt} f_t = -H(t) f_t = -[-\Delta + V(t)] f_t, \tag{3.2.8}$$

where

$$V(t,x) = V_t(x) - \frac{1}{2}\beta'(t)[U(t,x) - \langle U(t,\cdot)\rangle_t] - \frac{\beta}{2}[U'(t,x) - \langle U'(t,\cdot)\rangle_t]. \tag{3.2.9}$$

With the notation of Section 2.3 we say that $u \in C([0,T],\mathrm{L}^2)$ is a strict (respectively classical) solution of (3.2.8) if $u \in C^1([0,T],\mathrm{L}^2) \cap C([0,T],D[H(t)])$ (respectively $u \in C^1((0,T],\mathrm{L}^2) \cap C((0,T],D[H(t)])$). Proceeding exactly as in Proposition 2.14, we can show the following

Proposition 3.15. *Let $f_0 \in D(H(0))$ (respectively $f_0 \in L^2$). Then there exist a unique strict (respectively classical) solution to* (3.2.8) *with initial condition f_0.*

Remark 3.16. By ([61], [62, Theorem B.7.1), the operators $\exp[-sH_t]$, $t \geq 0$, $s > 0$ have integral kernels $K_{(t)}(s,x,y)$, jointly continuous in (s,x,y). In the sequel, for the sake of simplicity, we shall take as initial condition for (3.2.8) functions of the form

$$f = \exp[-\varepsilon H_{t=0}]f_0, \tag{3.2.10}$$

where ε is a positive constant and f_0 is either an arbitrary function in L^2, or the Dirac delta measure δ_x for some $x \in \mathbb{R}^n$.

Remark 3.17. Suppose that $U(t) = U$ independent of t. Set $V_2 = V_{\beta=2}$ and assume that $\beta(t) > 2$ (possibly by changing the "energy units") and that

$$0 \leq U \leq aV_2 + c \tag{3.2.11}$$

for suitable positive constants a, c. It is easy to realize that $V_\beta \geq \beta V_2/2$, and so (3.2.11) implies that (3.2.1) holds. Moreover, the continuity assumption (3.2.7) is satisfied with $\gamma = 1$ whenever one has

$$D(H(t)) = H^2(\mathbb{R}^n) \cap D(V(t)), \tag{3.2.12}$$

as has been shown in section 2.3.

The following Lemma is proved exactly as in section 1.1

Lemma 3.18. *Let $Q_t(f)$ denote the Dirichlet form associated with $-H_t$, and let $N(t) = ||f_t - \Phi_t||_2^2$. Then*

$$\frac{d}{dt}N(t) = -2Q_t(f_t) + \beta'(f_t,[U - \langle U\rangle_t]f_t) + \beta(f_t,[U' - \langle U'\rangle_t]f_t) \tag{3.2.13}$$

Moreover, one has

Lemma 3.19. *The following differential inequality holds:*

$$\frac{d}{dt}N(t) \leq -N(t)\,[(2 - 2d - \beta'(t)a)\gamma_t - \beta'(t)b - 2h(t)] + \beta'(t)b + 2h(t). \tag{3.2.14}$$

Here γ_t is the spectral gap of $-H_t$, whereas $h(t)$ and the constants involved in (3.2.14) *are as in Assumption* 3.10.

Proof. The relative boundedness inequalities (3.2.1),(3.2.2) are form bounds for the corresponding multiplication operators. Hence (3.2.13) implies that

$$\frac{d}{dt}N(t) \leq -Q_t(f_t)(2-\beta'(t)a-2d)+||f_t||_2^2(\beta'(t)b+2h(t)),$$

because $\langle U(t)\rangle_t$, $\langle U'_+(t)\rangle_t$, $U'_-(t)$ are non–negative. Since $Q_t(f_t)=Q_t(f_t-\Phi_t)$ and $||f_t||_2^2=N(t)+1$, this can be rewritten, by the spectral gap condition, as

$$\frac{d}{dt}N(t) \leq -(2-d-\beta'(t)a)\gamma_t N(t)+(N(t)+1)(\beta'(t)b+2h(t)),$$

which is (3.2.14).

□

Finally, the latter Lemma can be used to compare (in the total variation norm or weakly) the difference between the distribution of the process and the equilibrium distributions μ_t. See remark 3.7 as concerns the existence of a weak limit of μ_t as $t\to\infty$. In fact, we have the following

Theorem 3.20. *Assume that there exists $m>0$ such that $m(t)<m$, $m(t)$ being as in* (3.2.4). *Let $c>m$, $\varepsilon=1-m/c$ and $\beta(t)=c^{-1}\log(t_0+t)$ ($t_0>0$). Finally suppose that, for some $\delta\in(0,\epsilon)$, we have $h(t)\leq \mathrm{const.}\,(t_0+t)^{\delta-1}$. Then, for each $\delta\in(0,1)$, there exists $C>0$ such that*

$$|\mathrm{Prob}[X_t\in A]-\mu_t(A)| \leq C\frac{\mu_t(A)^{1/2}}{(t_0+t)^{(\delta-1)\varepsilon/2}}. \tag{3.2.15}$$

In particular, if μ_t has a limit μ in the weak topology as $t\to\infty$, then the distribution of the process X_t converges to μ weakly.

Proof. It suffices to prove that

$$N(t)\leq (t_0+t)^{(\delta-1)\varepsilon}/C^2. \tag{3.2.16}$$

In fact, (3.2.16) follows from Lemmas 3.18 and 1.18 by noting that, by assumption, the coefficients of the differential inequality (3.2.14) satisfy (1.4.20), (1.4.21) with ε,δ as above.

□

Chapter 4:

THE ALGEBRAIC APPROACH

In the previous chapters, we have discussed the Langevin algorithm on $\mathbb{R}^n$. This procedure is the continuous–time, continuous–space version of what is usually called *simulated annealing*. What we shall show here and in the following chapter is that simulated annealing corresponds to a particular case of a more general class of time–inhomogeneous evolutions on von Neumann algebras, and specifically to the case in which the algebra is commutative (and hence isomorphic to some L^∞ space). This can also be seen as a generalization of the theory of the asymptotic behaviour of dynamical semigroups with a faithful normal invariant state, cf. [63–69].

Here we introduce this class of evolutions, and study some of their main properties. We restrict to the case of discrete time evolutions and, after having recalled some general facts on von Neumann algebras and their modular theory in section 4.1, we discuss the class of evolution considered in section 4.2 which contains the main result of the chapter, whereas section 4.3 is devoted to rephrasing it in terms of conditions on the relative Hamiltonians between the invariant states at "time" n and $n+1$. Finally, in section 4.4 we discuss possible extension to C^*–algebras and non–normal states. Our general references for the theory of von Neumann algebras will be [41, 70].

4.1 Preliminaries

Let $\mathcal{M}$ be a von Neumann algebra of operators on a separable Hilbert space $\mathcal{H}$, which admits a cyclic and separating vector Ψ in $\mathcal{H}$.

Denote by ψ and by ψ' the faithful normal states on $\mathcal{M}$ and on the commutant $\mathcal{M}'$ respectively defined by

$$\psi(a) = \langle \Psi, a\Psi\rangle \ : \quad a \in \mathcal{M}\ ; \tag{4.1.1}$$

$$\psi'(a') = \langle \Psi, a'\Psi\rangle \ : \quad a' \in \mathcal{M}'. \tag{4.1.2}$$

Let Δ_Ψ and J be the modular operator and the modular involution canonically associated with the pair $(\mathcal{M},\Psi)$ by the Tomita–Takesaki theory, and let $V = \overline{\Delta^{1/4}\mathcal{M}_+\Psi}$ be the natural positive cone. For each normal state φ on $\mathcal{M}$ there exists a unique vector Φ in V such that

$$\varphi(a) = \langle \Phi, a\Phi\rangle \ : \ a \in \mathcal{M}. \tag{4.1.3}$$

The relative modular operator $\Delta_{\Phi,\Psi}$ is defined by

$$\Delta_{\Phi,\Psi} = S^*_{\Phi,\Psi}\overline{S}_{\Phi,\Psi}, \tag{4.1.4}$$

where

$$S_{\Phi,\Psi}a\Psi = a^*\Phi \ : \quad a \in \mathcal{M}. \tag{4.1.5}$$

Denote by $S(\mathcal{M})$ the set of all normal states on $\mathcal{M}$ and by $S_\psi(\mathcal{M})$ the set of those normal states on $\mathcal{M}$ which are majorized by a scalar multiple of ψ. The following Lemma is well–known (cf. [41]).

Lemma 4.1. *For any $\varphi \in S(\mathcal{M})$ the following conditions are equivalent:*
i) $\varphi \in S_\psi(\mathcal{M})$;
ii) there exist a (unique) element $x = x_\varphi$ of $\mathcal{M}'_+$ such that

$$\varphi(a) = \langle x_\varphi \Psi, a\,\Psi\rangle \ : \quad a \in \mathcal{M}; \tag{4.1.6}$$

iii) the Connes cocycle $\{(D\varphi : D\psi)_t = \Delta^{it}_{\Phi,\Psi}\Delta_\Psi^{-it} : t \in \mathbb{R}\} \subseteq \mathcal{M}$ extends to an analytic function on the strip $z \in \mathbb{C}: -1/2 < \mathrm{Im}\ z < 0$, continuous on the boundaries, with values in $\mathcal{M}$.
Moreover, one has

$$\Phi = (D\varphi : D\psi)_{-i/2}\Psi, \tag{4.1.7}$$

$$x_\varphi = J\left[(D\varphi : D\psi)_{-i/2}\right]^*(D\varphi : D\psi)_{-i/2}J. \tag{4.1.8}$$

An immediate consequence of the equivalence $i) \Longleftrightarrow ii)$ is the following

Corollary 4.2. *$S_\psi(\mathcal{M})$ is norm–dense in $S(\mathcal{M})$.*
Proof. A state φ in $S(\mathcal{M})$ can be written in the form

$$\varphi(a) = \sum_{j=1}^{\infty}\langle\zeta_j, a\zeta_j\rangle \ : \quad a \in \mathcal{M},$$

with $\zeta_j \in \mathcal{H}$, such that $\sum_{j=1}^{\infty}||\zeta_j||^2$. Given $\varepsilon > 0$, there exist a positive integer n and elements $x_1, \dots, x_n$ of $\mathcal{M}'$ such that, letting

$$\varphi_n(a) = \sum_{j=1}^{n}\langle x_j\Psi, a\,x_j\Psi\rangle \ : \quad a \in \mathcal{M}$$

one has

$$||\varphi_n(\mathbf{1})^{-1}\varphi_n - \varphi|| \le \varepsilon.$$

Moreover, $\varphi_n(\mathbf{1})^{-1}\varphi_n$ is a state in $S_\psi(\mathcal{M})$, since

$$\varphi_n(\mathbf{1})\varphi_n(a) = \langle x\,\Psi, a\,\Psi\rangle,$$

where

$$x = \left(\langle\Psi, \sum_{j=1}^{n} x_j^* x_j\Psi\rangle\right)^{-1}\sum_{j=1}^{n} x_j^* x_j.$$

□

Let $h = h^* \in \mathcal{M}$. Then the expression

$$\Psi(h) = \sum_{k=0}^{\infty}(-1)^k \int_0^{1/2} dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{k-1}} dt_k \Delta_\Psi^{t_k} h \Delta_\Psi^{t_{k-1}-t_k} h \ldots \Delta_\Psi^{t_1-t_2} h\Psi \qquad (4.1.9)$$

is norm–convergent, and h is said to be the relative Hamiltonian between the state ψ^h given by

$$\psi^h(a) = \frac{\langle \Psi(h), a\,\Psi(h)\rangle}{\langle \Psi(h), \Psi(h)\rangle} \quad : \quad a \in \mathcal{M} \qquad (4.1.10)$$

and ψ. Note the different conventions concerning sign and normalization as compared to Araki [71].

Given ψ and h, the perturbed state ψ^h is the unique faithful state in $S(\mathcal{M})$ maximizing the functional

$$\varphi \mapsto \langle \Phi, \log \Delta_{\Psi,\Phi} \Phi\rangle - \varphi(h), \qquad (4.1.11)$$

where $\langle \Phi, \log \Delta_{\Psi,\Phi}\Phi\rangle \leq 0$ is known as the relative entropy of φ with respect to ψ (the opposite sign convention is also used in the literature). This variational characterization of ψ^h makes sense also for self–adjoint operators h affiliated with $\mathcal{M}$ which are bounded from below but unbounded from above and may possibly have $+\infty$ as an eigenvalue (see Donald [72, 73]). With this extended notion of ψ^h, for each φ in $S_\psi(\mathcal{M})$ there exists a unique h such that $\varphi = \psi^h$. However, given ψ and h, the state ψ^h need not be in $S_\psi(\mathcal{M})$; a sufficient condition for $\psi^h \in S_\psi(\mathcal{M})$ is that $\Delta_\Psi^t h \Delta_\Psi^{-t} \in \mathcal{M}$ for all $t \in [0, 1/2]$ (cf. Lemma 4.4 below).

We recall that a positive map τ on $\mathcal{M}$ is said to be completely positive is $\tau \otimes \mathbf{1}_n$ is positive on $\mathcal{M} \otimes \mathbb{C}^n$ for each $n = 1, 2, \ldots$.

Definition 4.3. A *dynamical map* τ on $\mathcal{M}$ is a completely positive identity preserving weakly*–continuous linear map of $\mathcal{M}$ into itself.

Lemma 4.4. *Let τ be a dynamical map on $\mathcal{M}$, leaving ψ invariant. Then there exists a dynamical map τ' on $\mathcal{M}'$, leaving ψ' invariant, such that*

$$\langle \tau'(a')\,\Psi, a\,\Psi\rangle = \langle a'\Psi, \tau(a)\,\Psi\rangle \quad : \quad a \in \mathcal{M},\ a' \in \mathcal{M}'. \qquad (4.1.12)$$

Proof. If φ is in $S_\psi(\mathcal{M})$ and τ leaves ψ invariant, then $\varphi \circ \tau$ is in $S_\psi(\mathcal{M})$. Define τ' by linear extension of

$$\tau'(x_\varphi) = x_{\varphi\circ\tau} \quad : \quad \varphi \in S_\psi(\mathcal{M}). \qquad (4.1.13)$$

Then τ' is a positive weakly*–continuous linear map of $\mathcal{M}'$ into itself, satisfying (4.1.13), and $\tau'(\mathbf{1}) = \mathbf{1}$ since

$$\langle \tau'(\mathbf{1})\Psi, a\,\Psi\rangle = \langle \Psi, \tau(a)\Psi\rangle = \langle \Psi, a\,\Psi\rangle \quad : \quad a \in \mathcal{M},$$

and $\psi' \circ \tau' = \psi'$ since

$$\langle \Psi, \tau'(a')\Psi\rangle = \langle \tau(\mathbf{1})\Psi, a'\Psi\rangle = \langle \Psi, a'\Psi\rangle : \quad a' \in \mathcal{M}'.$$

Complete positivity is shown as follows: let $a_1, \ldots, a_n \in \mathcal{M}$, $x_1, \ldots, x_n \in \mathcal{M}'$. Since τ is completely positive, one has

$$0 \leq \sum_{i,j=1}^{n} \langle x_i \Psi, \tau(a_i^* a_j) x_j \Psi \rangle = \sum_{i,j=1}^{n} \langle x_j^* x_i \Psi, \tau(a_i^* a_j) \Psi \rangle$$

$$= \sum_{i,j=1}^{n} \langle \tau'(x_j^* x_i) \Psi, a_i^* a_j \Psi \rangle = \sum_{i,j=1}^{n} \langle a_i \Psi, \tau'(x_i^* x_j) a_j \Psi \rangle.$$

Since $\mathcal{M}\Psi$ is dense in $\mathcal{H}$, also τ' is completely positive.

□

Lemma 4.5. *Let τ be a dynamical map on $\mathcal{M}$, leaving ψ invariant. Then there exists a contraction T on $\mathcal{H}$ such that*

$$T(a\Psi) = \tau(a)\Psi \ : \quad a \in \mathcal{M}, \tag{4.1.14}$$

$$T^*(a'\Psi) = \tau'(a')\Psi \ : \quad a' \in \mathcal{M}', \tag{4.1.15}$$

Proof. By the Kadison–Schwarz inequality, $\tau(a^* a) - \tau(a^*)\tau(a) \geq 0$ for all $a \in \mathcal{M}$, so that we have

$$\begin{aligned} ||\tau(a)\Psi||^2 &= \psi(\tau(a^*)\tau(a)) \leq \psi(\tau(a^* a)) \\ &= \psi(a^* a) = ||a\Psi||^2 \ : \quad a \in \mathcal{M}. \end{aligned}$$

Then the linear operator T defined on $\mathcal{M}\Psi$ by (2.14) extends to a contraction on $\mathcal{H}$. For $a \in \mathcal{M}$, $a' \in \mathcal{M}'$, we have

$$\langle a'\Psi, T(a\Psi) \rangle = \langle a'\Psi, \tau(a)\Psi \rangle = \langle \tau'(a')\Psi, a\Psi \rangle$$

so that (4.1.15) holds.

□

Lemma 4.6. *In the situation of Lemma 4.5, the following are equivalent (for real $\gamma > 0$):*

i)

$$||\tau(a)\Psi|| \leq e^{-\gamma} ||a\,\Psi|| \quad \textit{for all } a \textit{ in } \mathcal{M} \textit{ with } \psi(a) = 0; \tag{4.1.16}$$

ii)

$$||T\Lambda|| \leq e^{-\gamma} ||\Lambda|| \quad \textit{for all } \Lambda \textit{ in } \mathcal{H} \textit{ with } \langle \Psi, \Lambda \rangle = 0; \tag{4.1.17}$$

iii)

$$||\tau'(a')\Psi|| \leq e^{-\gamma} ||a'\,\Psi|| \quad \textit{for all } a' \textit{ in } \mathcal{M}' \textit{ with } \psi'(a') = 0; \tag{4.1.18}$$

iv)

$$||T^*\Phi|| \leq e^{-\gamma} ||\Phi|| \quad \textit{for all } \Phi \textit{ in } \mathcal{H} \textit{ with } \langle \Psi, \Phi \rangle = 0. \tag{4.1.19}$$

Proof. Let $\mathcal{K}$ be the orthogonal complement of Ψ in $\mathcal{H}$. Since $T\Psi = T^*\Psi = \Psi$, T and T^* map $\mathcal{K}$ into itself, and $(T|_{\mathcal{K}})^* = T^*|_{\mathcal{K}}$. Then *ii)* and *iv)* are equivalent. Clearly *i)* is a special case of *ii)* and *iii)* is a special case of *iv)*. Conversely, $T|_{\mathcal{K}}$ is the closure of the map $a\,\Psi \mapsto \tau(a)\Psi$ with $\psi(a) = \langle\Psi, a\,\Psi\rangle = 0$, so that *i)* implies *ii)*, and $T^*|_{\mathcal{K}}$ is the closure of the map $a'\,\Psi \mapsto \tau'(a')\Psi$ with $\psi'(a') = \langle\Psi, a'\,\Psi\rangle = 0$, so that *ii)* implies *iv)*.

□

Remark 4.7. The equivalent conditions of Lemma 4.6 imply that

$$\lim_{k\to\infty} \varphi \circ \tau^k(a) = \psi(a) \quad \forall a \in \mathcal{M}, \quad \varphi \in S(\mathcal{M}). \tag{4.1.20}$$

The converse implication is not true in general.

4.2 Time–inhomogeneous evolutions on von Neumann algebras and non–commutative annealing

By "time–inhomogeneous evolution" on a von Neumann algebra $\mathcal{M}$ we mean a sequence $\{\tau_n : n = 1, 2, \ldots\}$ of dynamical maps, in the sense of Definition 4.3. To help intuition, τ_n may be regarded as the map describing evolution of the observables of a physical system from time t_{n-1} to time t_n, where $0 = t_0 < t_1 < \ldots < t_n \to \infty$. We assume that each τ_n has a unique (faithful normal) invariant state ψ_n, with cyclic and separating vector Ψ_n, and we investigate under which conditions, for any initial normal state φ_0 on $\mathcal{M}$, the time–evolved state $\varphi_n = \varphi_0 \circ \tau_1 \ldots \circ \tau_n$ becomes indistinguishable from ψ_n in the limit as $n \to \infty$.
Several results exist in the literature for the case when all τ_n are the same map with a faithful normal invariant state, or $\tau_n = \exp[(t_n - t_{n-1})\mathcal{L}]$, $\mathcal{L}$ being the generator of a dynamical semigroup (asymptotic behaviour of dynamical semigroups with a faithful normal invariant state) [63–69].

In the usual simulated annealing procedure, a time–inhomogeneous evolution of a (fictitious classical) physical system is used to minimize a nonnegative function U on a space X (interpreted as the energy function of the system); then the instantaneous invariant states ψ_n are Gibbs states with energy function U and inverse temperatures β_n diverging to $+\infty$.

In the usual commutative situation, the maps τ_n are in many cases symmetric with respect to their invariant states ψ_n, so that $\psi_n(a\tau_n(b)) = \psi_n(\tau_n(a)b)$ (*detailed balance*). If this happens, then the contraction T_n on $\mathcal{H}$ associated to τ_n as in Lemma 4.5 is self–adjoint. However, we shall by no means use this condition, but only assumptions of the following two kinds:
i) an estimate on the spectral gap of τ_n extended to a contraction operator on the GNS space of $(\mathcal{M}, \psi_n)$;
ii) an estimate on the difference (in a suitable sense to be defined below) between ψ_n and ψ_{n-1}.
In particular, we need an assumption of absolute continuity in the form $\psi_{n-1} \leq \lambda_n \psi_n$ for suitable constants $\lambda_n > 0$ for all n.

We do not address ourselves to the question *i*), and we just remind the reader that some partial result have been obtained in [44] for finite quantum systems (see section 5.3 and the Appendix), and in [45] for some infinite quantum systems. Concerning *ii*), we shall give sufficient conditions on the sequence of relative Hamiltonians h_n between ψ_n and ψ_{n-1} ensuring that the above mentioned difference is small enough (in norm) to allow application of our general argument.

Note that by Lemma 4.4, there exists a sequence $\{\tau_n' : n = 1, 2, \ldots\}$ of dynamical maps on $\mathcal{M}'$ such that

$$\langle \tau_n'(a')\Psi_n, a\,\Psi_n\rangle = \langle a'\Psi_n, \tau_n(a)\Psi_n\rangle \;:\quad a \in \mathcal{M},\; a' \in \mathcal{M}'. \tag{4.2.1}$$

Assume that each τ_n has a spectral gap, in the sense that there exist strictly positive constants $\gamma_n : n = 1, 2 \ldots$ such that, for all $n = 1, 2, \ldots$,

$$||\tau_n(a)\Psi_n|| \leq e^{-\gamma_n} ||a\,\Psi_n|| \quad \text{for all } a \text{ in } \mathcal{M} \text{ with } \psi_n(a) = 0. \tag{4.2.2}$$

By Lemma 4.6, a similar spectral gap holds also for τ'_n.

Assume also that there exists a sequence $R_n : n = 1, 2, \dots$ of elements of $\mathcal{M}'$ such that

$$R_n \Psi_n = \Psi_{n-1} \quad : n = 2, 3, \dots . \tag{4.2.3}$$

Equivalently (see Lemma 4.1), for $n = 2, 3, \dots$, ψ_{n-1} is majorized by a scalar multiple $\lambda_n \psi_n$ of ψ_n, and $R_n \in \mathcal{M}'$ is such that

$$R_n^* R_n = x_{\psi_{n-1}, n}, \tag{4.2.4}$$

where $x_{\psi_{n-1}, n}$ is the unique positive element of $\mathcal{M}'$ such that

$$\psi_{n-1}(a) = \langle x_{\psi_{n-1}, n} \Psi_n, a \Psi_n \rangle : \quad a \in \mathcal{M}. \tag{4.2.5}$$

Our problem is to find conditions on $\{\gamma_n\}$ and on $\{R_n\}$ ensuring that, for any initial state $\varphi_0 \in S(\mathcal{M})$, letting $\varphi_n = \varphi_{n-1} \circ \tau_n : n = 1, 2, \dots$, one has

$$\lim_{n \to \infty} \|\varphi_n - \psi_n\| = 0 \quad . \tag{4.2.6}$$

By Corollary 4.2, it suffices to prove (4.2.6) for φ_0 in the dense set $S_{\psi_1}(\mathcal{M})$. Then

$$\varphi_0(a) = \langle x_{\varphi_0, 1} \Psi_1, a \Psi_1 \rangle : \quad a \in \mathcal{M} \tag{4.2.7}$$

for a suitable positive element $x_{\varphi_0, 1}$ of $\mathcal{M}'$, and

$$\varphi_1(a) = \varphi_0(\tau_1(a)) = \langle x_{\varphi_0, 1} \Psi_1, \tau_1(a) \Psi_1 \rangle$$

$$= \langle \tau'_1(x_{\varphi_0, 1}) \Psi_1, a \Psi_1 \rangle = \langle x_1 \Psi_1, a \Psi_1 \rangle = \langle \Lambda_1, a \Psi_1 \rangle : \quad a \in \mathcal{M}, \tag{4.2.8}$$

where

$$x_1 = \tau'_1(x_{\varphi_0, 1}) \in \mathcal{M}'_+; \quad \Lambda_1 = x_1 \Psi_1 = T_1^* x_{\varphi_0, 1} \Psi_1. \tag{4.2.9}$$

Lemma 4.8. *Let φ_1 be given by (4.2.8), (4.2.9). Under the above conditions, for each $n = 2, 3, \dots$, φ_n is a normal state on $\mathcal{M}$ (actually, $\varphi_n \in S_{\psi_n}(\mathcal{M})$), which can be represented in the form*

$$\varphi_n(a) = \langle x_n \Psi_n, a \Psi_n \rangle = \langle \Lambda_n, a \Psi_n \rangle : \quad a \in \mathcal{M}, \tag{4.2.10}$$

where

$$x_n = \tau'_n(R_n^* x_{n-1} R_n) \in \mathcal{M}'_+; \quad \Lambda_n = x_n \Psi_n = T_n^* R_n^* \Lambda_{n-1}. \tag{4.2.11}$$

Proof. We proceed by induction. Indeed, (4.2.10) holds for $n = 1$ and we have, for all $a \in \mathcal{M}$,

$$\varphi_n(a) = \varphi_{n-1}(\tau_n(a)) = \langle \Lambda_{n-1}, \tau_n(a) \Psi_{n-1} \rangle$$

$$= \langle \Lambda_{n-1}, \tau_n(a) R_n \Psi_n \rangle = \langle R_n^* \Lambda_{n-1}, \tau_n(a) \Psi_n \rangle$$

$$= \langle R_n^* \Lambda_{n-1}, T_n a \Psi_n \rangle = \langle T_n^* R_n^* \Lambda_{n-1}, a \Psi_n \rangle = \langle \Lambda_n, a \Psi_n \rangle,$$

with Λ_n given by (4.2.11). However, $\Lambda_{n-1} = x_{n-1}\Psi_{n-1} = x_{n-1}R_n\Psi_n$, so that also

$$\varphi_n(a) = \langle R_n^* x_{n-1} R_n \Psi_n, \tau_n(a)\Psi_n\rangle$$
$$= \langle \tau_n'(R_n^* x_{n-1} R_n)\Psi_n, a\,\Psi_n\rangle = \langle x_n\,\Psi_n, a\,\Psi_n\rangle,$$

with x_n given by (4.2.11), since R_n and x_{n-1} are in $\mathcal{M}'$.

□

Lemma 4.9. *Under the above assumptions, let*

$$\alpha_n = \gamma_n - \log||R_n||, \tag{4.2.12}$$

$$\beta_n = e^{-\gamma_n}||(R_n^* R_n - 1)\Psi_n||. \tag{4.2.13}$$

Then, for all $n = 2, 3, \ldots$,

$$||\Lambda_n - \Psi_n|| \leq e^{-\alpha_n}||\Lambda_{n-1} - \Psi_{n-1}|| + \beta_n. \tag{4.2.14}$$

Proof. We have

$$\Lambda_n - \Psi_n = T_n^* R_n^* \Lambda_{n-1} - \Psi_n = T_n^*(R_n^*\Lambda_{n-1} - \Psi_n).$$

In addition, since

$$\langle R_n^*\Lambda_{n-1} - \Psi_n, \Psi_n\rangle = \langle \Lambda_{n-1}, \Psi_{n-1}\rangle - \langle \Psi_n, \Psi_n\rangle$$
$$= \varphi_{n-1}(\mathbf{1}) - \psi_n(\mathbf{1}) = 0,$$

the spectral gap assumption and Lemma 4.6 imply that

$$||\Lambda_n - \Psi_n|| \leq e^{-\gamma_n}||R_n^*\Lambda_{n-1} - \Psi_n||.$$

Finally, note that

$$R_n^*\Lambda_{n-1} - \Psi_n = R_n^*(\Lambda_{n-1} - \Psi_{n-1}) + (R_n^* R_n - 1)\Psi_n.$$

□

The estimates in the following Theorem are very similar to those in Lemma 1.18.

Theorem 4.10. *Under the above assumptions, suppose also that there exist real constants* $\alpha > 0$, $\beta \geq 0$, $1 > \delta > \varepsilon \geq 0$ *such that*

$$\alpha_n \geq \alpha n^{\delta - 1}, \quad \beta_n \leq \beta n^{\varepsilon - 1}: \quad n = 1, 2, \ldots. \tag{4.2.15}$$

Then there is a constant C (depending on φ_0), such that

$$|\varphi_n(a) - \psi_n(a)| \leq C||a||n^{\varepsilon - \delta} \to 0 \text{ as } n \to \infty. \tag{4.2.16}$$

Proof. By Lemma 4.8, we have

$$|\varphi_n(a) - \psi_n(a)| \leq ||\Lambda_n - \Psi_n||\, ||a||,$$

and hence, by Lemma 4.9, we compute

$$||\Lambda_n - \Psi_n|| \leq \exp\left[-\sum_{k=2}^{n} \alpha_k\right] ||\Lambda_1 - \Psi_1||$$

$$+\sum_{m=2}^{n} \exp\left[-\sum_{k=m+1}^{n} \alpha_k\right] \beta_m$$

(with $\sum_{k=n+1}^{n} \alpha_k = 0$). Eq.(4.2.15) implies that

$$||\Lambda_n - \Psi_n|| \leq \exp\left[-\alpha \sum_{k=2}^{n} k^{\delta-1}\right] ||\Lambda_1 - \Psi_1||$$

$$+\beta \sum_{m=2}^{n} \exp\left[-\alpha \sum_{k=m+1}^{n} k^{\delta-1}\right] m^{\varepsilon-1}.$$

Therefore, since $\delta < 1$,

$$\sum_{k=m+1}^{n} k^{\delta-1} \geq \int_m^n (x+1)^{\delta-1} dx = \frac{1}{\delta}\left\{(n+1)^\delta - (m+1)^\delta\right\}$$

and finally

$$||\Lambda_n - \Psi_n|| \leq \exp\left\{-\frac{\alpha}{\delta}\left((n+1)^\delta - 2^\delta\right)\right\} ||\Lambda_1 - \Psi_1||$$

$$+\beta \exp\left(-\frac{\alpha}{\delta}(n+1)^\delta\right) \sum_{m=2}^{n} \exp\left(\frac{\alpha}{\delta}(m+1)^\delta\right) m^{\varepsilon-1}. \tag{4.2.16}$$

The first term in the r.h.s. of (4.2.16) tends to 0 as $n \to \infty$ exponentially fast. Concerning the second term, we have

$$\sum_{m=2}^{n} \left[\frac{\alpha}{\delta}(m+1)^\delta\right] m^{\varepsilon-1} \leq \int_1^n \exp\left[\frac{\alpha}{\delta}(x+2)^\delta\right] x^{\varepsilon-1} dx.$$

Moreover, $x(x+2)^{-1}$ is an increasing function of x in $[1,\infty)$, so that the r.h.s of the latter inequality is smaller than

$$\left(\frac{1}{3}\right)^{\varepsilon-1} \int_1^n \exp\left(\frac{\alpha}{\delta}(x+2)^\delta\right) (x+2)^{\varepsilon-1} dx$$

$$= \frac{3^{1-\varepsilon}}{\delta}\left(\frac{\delta}{\alpha}\right)^{(\varepsilon/\delta)-1}\int_u^v e^y y^{(\varepsilon-\delta)/\delta}dy,$$

where we have performed the change of variable $y = \alpha(x+2)^\delta/\delta$ and where $u = \alpha 3^\delta/\delta$, $v = \alpha(n+2)^\delta/\delta$,
For large v, the integral is of the order of $e^v v^{(\varepsilon-\delta)/\delta}$. Indeed, for $0 < \gamma \leq 1$, letting $\mu = (\delta-\varepsilon)/\delta$, we have

$$\int_2^v e^y y^{-\mu}dy = e^v v^{-\mu} - e^2 2^{-\mu} + \mu\int_2^v e^y y^{-v-1}dy$$

$$< e^v v^{-\mu} + \frac{\mu}{2}\int_2^v e^y y^{-\mu}dy,$$

so that

$$\int_2^v e^y y^{-\mu}dy < \left(1-\frac{\mu}{2}\right)^{-1} e^v v^{-\mu}.$$

In conclusion, the second term on the r.h.s. of (4.2.16) is bounded by (constant)$\times n^{\varepsilon-\delta}$.

□

Remark 4.11. Suppose that the additional assumption that, for each $n = 1, 2, \ldots$, one has

$$\psi_n(a_n\tau_n(b_n)) = \psi_n(\tau_n(a)b_n): \quad a, b \in \mathcal{M} \quad (\textit{detailed balance}), \tag{4.2.17}$$

is satisfied. Then one simply has

$$\tau_n'(a') = J_n\tau_n(J_n a' J_n)J_n,$$

where J_n is the modular involution associated with the pair $(\mathcal{M}, \Psi_n)$ in $\mathcal{H}$ (it may be the case that J_n is independent of n, as happens when the Ψ_n are in the same natural positive cone). However, detailed balance is not really needed (cf [8] for the case of the algebra of functions on a finite space), provided one can prove a spectral gap condition without it.

4.3 Relative entropy and relative Hamiltonians

Here we assume that the sequence of states $\{\psi_n : n = 1, 2, \ldots\}$ is constructed starting from ψ_1 and from a sequence $\{h_n : n = 2, 3, \ldots\}$ of relative Hamiltonians in such a way that

$$\psi_n = (\psi_{n-1})^{h_n} : \quad n = 2, 2, \ldots, \tag{4.3.1}$$

in the sense of eq. (4.1.10), and we estimate the quantities $||R_n||$ and $||(R_n^* R_n - 1)\Psi_n||$ in terms of $h_n : n = 2, 3, \ldots$.
We restrict to bounded $h_n = h_n^* \in \mathcal{M}$, although ψ^h can be defined also for $h = h^*$ bounded from below, since we need $\psi_{n-1} \in S_{\psi_n}(\mathcal{M})$ in order to have the operators $R_n \in \mathcal{M}'$ on which our analysis is based, and this in turn implies that $-h_n$ is bounded from below, so that h_n is bounded.
We are able to prove that $\psi_{n-1} \in S_{\psi_n}(\mathcal{M})$ under the assumption that the function $t \mapsto \sigma_t^{(n-1)}(h_n) = \Delta_{\Psi_{n-1}}^{it} h_n \Delta_{\Psi_{n-1}}^{-it}$ extends to an analytic function on the strip $\{z \in \mathbb{C} : -1/2 < \operatorname{Im} z < 0\}$, continuous on the boundaries, with values in $\mathcal{M}$; we believe that this is only a sufficient condition. The following Lemma will be used repeatedly in the sequel.

Lemma 4.12. *Let ψ be a faithful normal state in $\mathcal{M}$, with $\psi(a) = \langle \Psi, a\Psi \rangle : a \in \mathcal{M}$, and with associated modular automorphism group $\sigma_t = \Delta_\Psi^{it} \cdot \Delta_\Psi^{-it} : t \in \mathbb{R}$. Let $h = h^* \in \mathcal{M}$ be such that the function $t \mapsto \sigma_t(h)$ extends to an analytic function on the strip $\{z \in \mathbb{C} : -1/2 < \operatorname{Im} z < 0\}$, continuous on the boundaries, with values in $\mathcal{M}$, and let ψ^h be defined by eq. (4.1.10). Denote by Φ the normalized vector $\Psi(h)/||\Psi(h)||$. Then there exists a unique R in $\mathcal{M}'$ such that*

$$\Psi = R\Phi; \tag{4.3.2}$$

R is invertible, and

$$||R||,\ ||R^{-1}|| \leq \exp[|||h|||], \tag{4.3.3}$$

where

$$|||h||| = \sup\{||\sigma_{-is}(h)|| : 0 \leq s \leq 1/2\}.$$

Proof. Consider the differential equations

$$\begin{cases} \dfrac{d}{ds} V(s) = -V(s)\sigma_{-is}(h) : \quad 0 \leq s \leq \dfrac{1}{2} \\ V(0) = \mathbf{1} \end{cases} . \tag{4.3.4}$$

and

$$\begin{cases} \dfrac{d}{ds} \tilde{V}(s) = \sigma_{-is}(h)\tilde{V}(s) : \quad 0 \leq s \leq \dfrac{1}{2} \\ V(0) = \mathbf{1} \end{cases} . \tag{4.3.5}$$

Both equations have unique solutions in $\mathcal{M}$ satisfying the bounds

$$||V(s)||,\ ||\tilde{V}(s)|| \leq \exp\{s|||h|||\} : \quad 0 \leq s \leq \frac{1}{2}.$$

Moreover, $\tilde{V}(s) = V(s)^{-1}$ for all $s \in [0,1/2]$. Indeed,

$$\frac{d}{ds}[V(s)\tilde{V}(s)] = V(s)[-\sigma_{-is}(h) + \sigma_{-is}(h)]\tilde{V}(s) = 0$$

so that $V(s)\tilde{V}(s) = \mathbf{1}$ for all s; and in addition the constant $\mathbf{1}$ solves the differential equation for $\tilde{V}(s)V(s)$, which reads

$$\begin{cases} \frac{d}{ds}\left[\tilde{V}(s)V(s)\right] = \sigma_{-is}(h)\left[\tilde{V}(s)V(s)\right] - \left[\tilde{V}(s)V(s)\right]\sigma_{-is}(h): \quad 0 \le s \le \frac{1}{2} \\ \tilde{V}(0)V(0) = \mathbf{1} \end{cases}$$

By the uniform boundedness of $\sigma_{-is}(h)$ on $[0,1/2]$, the solution to the latter equation is unique, so that $\tilde{V}(s)V(s) = \mathbf{1}$ for all s.
By considering the iterated series

$$V(s) = \sum_{k=0}^{\infty}(-1)^k \int_0^s ds_1 \int_0^{s_1} ds_2 \ldots \int_0^{s_{k-1}} ds_k\, \sigma_{-is_k}(h)\ldots\sigma_{-is_2}(h)\sigma_{-is_1}(h), \qquad (4.3.6)$$

it is clear that

$$\begin{aligned}\Psi(h) &= V(1/2)\Psi = J\Delta_\Psi^{1/2}V(1/2)^*\Psi \\ &= J\Delta_\Psi^{1/2}V(1/2)^*\Delta_\Psi^{-1/2}J\Psi = JV(1/2)J\Psi,\end{aligned}$$

where the last equality follows from the explicit expression of $V(s)$. Hence

$$\Phi = ||V(1/2)\Psi||^{-1}JV(1/2)J\Psi$$

or

$$\Psi = ||V(1/2)\Psi||J\tilde{V}(1/2)J\Phi = R\Phi.$$

Since Φ is cyclic and separating for $\mathcal{M}$ and for $\mathcal{M}'$ as Ψ is, it follows that $R \in \mathcal{M}'$ is uniquely determined to be

$$R = ||V(1/2)\Psi||J\tilde{V}(1/2)J. \qquad (4.3.7)$$

An obvious estimate gives

$$\begin{aligned}||R|| &\le ||V(1/2)||\,||\tilde{V}(1/2)|| \\ &\le \exp\left\{\frac{1}{2}|||h|||\right\}\exp\left\{\frac{1}{2}|||h|||\right\} = \exp\{|||h|||\}.\end{aligned}$$

Moreover,

$$R^{-1} = ||V(1/2)\Psi||^{-1}JV(1/2)J.$$

We have

$$1 = ||\Psi|| = ||\tilde{V}(1/2)V(1/2)\Psi|| \le ||\tilde{V}(1/2)||\,||V(1/2)\Psi||,$$

so that

$$||V(1/2)\Psi||^{-1} \leq ||\tilde{V}(1/2)||.$$

Hence

$$||R^{-1}|| \leq ||\tilde{V}(1/2)|| \, ||V(1/2)|| \leq \exp\{|||h|||\}.$$

□

Proposition 4.13. *Suppose that the sequence $\{h_n : n = 2,3,\ldots\}$ of relative Hamiltonians is such that the functions $t \mapsto \sigma_t^{(n-1)}(h_n)$ extend to analytic functions on the strip $\{z \in \mathbb{C}: -1/2 < \operatorname{Im} z < 0\}$, continuous on the boundaries, with values in $\mathcal{M}$; let $|||h_n|||_{n-1} = \sup\{||\sigma_{-is}^{(n-1)}(h_n)|| : 0 \leq s \leq 1/2\}$. Then*

$$||R_n|| \leq \exp\{|||h_n|||_{n-1}\}, \tag{4.3.8}$$

$$||(R_n^* R_n - 1)\Psi_n|| \leq \left(\exp\{|||h_n|||_{n-1}\} - 1\right)^{1/2}. \tag{4.3.9}$$

Proof. (4.3.8) is proved in Lemma 4.1. Next,

$$||(R_n^* R_n - 1)\Psi_n||^2 = ||R_n^*\Psi_{n-1} - \Psi_n||^2$$

$$= ||R_n^*\Psi_{n-1}||^2 + ||\Psi_n||^2 - 2\mathrm{Re}\,\langle R_n^*\Psi_{n-1}, \Psi_n\rangle$$

$$= ||R_n^*\Psi_{n-1}||^2 + ||\Psi_n||^2 - 2||\Psi_{n-1}||^2 = ||R_n^*\Psi_{n-1}|| - 1$$

$$\leq ||R_n^*|| \, ||\Psi_{n-1}|| - 1 \leq \exp\{|||h_n|||_{n-1}\} - 1,$$

which is (4.3.9).

□

Remark 4.14. In applications where the spectral gap γ_n tends to 0 as $n \to \infty$, one needs $|||h_n|||_{n-1} \to 0$ faster than γ_n (at least). This implies that $||\psi_n - \psi_{n-1}|| \to 0$ as $n \to \infty$, but by no means does it necessarily follow that ψ_n converges to a limit as $n \to \infty$.

4.4 C*–algebras and generalizations

In the main application of the above results, simulated annealing, the states ψ_n represent thermal states at different temperatures of a fictitious finite (but large) physical system. For infinite physical systems thermal states at different temperatures are typically *disjoint* states on a C^*–algebra $\mathcal{A}$, meaning that for each n there is a GNS triple $(\mathcal{H}_n, \pi_n, \Psi_n)$ associated with the pair $(\mathcal{A}, \psi_n)$, the cyclic vector Ψ_n is also separating for the von Neumann algebra $\pi_n(\mathcal{A})''$, but no subrepresentation of π_n is unitarily equivalent to a subrepresentation of π_m for $n \neq m$. Unfortunately, there is no simple generalization of the above techniques to this new situation, in view of the following

Lemma 4.15. *Let ψ_1, ψ_2 be states on a C^*–algebra $\mathcal{A}$, with GNS triples $(\mathcal{H}_1, \pi_1, \Psi_1)$, $(\mathcal{H}_2, \pi_2, \Psi_2)$ such that Ψ_i is also separating for the bicommutant $\pi_i(\mathcal{A})''$ of $\pi_i(\mathcal{A})$ in $\mathcal{B}(\mathcal{H}_i)$: $i = 1, 2$. If the operator $R : \pi_2(\mathcal{A})\Psi_2 \subseteq \mathcal{H}_2 \to \mathcal{H}_1$ defined by*

$$R\pi_2(a)\Psi_2 = \pi_1(a)\Psi_1 : \quad a \in \mathcal{A} \tag{4.4.1}$$

is closable, then π_1 is unitarily equivalent to a subrepresentation of π_2.
Proof. Let q be the densely defined quadratic form on $\pi_2(\mathcal{A})\Psi_2 \subseteq \mathcal{H}_2$ given by

$$q(\pi_2(a)\Psi_2) = ||\pi_1(a)\Psi_1||^2 = ||R\pi_2(a)\Psi_2||^2 : \quad a \in \mathcal{M}. \tag{4.4.2}$$

Since R is closable, q is closable. Let x be the positive self–adjoint operator in $\mathcal{H}_2$ associated with the closure $\overline{q}$ of q: then $\pi_2(\mathcal{A})\Psi_2 \subseteq \mathcal{D}(x^{1/2})$ and

$$||x^{1/2}\pi_2(a)\Psi_2||^2 = ||\pi_1(a)\Psi_1||^2. \tag{4.4.3}$$

For unitary $u \in \mathcal{A}$, one has

$$\begin{aligned} ||x^{1/2}\pi_2(u)\pi_2(a)\Psi_2||^2 &= ||\pi_1(u)\pi_1(a)\Psi_1||^2 \\ &= ||\pi_1(a)\Psi_1||^2 = ||x^{1/2}\pi_2(a)\Psi_2||^2 \end{aligned}$$

so that $\pi_2(u)^* x \pi_2(u) = x$ in the sense of quadratic forms, and the spectral projections of x commute with $\pi_2(u)$. Since a Banach *–algebra is generated as a linear space by its unitary elements, x is affiliated with the commutant $\pi_1(\mathcal{A})'$, and

$$||\pi_2(a)x^{1/2}\Psi_2||^2 = ||x^{1/2}\pi_2(a)\Psi_2||^2 = ||\pi_1(a)\Psi_1||^2 : \quad a \in \mathcal{A}. \tag{4.4.4}$$

Let $\mathcal{K}$ be the closed subspace of $\mathcal{H}_2$ given by $\overline{\pi_2(\mathcal{A})x^{1/2}\Psi_2}$; $\mathcal{K}$ is stable under $\pi_2(\mathcal{A})$. Let U be the linear operator mapping $\pi_1(\mathcal{A})\Psi_1 \subseteq \mathcal{H}_1$ into $\mathcal{K}$ defined by

$$U\pi_1(a)\Psi_1 = \pi_2(a)x^{1/2}\Psi_2 : \quad a \in \mathcal{A}. \tag{4.4.5}$$

By (4.4.4), U extends to an isometry of $\mathcal{H}_1$ into $\mathcal{K}$. Moreover, U is actually unitary from $\mathcal{H}_1$ onto $\mathcal{K}$. Indeed, for all a, b in $\mathcal{A}$, one has

$$\langle U\pi_1(a)\Psi_1, \pi_2(b)x^{1/2}\Psi_2\rangle = \langle \pi_2(a)x^{1/2}\Psi_2, \pi_2(b)x^{1/2}\Psi_2\rangle$$

$$= \langle \pi_2(a)\Psi_2, x\pi_2(b)\Psi_2\rangle = \langle \pi_1(a)\Psi_1, \pi_1(b)\Psi_1\rangle$$

where the last two equalities follow from the fact that $x \in \mathcal{M}'$ and by polarization from (4.4.3), respectively. Hence $U^*\pi_2(b)x^{1/2}\Psi_2 = \pi_1(b)\Psi_1$ and $UU^*\pi_2(b)x^{1/2}\Psi_2 = \pi_2(b)x^{1/2}\Psi_2$. By density, $UU^* = 1$ on $\mathcal{K}$.
Now it is an easy exercise to prove that

$$U\pi_1(a)U^* = \pi_2(a)|_{\mathcal{K}} \quad \forall a \in \mathcal{M} \tag{4.4.6}$$

which proves that π_1 is unitarily equivalent to a subrepresentation of π_2.

□

For this reason, the only generalization running on the same lines as the arguments of the previous sections can be obtained by assuming the following: we have dynamical maps $\tau_n : n = 1, 2, \ldots$, all defined on the *same* von Neumann algebra $\mathcal{M}$, each map with an invariant faithful normal state $\psi_n = \langle \Psi_n, \cdot\, \Psi_n\rangle$, and there exist *closed* operators $R_n : n = 2, 3, \ldots$, affiliated with $\mathcal{M}'$, such that

$$\Psi_n \in \mathcal{D}(R_n), \quad R_n\Psi_n = \Psi_{n-1} : \quad n = 2, 3 \ldots. \tag{4.4.7}$$

Conditions equivalent to (4.4.7) with closed unbounded R_n are discussed in Kosaki [70]. In particular, it is *not* true that, if ψ_n is a faithful normal state on $\mathcal{M}$, each normal state on $\mathcal{M}$ can be represented in this form.
In order to make sense of the formulas in the previous sections in this more general situation, it suffices to assume that

$$T^*_{n-1} \text{ maps } \mathcal{H} \text{ into } \mathcal{D}(R^*_n) \quad \forall n = 2, 3, \ldots. \tag{4.4.8}$$

However, something more is needed to imitate the estimates in Lemma 4.7 and Theorem 4.8. To be specific, we assume the following.

Assumption 4.16. *Each τ_n can be written as the product of two dynamical maps $\tilde{\tau}_n$ and $\hat{\tau}_n$*

$$\tau_n = \tilde{\tau}_n\hat{\tau}_n \tag{4.4.9}$$

with similar properties: i.e. $\tilde{\tau}_n$ and $\hat{\tau}_n$ leave ψ_n invariant, so that they are associated with contractions $\tilde{T}_n$ and $\hat{T}_n$ on $\mathcal{H}$ such that

$$\tilde{T}_n(a\,\Psi_n) = \tilde{\tau}_n(a)\Psi_n; \quad \hat{T}_n(a\,\Psi_n) = \hat{\tau}_n(a)\Psi_n; \tag{4.4.10}$$

moreover $\tilde{\tau}_n$ satisfies a spectral gap condition with a constant $e^{-\tilde{\gamma}_n}$, so that

$$||\tilde{T}^*_n\Lambda|| \leq e^{-\tilde{\gamma}_n}||\Lambda|| \quad \forall \Lambda \in \mathcal{H} \textit{ with } \langle \Psi_n, \Lambda\rangle = 0, \tag{4.4.11}$$

and finally

$$\hat{T}^*_{n-1} \textit{ maps } \mathcal{H} \textit{ into } \mathcal{D}(R^*_n) : \quad n = 2, 3, \ldots. \tag{4.4.12}$$

The above conditions are rather natural if

$$\tau_n = \exp\left[(t_n - t_{n-1})\mathcal{L}_n\right] \tag{4.4.13}$$

with $t_n > 0$, $\mathcal{L}_n$ being the generator of a semigroup of dynamical maps: one can take

$$\tilde{\tau}_n = \exp\left[(t_n - t_{n-1})(1-\zeta_n)\mathcal{L}_n\right], \quad \hat{\tau}_n = \exp\left[(t_n - t_{n-1})\zeta_n\mathcal{L}_n\right] \tag{4.4.14}$$

with $0 < \zeta_n < 1$. In the case of classical Langevin diffusion on $\mathbb{R}^n$ (cf. [18]), in which $\mathcal{L}_n$ is a differential operator of the form $-\Delta + \beta_n \nabla U \cdot \nabla$, a condition of the form (4.4.12) follows from suitable intrinsic hypercontractivity properties of the semigroup generated by $\mathcal{L}_n$ provided that U grows at infinity sufficiently fast (tipically, faster than (const.)$|x|^2$, cf. chapter 1).
As a consequence of (4.4.9), we have

$$T_n^* = \hat{T}_n^* \tilde{T}_n^* : \quad n = 1, 2, \ldots . \tag{4.4.15}$$

As a consequence of (4.4.12) and of the closed graph theorem, the operators $\tilde{R}_n^*$ defined by

$$\tilde{R}_n^* = R_n^* \hat{T}_{n-1}^* : \quad n = 2, 3, \ldots \tag{4.4.16}$$

are everywhere defined and bounded.
Now define a sequence $\tilde{\Lambda}_n$ of vectors in $\mathcal{H}$ by

$$\begin{cases} \tilde{\Lambda}_1 = \tilde{T}_1^* x_{\varphi_0,1}\Psi_1 \\ \tilde{\Lambda}_n = \tilde{T}_n^* \tilde{R}_n^* \tilde{\Lambda}_{n-1} : \quad n = 2, 3 \ldots . \end{cases} \tag{4.4.17}$$

Then the vectors Λ_n such that $\varphi_n(a) = \langle \Lambda_n, a\Psi_n\rangle : \ a \in \mathcal{M}$ are given by

$$\Lambda_n = \hat{T}_n^* \tilde{\Lambda}_n : \quad n = 1, 2, \ldots . \tag{4.4.18}$$

Lemma 4.17. *Under the above assumptions, let*

$$\tilde{\alpha}_n = \tilde{\gamma}_n - \log ||\tilde{R}_n||, \tag{4.4.19)}$$

$$\tilde{\beta}_n = e^{-\tilde{\gamma}_n} ||(\tilde{R}_n^* \tilde{R}_n - 1)\Psi_n||. \tag{4.4.20}$$

Then , for all $n = 2, 3, \ldots$

$$||\tilde{\Lambda}_n - \Psi_n|| \le e^{-\tilde{\alpha}_n} ||\tilde{\Lambda}_{n-1} - \Psi_{n-1}|| + \tilde{\beta}_n. \tag{4.4.21}$$

Proof. We have

$$\tilde{\Lambda}_n - \Psi_n = \tilde{T}_n^* \tilde{R}_n^* \tilde{\Lambda}_{n-1} - \Psi_n = \tilde{T}_n^* (\tilde{R}_n^* \tilde{\Lambda}_{n-1} - \Psi_n).$$

Moreover, $R_n^* \tilde{\Lambda}_{n-1} - \Psi_n$ is orthogonal to Ψ_n since

$$\langle \tilde{R}_n^* \tilde{\Lambda}_{n-1} - \Psi_n, \Psi_n \rangle = \langle R_n^* \hat{T}_{n-1}^* \tilde{\Lambda}_{n-1}, \Psi_n \rangle - \langle \Psi_n, \Psi_n \rangle$$

$$= \langle \Lambda_{n-1}, \Psi_{n-1} \rangle - \langle \Psi_n, \Psi_n \rangle = \varphi_{n-1}(1) - \psi_n(1) = 0.$$

Then

$$||\tilde{\Lambda}_n - \Psi_n|| \leq e^{-\tilde{\gamma}_n} ||\tilde{R}_n^* \tilde{\Lambda}_{n-1} - \Psi_n||.$$

Note that

$$\tilde{R}_n^* \tilde{\Lambda}_{n-1} - \Psi_n = \tilde{R}_n^* (\tilde{\Lambda}_{n-1} - \Psi_{n-1}) + \tilde{R}_n^* \Psi_{n-1} - \Psi_n$$

and

$$\tilde{R}_n^* \Psi_{n-1} = \tilde{R}_n^* \hat{T}_{n-1} \Psi_{n-1} = \tilde{R}_n^* \hat{T}_{n-1} R_n \Psi_n = \tilde{R}_n^* \tilde{R}_n \Psi_n.$$

□

Theorem 4.18. *Under the above assumptions, suppose also that there exist real constants* $\alpha > 0,\ \beta \geq 0,\ 1 > \delta > \varepsilon \geq 0$ *such that*

$$\tilde{\alpha}_n \geq \alpha n^{\delta - 1}, \quad \tilde{\beta}_n \leq \beta n^{\varepsilon - 1}: \quad n = 1, 2, \ldots. \tag{4.4.22}$$

Then there is a constant C (depending on φ_0), such that

$$|\varphi_n(a) - \psi_n(a)| \leq C ||a|| n^{\varepsilon - \delta} \to 0 \text{ as } n \to \infty. \tag{4.4.23}$$

Proof. We have

$$|\varphi_n(a) - \psi_n(a)| = |\langle \Lambda_n - \Psi_n, a \Psi_n \rangle|$$

$$\leq ||\Lambda_n - \Psi_n|| \, ||a \Psi_n|| = ||\hat{T}_n(\tilde{\Lambda}_n - \Psi_n)|| \, ||a \Psi_n|| \leq ||\tilde{\Lambda}_n - \Psi_n|| \, ||a||.$$

It suffices to prove that

$$||\tilde{\Lambda}_n - \Psi_n|| \leq C n^{\varepsilon - \delta},$$

and this is accomplished exactly as in Theorem 4.8, taking advantage of Lemma 4.17.

□

Chapter 5:

APPLICATIONS

We have discussed in chapter 4 a general algebraic framework which, we claimed, is a natural generalization of the well-known stochastic algorithm known as *simulated annealing*. In this chapter we prove that this is the case indeed and that, as a bonus, we are able to discuss without further technical complications also the case of time dependent energy function, which we hope could be of some interest in connection with such problem as adaptive algorithms and "estimation and annealing" algorithms (see [40]).

We discuss in section 5.1, in terms of the theory of section 4, a time–inhomogeneous Markov chain on a finite set which corresponds to a simulated annealing procedure with time–dependent energy function. Then we consider once more in section 4.2 the Langevin algorithm on compact manifolds, but for the case of a piecewise time–independent generator, and show that it is falls within the general framework of chapter 4. Finally, section 5.3 shows some applications to finite quantum systems.

5.1 Time–inhomogeneous Markov chains: simulated annealing with time–dependent energy function

We consider the following variant of the well–known simulated annealing algorithm. Let X be a finite set with $|X|$ points, let μ_0 be a probability measure on X charging any point, and consider a Markov chain $\{Y_n : n \in \mathbb{N}\}$ with arbitrary starting point and with transition probabilities given by

$$P[Y_n = y|Y_{n-1} = x] = q_0(x,y)\exp[-\beta_n[U_n(y) - U_n(x)]_+] \tag{5.1.1}$$

for $y \neq x$, where

i) $q_0(x,y)$ is an irreducible transition matrix which is μ_0 reversible in the sense that

$$\alpha(x,y) := \mu_0(x)q_0(x,y) = \alpha(y,x); \tag{5.1.2}$$

ii) $\{\beta_n : n \in \mathbb{N}\}$ is a sequence of positive real numbers (inverse temperatures) with $\lim_{n\to\infty}\beta_n = +\infty$;

iii) $\{U_n : n \in \mathbb{N}\}$ is a sequence of real functions on X (energy functions).

The usual simulated annealing algorithm is obtained for $U_n = U$ independent of n. It is well–known that, if $\beta_n \leq c^{-1}\log(1+n)$, c being a sufficiently large constant, the distribution of Y_n becomes concentrated, as $n \to \infty$, on the set of global minima of U. Here we want to investigate which kind of conditions on the time dependence of U_n have to be assumed if one wants to prove similar results for the more general Markov chain (5.1.1).

As usual, our estimates are based on a comparison of the distribution Q_n of Y_n with the probability distribution

$$P_n(x) = Z_n^{-1}\exp[-\beta_n U_n(x)]\mu_0(x), \tag{5.1.3}$$

where

$$Z_n = \sum_{x \in X} \exp[-\beta_n U_n(x)]\mu_0(x). \tag{5.1.4}$$

Under suitable assumptions, we shall prove that the two distributions become indistinguishable in the course of time, so that we obtain a limiting distribution as $n \to \infty$ for Y_n if P_n has a weak limit P_∞; we may suppose that U_n tends to a limit function U_∞, so that P_∞ is concentrated on the absolute minima of U_∞. The techniques of the present work shed no light on the problem of the limiting behaviour of Y_n when P_n does not have a limit as $n \to \infty$.

In the space $\mathbb{C}^{|X|}$ of complex–valued functions on X we shall consider the scalar product

$$\langle f, g\rangle = \sum_{x \in X} \overline{f(x)} g(x)\mu_0(x),$$

the corresponding norm $||f||_2 = \langle f, f\rangle^{1/2}$ and the supremum norm $||f||_\infty = \max_{x\in X} |f(x)|$. Of course, here the von Neumann algebra $\mathcal{M}=\mathrm{L}^\infty(X)$ is identified with $\mathbb{C}^{|X|}$ endowed with $||\cdot||_\infty$, and the Hilbert space $\mathcal{H} = \mathrm{L}^2(X)$ on which it acts is identified with $\mathbb{C}^{|X|}$ endowed with $||\cdot||_2$.

For $n \in \mathbb{N}$, let P_n be the probability distribution defined by (5.1.2) and let Q_n be given by

$$\begin{aligned} &Q_0[y] = \delta_{y,x_0}, \quad x_0 \in X, \\ &Q_n[y] = \sum_{x \in X} P[Y_n = y | Y_{n-1} = x] Q_{n-1}[x] : \quad n = 1, 2, \ldots, \end{aligned} \tag{5.1.5}$$

where the transition probabilities are given by (5.1.1). Define the following maps:

$$\begin{aligned} &\tau_n : \mathbb{C}^{|X|} \mapsto \mathbb{C}^{|X|}; \\ &[\tau_n(f)](x) = \sum_{y \in X} P[Y_n = y | Y_{n-1} = x] f(y); \quad n = 1, 2, \ldots; \end{aligned} \tag{5.1.6}$$

$$\begin{aligned} &\psi_n : \mathbb{C}^{|X|} \mapsto \mathbb{C}; \\ &\psi_n(f) = \sum_{x \in X} P_n[x] f(x); \quad n = 1, 2, \ldots; \end{aligned} \tag{5.1.7}$$

$$\begin{aligned} &\varphi_n : \mathbb{C}^{|X|} \mapsto \mathbb{C}; \\ &\varphi_n(f) = \sum_{x \in X} Q_n[x] f(x); \quad n \in \mathbb{N}. \end{aligned} \tag{5.1.8}$$

Then each τ_n is a dynamical map on $\mathbb{C}^{|X|}$, in the sense of Definition 4.3. Moreover ψ_n and φ_n are states on $\mathcal{M}$; here $\mathbf{1}$ is the function which is identically equal to one on X. As regards the notation we remark that the Gibbs states are denoted by ψ instead of μ for the sake of greater homogeneity of notation with the previous chapter.

It is clear from the above definitions that, for $n = 1, 2, \ldots$ we have

$$\varphi_n = \varphi_{n-1} \circ \tau_n; \tag{5.1.9}$$

$$\psi_n(f\tau_n(g)) = \psi_n(\tau_n(f)g): \quad f, g \in \mathbb{C}^{|X|}. \tag{5.1.10}$$

So (5.1.9) shows that φ_n is the time–evolved of φ_0 in the sense of section 4.2, and (5.1.10) shows that the detailed balance condition holds.

Our aim is to control the differences $\psi_n(f) - \varphi_n(f) : f \in \mathbb{C}^{|X|}$, in the limit as $n \to \infty$. To this end, it is convenient to define Ψ_n, Λ_n in $\mathbb{C}^{|X|}$, $n = 1, 2, \ldots$, by

$$\Psi_n(x) = P_n[x]^{1/2}, \tag{5.1.11}$$

$$\Lambda_n(x) = Q_n[x] P_n[x]^{-1/2}, \tag{5.1.12}$$

so that, for all $n = 1, 2, \ldots, f \in \mathbb{C}^{|X|}$,

$$\psi_n(f) = \langle \Psi_n, f\Psi_n \rangle, \tag{5.1.13}$$

$$\varphi_n(f) = \langle \Lambda_n, f\Psi_n \rangle. \tag{5.1.14}$$

Lemma 5.1. *For each $n = 1, 2, \ldots$, there exists a self–adjoint contraction T_n on $\mathbb{C}^{|X|}$ such that*

$$T_n(f\Psi_n) = \tau_n(f)\Psi_n : \quad f \in \mathbb{C}^{|X|}. \tag{5.1.15}$$

Moreover, the eigenvalue $\lambda = 1$ of T_n is simple, with eigenvector Ψ_n, and for each $w \in \mathbb{C}^{|X|}$ with $\langle w, \Psi \rangle_n = 0$ one has

$$||T_n w||_2 \leq \exp[-\gamma_n] ||w||_2, \tag{5.1.16}$$

where $\gamma_n > 0$.

Proof. The operator T_n defined by (5.1.15) is a contraction by Lemma 4.5. Self–adjointness follows from (5.1.10). It is clear that $S\Psi_n = \Psi_n$. (5.1.16) follows for example from the work of Holley and Stroock [6], keeping into account the form (5.1.1) of the transition probabilities and assumption *i)* following (5.1.1). More precisely, it follows from [6] that

$$\gamma_n \geq \gamma \exp[-\beta_n m_n], \tag{5.1.17}$$

where $\gamma > 0$ depends only on the matrix $q_0(x, y)$ and where m_n may be informally described as the maximum depth of a local minimum of U_n which is not a global minimum (see the Appendix).

□

Lemma 5.2. *For $n = 1, 2, \ldots$, define the (self–adjoint) linear operator R_n on $\mathbb{C}^{|X|}$ by*

$$R_n(f\Psi_n) = f\Psi_{n-1} : \quad f \in \mathbb{C}^{|X|}. \tag{5.1.18}$$

Then

$$\Lambda_n = T_n R_n \Lambda_{n-1} : \quad n = 2, 3, \ldots. \tag{5.1.19}$$

Proof. An application of Lemma 4.8. However we write down once more the computation for the sake of clearness. In fact, for all $f \in \mathbb{C}^{|X|}$,

$$\langle \Lambda_n, f\Psi_n \rangle = \varphi_n(f) = \varphi_{n-1}(\tau_n(f))$$

$$= \langle \Lambda_{n-1}, \tau_n(f)\Psi_{n-1} \rangle = \langle \Lambda_{n-1}, R_n[\tau_n(f)\Psi_n] \rangle$$
$$= \langle R_n \Lambda_{n-1}, T_n(f\Psi_n) \rangle = \langle T_n R_n \Lambda_{n-1}, f\Psi_n \rangle.$$

□

Lemma 5.3. *For $n = 2, 3, \ldots$, let*

$$a_n = \gamma_n - \log ||R_n||, \tag{5.1.20}$$

$$b_n = \exp[-\gamma_n] ||(R_n^2 - 1)\Psi_n||_2. \tag{5.1.21}$$

Then

$$||\Lambda_n - \Psi_n||_2 \le \exp[-a_n] ||u_{n-1} - v_{n-1}||_2 + b_n. \tag{5.1.22}$$

Proof. This is nothing else than Lemma 4.9, since each R_n is self–adjoint by construction.

□

Proposition 5.4. *Suppose that there exist constants $a > 0$, $b > 0$, $1 > \delta > \varepsilon \ge 0$ such that*

$$a_n \ge a n^{\delta - 1}, \quad b_n \le b n^{\varepsilon - 1} : \quad n = 2, 3, \ldots. \tag{5.1.23}$$

Then, for all $f \in \mathbb{C}^{|X|}$,

$$|\varphi_n(f) - \psi_n(f)| \le (\text{constant}) \; ||f||_\infty n^{\varepsilon - \delta} \to 0 \quad \text{as } n \to \infty. \tag{5.1.24}$$

Proof. See Theorem 4.10

□

It is now necessary to make further assumptions on β_n, U_n in order to be allowed to use Proposition 5.4 In addition to *i)*–*iii)*, we shall make the following

Assumption 5.5. There exist positive constants c, h, γ, m, k, η such that $c > m$ and $m/c < \eta \le 1$, and such that

$$0 \le \beta_n - \beta_{n-1} \le \frac{1}{cn}, \quad \forall n = 1, 2, \ldots; \tag{5.1.25}$$

$$0 \le U_n(x) \le h, \quad \forall n = 1, 2, \ldots, \; \forall x \in X; \tag{5.1.26}$$

$$\gamma_n \geq \gamma \exp[-\beta_n m], \qquad \forall n = 1, 2, \ldots; \tag{5.1.27}$$

$$|U_n(x) - U_{n-1}(x)| \leq \frac{k}{n^\eta}, \qquad \forall n = 1, 2, \ldots, \ \forall x \in X. \tag{5.1.28}$$

Then (5.1.25) implies that

$$\beta_n \leq (1/c) \log n + d, \qquad \forall n = 1, 2, \ldots, \tag{5.1.29}$$

with $d = \beta_0 + (1/c)$. Note that (5.1.28) does not necessarily imply that $U_n(x)$ has a limit for $n \to \infty$.

Theorem 5.6. *Under assumptions* (5.1.25)–(5.1.28), *one has*

$$|\varphi_n(f) - \psi_n(f)| \leq \text{const.} ||f||_\infty n^{-\delta'} \to 0 \text{ as } n \to \infty, \tag{5.1.30}$$

where δ' is any positive number strictly smaller than $\eta - m/c$.

Proof. It suffices to compute a_n, b_n given in (5.1.20), (5.1.21), and prove that they satisfy (5.1.23) with $\delta = 1 - m/c$, $\varepsilon \in (1 - \eta, 1 - m/c)$ arbitrary. Then (5.1.30) follows from Proposition 2.4 for any δ' of the form $\delta' = \delta - \varepsilon$. In fact, we have from (5.1.28), (5.1.29) that

$$\gamma_n \geq \gamma e^{-md} n^{-m/c}, \tag{5.1.31}$$

whereas R_n given in (5.1.18) is the operator of multiplication by the function

$$R_n(x) = \left[\frac{P_{n-1}(x)}{P_n(x)}\right]^{1/2}. \tag{5.1.32}$$

The r.h.s. of (5.1.32) can be evaluated as follows: note that

$$R_n(x) = \exp[\{-\beta_{n-1} U_{n-1}(x) + \beta_n U_n(x)\}/2] \left(\frac{Z_n}{Z_{n-1}}\right)^{1/2},$$

and that

$$\begin{aligned}
&\sup_{x \in X} \exp[\{-\beta_{n-1} U_{n-1}(x) + \beta_n U_n(x)\}/2] \\
&\leq \exp[\{(\beta_n - \beta_{n-1}) \sup_{x \in X} [U_n(x)] + \beta_{n-1} \sup_{x \in X} |U_n(x) - U_{n-1}(x)|\}/2] \\
&\leq \exp[h/(2cn) + [(\log n)/c + d] k/(2n^\eta)] \\
&\leq \exp[k'(\log n)/n^\eta], \qquad \forall n = 1, 2, \ldots,
\end{aligned}$$

where $k' = \text{const.} > 0$; in addition

$$\exp[-\beta_{n-1} U_{n-1}(x)] = \exp[-\beta_n U_n(x) + \beta_{n-1}[U_n(x) - U_{n-1}(x)] + (\beta_n - \beta_{n-1}) U_n(x)]$$

$$\geq \exp[-\beta_n U_n(x) - [(\log n)/c + d]k/n^\eta]$$
$$\geq \exp[-\beta_n U_n(x)] \exp[-k''(\log n)/n^\eta)], \quad \forall x \in X, \ \forall n = 1, 2, \ldots,$$

for a suitable $k'' > 0$. This implies that $[Z_n/Z_{n-1}]^{1/2}$ is bounded from above as a function of n by $\exp[k''(\log n)/(2n^\eta)]$. Therefore

$$a_n = \gamma_n - \log ||R_n|| \geq \gamma e^{-md} n^{-m/c} - (k' + k''/2)\frac{\log n}{n^\eta} \geq \gamma' n^{-m/c},$$

with $\gamma' = \text{const.} > 0$.
On the other hand, we have

$$||(R_n^2 - 1)\Psi_n||_2 \leq \sup_{x \in X} |R_n^2(x) - 1|.$$

By the same kind of estimates as above we can see that there exist positive constants α', α'' such that

$$\exp[\alpha'(\log n)/n^\eta] < R_n^2(x) < \exp[\alpha''(\log n)/n^\eta],$$

for all $x \in X$ and for all $n = 1, 2, \ldots$. Hence, for a suitable positive constant h',

$$b_n = \exp[-\gamma_n]||(R_n^2 - 1)\Psi_n||_2 \leq h'\frac{\log n}{n^\eta} < bn^{\varepsilon - 1}$$

for all ε in $(1 - \eta, 1 - m/c)$ and for a suitable positive constant $b = b(\varepsilon)$.

□

Corollary 5.7. *If ψ_n has a weak limit ψ_∞ as $n \to \infty$, then φ_n converges weakly to ψ_∞ as $n \to \infty$.*

Remark 5.8. If $U_n = U$ independent of n, then ψ_n has a weak limit as $n \to \infty$, which is the uniform measure on the set of the absolute minima of U. The same conclusion holds if U_n converges pointwise to U sufficiently fast so that $\beta_n |U_n(x) - U(x)| \to 0$ as $n \to \infty$ for all x. This is compatible with (5.1.28), (5.1.29).

Remark 5.9. The above calculation are easily generalizable to a class of jump processes over compact connected finite dimensional Riemannian manifolds $\mathbb{M}$. In fact, let ν be the Riemannian measure of $\mathbb{M}$, U_n be a sequence of C^∞ nonnegative functions on $\mathbb{M}$, and q_0 be a symmetric bistochastic kernel on $\mathbb{M}$. Let

$$Z_n = \int \exp[-\beta_n U_n(x)]\nu(dx), \quad n \in \mathbb{N}; \tag{5.1.33}$$

$$P_n(dx) = Z_n^{-1} \exp[-\beta_n U_n(x)]\nu(dx), \quad n \in \mathbb{N}; \tag{5.1.34}$$

Consider the jump process over $\mathbb{M}$ with transition probabilities

$$q_n(x, y, dy) = K_n^{-1}(x) q_0(x, y) \exp[\beta_n [U_n(y) \wedge U_n(x)]_+] P_n(dy) \tag{5.1.35}$$

and jump intensities $K_n(x)$, $K_n(x)$ being the appropriate normalization factor. Intuitively, if the process starts from $x \in \mathbb{M}$, it stays at x with probability $1 - K(x)(\in [0,1))$ and jumps in a neighbourhood dy of y with probabilities given by (5.1.35), conditional upon jumping at all. The generator of the jump process is then

$$Lf(x) = \int (f(y) - f(x))q_0(x,y)\exp[\beta_n[U_n(y) \wedge U_n(x)]_+]P_n(dy), \tag{5.1.36}$$

whose bilinear form Q is given by

$$Q(f,g) = \frac{1}{2}\int (f(y) - f(x))(g(y) - g(x))q_0(x,y)e^{\beta_n(U_n(x)\wedge U_n(y))}P_n(dx)P_n(dy). \tag{5.1.37}$$

Define a family of states on $\mathcal{M} = \mathrm{L}^\infty(\mathbb{M})$ by

$$\psi_n(f) = \int f(x)P_n(dx), \qquad f \in \mathcal{L}^\infty,\ n \in \mathbb{N}; \tag{5.1.38}$$

finally, define a family of dynamical maps by

$$[\tau_n(f)](x) = \int f(y)q_n(x,y,dy)\nu(dy), \quad f \in \mathcal{L}^\infty, \quad n = 1,2,\cdots. \tag{5.1.39}$$

The time evolution is given in terms of the family of states given by

$$\varphi_0(f) = \int K(x)f(x)\nu(dx), \quad f \in \mathcal{L}^\infty, \quad x_o \in \mathbb{M}; \tag{5.1.40}$$

$$\varphi_n(f) = \varphi_{n-1} \circ \tau_n(f), \quad f \in \mathcal{L}^\infty,\ n = 1,2,\cdots, \tag{5.1.41}$$

$K(x)$ being an initial probability distribution on $\mathbb{M}$. Then everything goes through under exactly the same assumptions, provided in addition that the behaviour of the spectral gap of the self-adjoint contraction R_n defined as in Lemma 5.2 is of the usual form. Some partial results in this direction have been obtained in [75, sect. 5].

5.2 Langevin diffusions revisited: the case of a piecewise time–independent generator on compact manifolds.

Here we apply the general scheme of section 3 to a class of diffusions on compact connected finite dimensional Riemannian manifolds $\mathbb{M}$ which are described as follows: with the notation of section 3.1, we consider the Polish space $\Omega = C([0,+\infty), \mathbb{M})$, and the coordinate process X_t on Ω given by $X_t(\omega) = \omega(t), \omega \in \Omega$. Let $\mathcal{F}_t$ be the filtration generated by X_t, so that $\mathcal{F} = \cup_{t\geq 0}\mathcal{F}_t$ generates the Borel field of Ω endowed with the topology of uniform convergence on finite intervals. Finally, let $L_t : C^\infty(\mathbb{M}) \to C^\infty(\mathbb{M})$ be defined by

$$L_t f = e^{\beta(t)U(t)}\nabla \cdot \left(e^{-\beta(t)U(t)}\nabla f\right); \tag{5.2.1}$$

here we have denoted by $U(t)$ the function $U(t,\cdot)(\geq 0)$ on $\mathbb{R}^n$. Assume moreover that both U and β are piecewise constant as functions of t, and that $U(t,\cdot)$ is a function of class C^∞ for all t, so that there is a sequence $\{t_n\}_{n\in\mathbb{N}}$ with $0 = t_0 < t_1 < t_2 < \ldots < t_n \to \infty$ as $n \to \infty$ and a sequence of nonnegative functions U_n of class C^∞ on $\mathbb{M}$ with

$$U(t,x) = U_n(x) \quad \text{for } t \in [t_{n-1}, t_n), \qquad (5.2.2),$$

$$\beta(t) = \beta_n(t) \quad \text{for } t \in [t_{n-1}, t_n), \qquad (5.2.3).$$

Denote also by g and ν the Riemannian metric and the normalized Riemannian measure on $\mathbb{M}$, respectively.

Then, for each fixed t, L_t is essentially self–adjoint in $\mathrm{L}^2(\mathbb{M}, \mu_{t,\beta(t)})$ where

$$\mu_{t,\beta(t)}(dx) = Z^{-1}_{t,\beta(t)} e^{-\beta(t)U(t,x)}\nu(dx), \tag{5.2.4}$$

$$Z_{t,\beta(t)} = \int e^{-\beta(t)U(t,x)}\nu(dx); \tag{5.2.5}$$

hereafter, all integrals will be performed on $\mathbb{M}$.
Furthermore, L_t is a non-positive operator, which has $\lambda = 0$ as simple, non–degenerate, highest eigenvector, and there is a gap γ_t between $\lambda = 0$ and the rest of its spectrum.

Proposition 5.9. For each $(s,x) \in [0,+\infty) \times \mathbb{M}$, there exists a unique probability measure $P_{s,x}$ on $(\Omega, \mathcal{F})$ such that

$$\left(f(X_t) - f(x) - \int_s^t L_u f(X_u)du,\ \mathcal{F}_t\ ,\ P_{s,x}\right) \tag{5.2.6}$$

is a mean zero martingale for all $f \in C^\infty(\mathbb{M})$. Moreover, the family $P_{s,x}$ is Feller continuous and strongly Markov.

We remark that similar properties hold also if the process does not start from a fixed $x \in \mathbb{M}$, but has a (measurable) arbitrary initial probability distribution.

Let $\mathcal{H} = \mathrm{L}^2(\mathbb{M}, d\nu)$ with the usual scalar product, and let $\mathcal{M} = \mathrm{L}^\infty(\mathbb{M}, du)$ with the usual norm. We define faithful normal states ψ_n on $\mathbb{M}$ by

$$\psi_n(f) = \int f(x) d\mu_n(x) =$$

$$= \int f(x) \frac{\exp[-\beta_n U_n]}{Z_n(\beta_n)} dx = \langle \Psi_n, f \Psi_n \rangle, \tag{5.2.7}$$

where

$$\Psi_n(x) = \frac{\exp[-\beta_n U_n(x)/2]}{Z_n(\beta_n)^{1/2}}. \tag{5.2.8}$$

For each $n = 1, 2, \ldots$, define the positivity and identity preserving weakly*–continuous linear map τ_n of $\mathcal{M}$ into itself by

$$[\tau_n(f)](x) = \mathbf{E}[f(x(t_n)) | x_{t_{n-1}} = x], \tag{5.2.9}$$

where $\mathbf{E}[\,\cdot\,| x_{t_{n-1}} = x]$ denotes expectation over the paths of the Wiener process, conditional upon $x_{t_{n-1}} = x$. Let $|\cdot|$ denote the Riemannian norm of a vector field, ∇ the Riemannian gradient and $\triangle$ the Laplace–Beltrami operator on $\mathbb{M}$. Then,

$$\tau_n = \exp[(t_n - t_{n-1}) L_n], \tag{5.2.10}$$

where

$$L_n = \frac{1}{\beta_n} \triangle - \nabla U_n \cdot \nabla \tag{5.2.11}$$

Using the unitary equivalence of $\mathrm{L}^2(\mathbb{M}, d\psi_n)$ with $\mathcal{H}$ given by $f \mapsto f \Psi_n$, we can write

$$T_n(u) = \tau_n(u \Psi_n^{-1}) \Psi_n, \quad \forall u \in \mathcal{H}, \tag{5.2.12}$$

where the self–adjoint contraction T_n on $\mathcal{H}$ is given by

$$T_n = \exp\left[-\beta_n^{-1}(t_n - t_{n-1}) H_n\right], \tag{5.2.13}$$

and

$$H_n = -\triangle + \left[\beta_n |\nabla U_n|^2/4 - \beta_n \triangle U_n/2\right] = -\triangle + \Psi_n \tag{5.2.14}$$

are nonnegative self adjoint operators.

Note that H_n differs from $H_0 = -\triangle$ by a bounded multiplication operator, and hence that H_n is self–adjoint on the domain of H_0.

Finally, define a normal state φ_0 on $\mathcal{M}$ by

$$\varphi_0(f) = \int f(x) \phi_0(x) dx, \quad f \in \mathcal{M}, \tag{5.2.15}$$

ϕ_0 being a positive continuous function with $\int \phi_0(x) dx = 1$, and define faithful normal states φ_n, $n = 1, 2, \ldots$ by $\varphi_n = \varphi_{n-1} \circ \tau_n$ as in chapter 4. We now turn to verifying that the general theory of section 4.2 is applicable in such a case.

To this end, note that the operator H_n defined in (5.2.14) has a gap γ_n between $\lambda = 0$ and the rest of its spectrum. This follows for example from [12]. In particular it is known that

$$\gamma_n \geq \gamma \exp[-\beta_n m_n], \tag{5.2.16}$$

where the constant $\gamma > 0$ is independent of U_n and m_n has an informal description as the maximum depth of a local minimum of U_n which is not a global minimum.

In addition, note that

$$\varphi_1(f) = \varphi_0[T_1(f)] = \int [T_1(f)](y)\phi_0(y)dy$$

$$= \int\int K(t_1, x, y) f(x)\phi_0(y)dxdy \tag{5.2.17}$$

for a suitable (positive, jointly continuous) integral kernel K (see for example [61]). Then

$$\varphi_1(a) = \langle u_1, f v_1\rangle \tag{5.2.18}$$

where

$$u_1(x) = \left[\int K(t_1, x, y)\phi_0(y)dy\right] v_1(x)^{-1} \quad (\text{in } \mathrm{L}^2(\mathbb{M})). \tag{5.2.19}$$

Note that each map τ_n leaves invariant the norm–closed subspace $C(\mathbb{T}^d)$ of continuous functions. If we restrict to such a subspace, we may take as initial state φ_0 also $\varphi_0(f) = f(x_0)$ for arbitrary $x_0 \in \mathbb{M}$, whereby $u_1(x) = K(t_1, x, x_0)v_1(x)^{-1}$ (in $\mathrm{L}^2(\mathbb{M})$).

Finally, note that R_n is the (self–adjoint) operator of multiplication on $\mathcal{H}$ by the function $v_{n-1}\Psi_n^{-1}$, which is bounded by construction. Therefore, we have only to evaluate the quantities

$$a_n = \gamma_n - \log ||R_n||, \tag{5.2.20}$$

$$b_n = \exp[-\gamma_n]||(R_n^2 - 1)\Psi_n||_2; \tag{5.2.21}$$

we shall show that they satisfy (4.2.15). The calculation is similar to the case of a finite state space; in analogy with section 5.1, we shall make the following

Assumption 5.10. There exists a constant c such that

$$0 \leq \beta_n - \beta_{n-1} < \frac{1}{cn} \qquad \forall n = 1, 2, \ldots. \tag{5.2.22}$$

Moreover, there exist positive constants h, γ, m, k, η, with $m < c$, $m/c < \eta \leq 1$, such that:

$$0 \leq U_n(x) \leq h, \quad \forall n = 1, 2, \ldots, \ \forall x \in X; \tag{5.2.23}$$

$$\gamma_n \geq \gamma \exp[-\beta_n m], \qquad \forall n = 1, 2, \ldots; \tag{5.2.24}$$

$$|U_n(x) - U_{n-1}(x)| \leq \frac{k}{n^\eta}, \quad \forall n = 1, 2, \ldots, \ \forall x \in X. \tag{5.2.25}$$

Note that (5.2.22) implies that

$$\beta_n \leq \frac{1}{c} \log n + d_0 \quad \forall n = 1, 2, \ldots, \tag{5.2.26}$$

with $d_0 = \beta_0 + 1/c$.
Under these hypotheses, we have the following

Theorem 5.11. *Under Assumption 5.10,*

$$|\varphi_n(f) - \psi_n(f)| \leq \text{const.})||f||_\infty n^{-\delta'} \to 0 \quad \text{as } n \to \infty \tag{5.2.27}$$

for any δ' of the form $\delta' = \delta - \varepsilon$ with $\delta = 1 - m/c$ and $\varepsilon \in (1 - \eta, 1 - m/c)$.

Proof. Proceeding as in section 5.1, Theorem 5.5 it is easy to show that

$$||R_n|| \leq \alpha \exp[k'(\log n)/n^\eta], \tag{5.2.28}$$

for suitable positive constants α, k, and that

$$||(R_n^2 - 1)\Psi_n||_2 \leq h' \frac{\log n}{n^\eta}, \tag{5.2.29}$$

for a suitable positive constant h'. Noting that (5.2.24) and (5.2.26) imply that $\gamma_n \geq$ (const.) $\times\, n^{-m/c}$, it follows that there exist positive constants γ', b such that, for all $n = 1, 2, \ldots$ and for all ε in $(1 - \eta, 1 - m/c)$,

$$a_n \geq \gamma' n^{-m/c}, \quad b_n \leq b n^{\varepsilon - 1}.$$

The thesis follows from Theorem 4.10.

□

Corollary 5.12. *If ψ_n has a weak limit ψ_∞ as $n \to \infty$, then φ_n converges weakly to ψ_∞ as $n \to \infty$, and hence the distribution of the coordinate process with respect to the measure associated to L_t and to the initial condition considered converges weakly to ψ_∞.*

Remark 5.13.. As often remarked in the previous chapters, if $U_n = U$ independent of n, sufficient conditions are known in order that the weak limit ψ_∞ of ψ_n as $n \to \infty$ exists. In such a case the weak limit is a measure concentrated on the set K of absolute minima of U; however, it need not be the uniform one on K, unless $\psi_{\text{Leb}}(K) > 0$. Under suitable assumptions on U a limiting measure can be shown to exist and can be described in terms of the Hessian of U on K (see Hwang [10]).

More generally, if U_n converges pointwise to U sufficiently fast so that $\beta_n|U_n(x) - U(x)| \to 0$ as $n \to \infty$ uniformly in x, the same considerations hold. This is compatible with Assumption 4.9.

5.3 Time–inhomogeneous quantum evolutions.

Let $\mathcal{K}$ be a complex separable Hilbert space with scalar product (f,g) and norm $||f||_{\mathcal{K}} = (f,f)^{1/2}$. Let $\mathcal{M} = \mathcal{B}(\mathcal{K})$ be the von Neumann algebra of *all* bounded linear operators on $\mathcal{K}$ endowed with the operator norm $||A||$. Let $\mathcal{H} = \mathcal{J}_2(\mathcal{K})$ be the Hilbert space of all Hilbert–Schmidt operators on $\mathcal{K}$ with scalar product

$$\langle u, v\rangle = \mathrm{Tr}(u^* v) \tag{5.3.1}$$

and norm $||u||_2 = \langle u,u\rangle^{1/2}$. We shall identify $\mathcal{M}$ with its left regular representation π_L on $\mathcal{H}$, $\pi_L(A)u = Au$, $A \in \mathcal{M}$, $u \in \mathcal{K}$.

There is a bijection between the space of faithful normal states φ on $\mathcal{M}$ and positive trace–class operators ϱ^2 on $\mathcal{K}$ with unit trace and densely defined inverse so that, denoting by ϱ the positive square root of ϱ^2, one has

$$\varphi(a) = \mathrm{Tr}\,[\varrho^2 A] = \langle \varrho, a\,\varrho\rangle.$$

Moreover, the modular operator Δ_φ associated with φ is given by $\Delta_\varphi = \varrho^2$, and the modular involution is the involution on $\mathcal{K}$ so that, on the dense set $\mathcal{M}\varrho$ one has $(JAJ)(B\varrho) = B\varrho A^*$.

Let also $\psi_n : n = 1,2,\ldots$ be the sequence of faithful normal states on $\mathcal{M}$ defined by

$$\psi_n(A) = \frac{\mathrm{Tr}\,(\exp[-H_n]A)}{\mathrm{Tr}\,(\exp[-H_n])} = \langle \Psi_n, A\Psi_n\rangle \tag{5.3.2}$$

where the cyclic and separating vectors Ψ_n in $\mathcal{H}$ are given by

$$\Psi_n = \frac{\exp[-H_n/2]}{(\mathrm{Tr}\,(\exp[-H_n]))^{1/2}}, \quad n = 1,2,\ldots \tag{5.3.3}$$

and where H_n are self–adjoint operators in $\mathcal{K}$ such that

$$\mathrm{Tr}(\exp[-H_n]) < \infty, \quad n = 1,2,\ldots. \tag{5.3.4}$$

In particular each H_n is bounded from below and has a discrete spectrum. Then for each $n = 1,2,\ldots$ there exists a complete orthonormal set $f_r^{(n)}$: $r = 0,1,\ldots$ in $\mathcal{K}$ and positive numbers $\varrho_r^{(n)}$: $r = 0,1,\ldots$ such that

$$\Psi_n^2 f_r^{(n)} = \varrho_r^{(n)} f_r^{(n)}, \quad r = 0,1,\ldots, \quad \text{with } \varrho_0^{(n)} \le \varrho_1^{(n)} \le \ldots. \tag{5.3.5}$$

Let $\tau_n : n = 1,2,\ldots$ be completely positive, identity preserving, ultraweakly continuous linear maps of $\mathcal{M}$ into itself such that

$$\psi_n(A\tau_n(B)) = \psi_n(\tau_n(A)B). \tag{5.3.6}$$

Define linear operators $e^{(n)}_{rr'}$ on $\mathcal{K}$ by $e^{(n)}_{rr'} f^{(n)}_s = \delta_{r's} f^{(n)}_r$. It follows from [42] that the τ_n may be written as

$$\tau_n(A) = \sum_{rr'ss'} c^{(n)}_{rr'ss'} e^{(n)}_{rr'} A e^{(n)}_{s's}, \tag{5.3.7}$$

where the series converges ultraweakly and where:

i) (positivity)

$$\sum_{rr'ss'} \overline{x}_{rr'} c^{(n)}_{rr'ss'} x_{ss'} \geq 0 \quad \forall \{x_{ss'}\}_{s,s' \in \mathbb{N}} \subset \mathbb{C} \tag{5.3.8}$$

whenever the series converges;
ii) (normalization)

$$\sum_{rr'ss'} c^{(n)}_{rr'ss'} e^{(n)}_{rr'} e^{(n)}_{ss'} \left(= \sum_{rr's} c^{(n)}_{rr'sr'} e^{(n)}_{rs} \right) = 1 \tag{5.3.9}$$

in the sense of ultraweak convergence;
iii) (detailed balance)

$$c^{(n)}_{rr'ss'} \varrho^{(n)}_s = c^{(n)}_{s'sr'r} \varrho^{(n)}_{s'} \quad (\Longleftrightarrow c^{(n)}_{rr'ss'} \varrho^{(n)}_r = c^{(n)}_{s'sr'r} \varrho^{(n)}_{r'}). \tag{5.3.10}$$

Example 5.14. Let us consider a classical Markov process Y_n with discrete state space X and with transition probabilities

$$P[Y_n = y | Y_{n-1} = x] = q_0(x,y) \exp[-\beta_n [U_n(y) - U_n(x)]_+], \tag{5.3.11}$$

with U_n such that $\sum_{y \in X} \exp[-\beta_n U_n] < \infty$ for all $n \in \mathbb{N}$. Let H_n be the (self–adjoint) closure of the densely defined operator determined by

$$H_n e^{(n)}_{xx} = \beta_n U_n(x) e^{(n)}_{xx} : \quad n \in X, \tag{5.3.12}$$

and let

$$c_{xx'yy'} = \delta_{xy} \delta_{x'y'} P[Y_n = x' | Y_{n-1} = x]. \tag{5.3.13}$$

Then all conditions *i)*–*iii)* are satisfied, with

$$\varrho^{(n)}_x = \frac{\exp[-\beta_n U_n(x)]}{\sum_{y \in X} \exp[-\beta_n U_n(y)]}. \tag{5.3.14}$$

Note that $e^{(n)}_{xx}$ may depend on n, so that the H_n may not commute with one another.

Now we apply the general scheme outlined in chapter 4 to the above situation. We know that there exist self–adjoint contractions T_n on $\mathcal{H}$ such that

$$T_n(A\Psi_n) = \tau_n(A)\Psi_n, \quad \forall A \in \mathcal{M}, n = 1, 2, \ldots. \tag{5.3.15}$$

Moreover, in the situation of the above Example, with X finite, it has been shown in [44] that, under the same assumptions leading to (5.2.16), a spectral gap for T_n exists and satisfies the bound

$$\Gamma_1 \exp[-\beta_n m_n^*] \leq \gamma_n \leq \Gamma_2 \exp[-\beta_n m_n^*], \tag{5.3.16}$$

where the positive constants Γ_1, Γ_2 depend only on q_0, whereas the positive constant m_n^* depends only on q_0 and on U_n (see the Appendix). We also refer to [46] and references quoted therein for estimates on the spectral gap for generators satisfying detailed balance in some models of infinite quantum systems.

We note also that in the case that dim $\mathcal{K} < \infty$ then there exists $u_1 \in \mathcal{H}$ such that $\varphi_1(a) = \langle u_1, a v_1 \rangle : a \in \mathcal{M}$ without further assumptions. Otherwise, a sufficient condition for this to hold is that $\varphi_1(A) = \mathrm{Tr}[\varrho_1 A]$ with ϱ_1 trace class and given by $\varrho_1 = \sum_{i=1}^{\infty} c_i P_{\phi_i}$ (P_ϕ denoting the orthogonal projection onto the subspace generated by the unit vector $\phi \in \mathcal{K}$) with $\phi_i \in \mathcal{D}(v_1^{-1})$ and $\sum_i c_i P_{\phi_i} v_1^{-1}$ convergent in $\mathcal{J}_2(\mathcal{K})$; then $u_1 = \sum_{i=1}^{\infty} c_i P_{\phi_i} v_i^{-1}$. However, if this is not the case in general it suffices to restrict to φ_0 in the dense set $S_{\psi_1}(\mathcal{M})$ (see (4.2.8, 4.2.9).

Furthermore, we have the following

Lemma 5.15. *Let R_n be the (densely defined) linear operator in $\mathcal{H}$ defined as in section 4.2 with Ψ_n as in* (5.3.3). *Suppose that, for all $t \in (0, 1/2)$, the operator*

$$V_n(t) = \exp[-tH_{n-1}](H_n - H_{n-1}) \exp[tH_{n-1}], \tag{5.3.17}$$

is densely defined and bounded with $||\Psi_n(t)|| \leq \alpha_n$, with $\alpha_n > 0$ independent of $t \in (0, 1/2)$. Then R_n is continuous.

Proof. Note that, if $u \in \mathcal{D}(R_n)$, we have

$$R_n u = u \left(\Psi_{n-1} \Psi_n^{-1}\right)^* . \tag{5.3.18}$$

Therefore, it suffices to prove that $\Psi_{n-1}\Psi_n^{-1} \in \mathcal{B}(\mathcal{K})$ for all $n = 1, 2, \ldots$. This happens if and only if $\exp[-H_{n-1}/2] \exp[H_n/2] \in \mathcal{B}(\mathcal{K})$ for all $n = 1, 2, \ldots$. Set now

$$U_n(t) = \exp[-tH_{n-1}] \exp[tH_n] \tag{5.3.19}$$

(densely defined on $\mathcal{D}(\exp[tH_n])$), for all $n = 1, 2, \ldots$. Since one has

$$U_n(t) = 1 + \int_0^t \exp[-sH_{n-1}](H_n - H_{n-1}) \exp[sH_n] ds,$$

it follows that

$$U_n(t) = 1 + \int_0^t \exp[-sH_{n-1}](H_n - H_{n-1}) \exp[sH_{n-1}] \exp[-sH_{n-1}] \exp[sH_n] ds$$

$$= 1 + \int_0^t V_n(s) U_n(s) ds. \tag{5.3.20}$$

Iterating (5.3.20) one finds that

$$U_n(t) = 1 + \sum_{k=1}^{\infty} \int_{0<s_1<\ldots<s_k<t} V_n(s_k)\ldots V_n(s_1)ds_1\ldots ds_k, \tag{5.3.21}$$

where the series in the r.h.s. converges uniformly by (5.3.17) for all $t \in [0, 1/2]$. In particular,

$$||U_n(t)|| \le \exp(t\alpha_n) \tag{5.3.22}$$

for all $t \in [0, 1/2]$, for all $n = 1, 2, \ldots$.

□

We now turn our attention to the inequalities (4.2.15), under which our conclusions follow. We need estimates on the spectral gap of T_n in the following form:

$$\gamma_n \ge \gamma n^{-m/c}, \tag{5.3.23}$$

where γ, m, c are positive constants with $c > m$. Indeed, in the situation of the above Example we have

$$\gamma_n \ge \gamma \exp[-\beta_n m_n^*], \tag{5.3.24}$$

where γ depends only on the matrix q_0. If, as in the usual formulation of simulated annealing, $\beta_n \le [\log n]/c + d$, and if $m_n^* \le m$ independent of n, we have (5.3.23).

Theorem 5.16. *Define, as in section 4.2, a_n and b_n through*

$$a_n = \gamma_n - \log ||R_n||, \tag{5.3.25}$$

$$b_n = e^{-\gamma_n} ||(R_n^* R_n - 1)\Psi_n||_2 \tag{5.3.26}$$

for $n = 1, 2, \ldots$, where γ_n and R_n are as above. Suppose also that (5.3.23) holds, and that

$$||V_n(t)|| \le \frac{k}{n^\eta}, \quad \forall t \in (0, 1/2) \tag{5.3.27}$$

for some positive constant k and for some η with $m/c < \eta \le 1$, where V_n is defined in (5.3.17) and m, c are as in (5.3.23). Then there exist constants $a, b > 0$ such that

$$a_n \ge a n^{\delta-1}, \quad b_n \le b n^{\varepsilon-1} \quad \forall n = 1, 2, \ldots, \tag{5.3.28}$$

with $\delta = 1 - m/c$ and for all $\varepsilon \in (1-\eta, 1-m/c)$. In particular one has, for all $A \in \mathcal{M}$ and for all $n = 1, 2, \ldots$,

$$|\varphi_n(A) - \psi_n(A)| \le (\text{constant}) ||A|| n^{\varepsilon-\delta} \to 0 \quad \text{as } n \to \infty. \tag{5.3.29}$$

Proof. We start estimating a_n. To this end, set

$$\tilde{U}_n(t) = \exp[-tH_n] \exp[tH_{n-1}] \; (= U_n(t)^{-1}), \tag{5.3.30}$$

densely defined on $\mathcal{D}(\exp[tH_{n-1}])$, for all $n = 1, 2, \ldots$, $t \in [0, 1]$. Proceeding as in the proof of Lemma 5.15 we find:

$$\tilde{U}_n(t) = 1 - \int_0^t \tilde{U}_n(s) V_n(s) ds. \tag{5.3.31}$$

Iterating (5.3.31) one finds

$$\tilde{U}_n(t) = 1 + \sum_{k=1}^{\infty} (-1)^k \int_{0<s_1<\ldots<s_k<t} V_n(s_1) \ldots V_n(s_k) \; :ds_1 \ldots ds_k. \tag{5.3.32}$$

Eqs. (5.3.21) and (5.3.32) can also be written in terms of time–ordered exponentials as follows:

$$U_n(t) = \overleftarrow{T} \exp\left(\int_0^t V_n(s) ds \right), \tag{5.3.33}$$

$$\tilde{U}_n(t) = \overrightarrow{T} \exp\left(- \int_0^t V_n(s) ds \right). \tag{5.3.34}$$

It follows from (5.3.17) and (5.3.34) that, proceeding exactly as in Lemma 5.15, $||\tilde{U}_n(t)|| \leq \exp[tk/n^\eta]$ for all $t \in [0, 1/2]$. Finally, Lemma 4.1 shows that

$$R_n = ||U_n(1/2)\Psi_n||_2 || J\tilde{U}(1/2) J, \tag{5.3.35}$$

so that

$$||R_n|| \leq ||U_n(1/2)|| \; ||\tilde{U}_n(1/2)||^{1/2}, \tag{5.3.36}$$

and the previous estimates on the norms in the r.h.s. of (5.3.36) imply that there exists $A > 0$ such that $\log ||R_n|| \leq A/n^\eta$. Since the assumptions on the behaviour of the spectral gap γ_n and (5.25) imply in turn that, for some constant $B > 0$,

$$\gamma_n \geq B n^{-m/c}, \tag{5.3.37}$$

it follows that the first part of (5.3.28) holds with $\delta = 1 - m/c$.
In order to prove the second part of (5.3.28), it suffices to show that there exists a constant $C > 0$ such that

$$||(R_n^* R_n - 1)\Psi_n||_2 \leq \frac{C}{n^\eta}. \tag{5.3.38}$$

Hence, one computes

$$||(R_n^* R_n - 1)\Psi_n||_2^2 = ||R_n^* \Psi_{n-1} - \Psi_n||_2^2$$

$$||R_n^* \Psi_{n-1}||_2^2 + 1 - 2 \operatorname{Re}\langle R_n^* \Psi_{n-1}, \Psi_n \rangle = ||R_n^* \Psi_{n-1}||_2^2 - 1$$

$$\leq ||R_n^*|| - 1 \leq \text{ const. } (\exp[1/n^\eta] - 1)$$

by the previous step, and this implies that (5.3.38) holds.

Thus, the thesis follows from Lemma 3.16.

□

Corollary 5.17. *If ψ_n has a weak limit ψ_∞ as $n \to \infty$, then φ_n converges weakly to ψ_∞ as $n \to \infty$.*

Remark 5.18. If $H_n = \beta_n H$ with $\beta_n \to \infty$ as $n \to \infty$, then ψ_∞ exists and is the projection onto the subspace corresponding to the smallest eigenvalue of H. More generally, if H_n has the form (5.3.12), with $\beta_n \to \infty$ and $U_n(x) \to U(x)$ for all $x \in X$ sufficiently fast so that $\beta_n(U_n(x) - U(x)) \to 0$ as $n \to \infty$ uniformly in x, the same conclusion holds.

Appendix:

THE SPECTRAL GAP

This appendix is a (partial) review concerning some of the estimates which have been repeatedly used in the course of this thesis concerning the asymptotic behaviour of the spectral gap of the generators of the processes considered, and therefore concerning the identification of the critical constant in the optimal schedule for simulated annealing. This review is far from complete, and in particular we have chosen to discuss only the case of the generator of simulated annealing on finite sets [6] and its non–commutative generalization [44], to be dealt with in section A.1, and of Langevin diffusions on $\mathbb{R}^n$ [16, 17], which we consider in section A.2.

A.1. Simulated annealing on compact sets and non–commutative generalizations

Let X be a finite set with $|X|$ points, and μ_0 be a probability measure on X charging any point. Let q_0 be an irreducible transition matrix on X, which in addition is μ_0–reversible in the sense that

$$\alpha(x,y) := \mu_0(x)q_0(x,y) = \alpha(y,x). \tag{A.1.1}$$

The simplest example is that in which μ_0 is the normalized counting measure and q_0 is a symmetric matrix.
Let U be a nonnegative function on X; define the Gibbs measures μ_β at inverse temperature β by

$$\mu_\beta = Z_\beta^{-1}e^{-\beta U}\mu_0, \tag{A.1.2}$$

Z_β being the normalization constant. The transition probabilities at inverse temperature β are defined by

$$q_\beta(x,y) = \exp[-\beta[U(y)-U(x)]_+]q_0(x,y) \quad \text{if } y \neq x, \tag{A.1.3}$$

and by normalization if $y = x$. Define also linear operator L_β on $\mathbb{C}^{|X|}$ by

$$L_\beta f(x) = \sum_{y\in X}(f(y)-f(x))q_\beta(x,y), \quad x \in X. \tag{A.1.4}$$

Consider L_β as acting on the real Hilbert space $\mathcal{L}^2(\mu_\beta)$. Then its bilinear form Q_β is defined by

$$Q_\beta(f,g) = -\int fL_\beta g d\mu_\beta, \tag{A.1.5}$$

and is easily seen to be given by

$$Q_\beta(f,g) = (2Z_\beta)^{-1} \sum_{x,y \in X} e^{-\beta(U(x) \vee U(y))}(f(x) - f(y))(g(x) - g(y))\alpha(x,y). \qquad (A.1.6)$$

Note that $-L_\beta$ is a nonnegative self–adjoint operator, with $\lambda = 0$ as lowest eigenvalue with unique normalized eigenvector $f(x) = 1$ for all x. Let $||\cdot||_2$ denote the norm in $\mathcal{L}^2(\mu_\beta)$, and $\langle f \rangle_\beta = \int f d\mu_\beta$; we are interested in determining the behaviour for large β of the second eigenvalue $\gamma(\beta)$ of $-L_\beta$, which is characterized as

$$\gamma(\beta) = \inf\{Q_\beta(f,f) \mid ||f||_2 = 1, \langle f \rangle_\beta = 0\}. \qquad (A.1.7)$$

Equivalently, let

$$\mathrm{Var}_\beta(f) = \mu_\beta[(f - \langle f \rangle_\beta)^2]. \qquad (A.1.8)$$

Then

$$\gamma(\beta) = \inf\{-\mathrm{Var}_\beta^{-1} \sum f(x) L_\beta f(x) \mu_\beta(x) \ : \ f \in \mathbb{C}^{|X|}\}. \qquad (A.1.9)$$

We define a nonnegative number m, depending only on U, as follows. A path from $x \in X$ to $y \in X$ is a finite sequence $x_0 = x, x_1, \dots, x_n = y$ such that $q_0(x_i, x_{i+1}) > 0$ for all $i = 0, \dots, n-1$. Let $P_{x,y}$ denote the set of paths from x to y, and let a path $p = \{p_i\}_{i=0,\dots,n-1}$ be given. Define the elevation of p by $\mathrm{Elev}(p) = \max\{U(p_i), p_i \in p\}$, and the lowest possible elevation when travelling from x to y as $H(x,y) = \min\{\mathrm{Elev}(p) : p \in P_{x,y}\}$. Finally, let

$$m = \max\{H(x,y) - U(x) - U(y) : x,y \in X\}. \qquad (A.1.10)$$

It is easy to realize that, if x_0, y_0 are a couple of points at which the above maximum is attained, then at least one of them is an absolute minimum of U. The number m can thus be described as the maximum "energy height" which must be reached in order to reach a global minimum along a way which takes the lowest passes.

We shall prove in the sequel, following [6] that $\gamma(\beta) \sim \exp[-\beta m]$ as $\beta \to \infty$.

Theorem A.1. *There exist constants c, C with $0 < c \leq C < \infty$, independent of β, such that, for all $\beta \geq 0$,*

$$c\, e^{-\beta m} \leq \gamma(\beta) \leq C\, e^{-\beta m}. \qquad (A.1.11)$$

Proof. i) (upper bound).
First assume $m > 0$. Choose x_0, y_0 so that $H(x_0, y_0) - U(x_0) - U(y_0) = m$. Let$A = \{y \in X \mid H(y_0, y) < H(y_0, x_0)\}$. Then $x_0 \in A$, $x_0 \notin A$, since $m > 0$. Moreover, if $x \in A$, $y \notin A$, and $q_0(x,y) > 0$, then $U(y) > H(y_0, x_0)$, since it suffice to consider the path $p \in P_{y_0,x}$ for which $H(y_0, x) = \mathrm{Elev}(p)$, and extend it by one further step to y. This implies that for all such x, y and for all $\beta > 0$, one has $q_\beta(x,y)\mu_\beta(x) \leq \alpha(x,y) Z_\beta^{-1} \exp[-\beta H(y_0, x_0)]$ so that, letting $F(x) = I_A(x)$, one has

$$-\sum_{x,y} F(x) L_\beta F(x) \mu_\beta(x) = \frac{1}{2} \sum q_\beta(x,y)(F(y) - F(x))^2 \mu_\beta(x) \leq$$

$$\leq Z_\beta^{-1} \sum_{x \in A} \sum_{y \notin A} \alpha(x,y) e^{-\beta H(y_0,x_0)}. \tag{A.1.12}$$

Finally,

$$\mathrm{Var}_\beta(F) \geq Z_\beta^{-2} \mu_0(x_0)\mu_0(y_0) e^{-\beta(U(x_0)+U(y_0))}, \tag{A.1.13}$$

so that, for some $C > 0$ independent of β,

$$-\mathrm{Var}_\beta^{-1}(F) \sum F(x) L_\beta F(x) \mu_\beta(x) \leq C\, e^{-\beta m}. \tag{A.1.14}$$

If $m = 0$, the above argument holds upon substituting F with the indicator function of the set of the absolute maxima of U.
ii) (lower bound).
We shall prove that

$$-\mathrm{Var}_\beta^{-1}(f) \sum f(x) L_\beta f(x) \mu_\beta(x) \geq c\, e^{-\beta m}, \tag{A.1.15}$$

for all $f \in \mathbb{C}^{|X|}$, for all $\beta \geq 0$ and for a suitable constant $c > 0$.
Fix $x, y \in X$, and choose a path $p \in P_{x,y}$ such that $\mathrm{Elev}(p) = H(x,y)$. Let $n(x,y)$ be the length of p and define a positive number N as $N = \max_{x,y} n(x,y) (< \infty)$. Take $z, w \in X$ and define $\chi_{z,w}(x,y) = 1$ whenever for some index i one has $p_i^{x,y} = z$, $p_{i+1}^{x,y} = w$, and $\chi_{z,w} = 0$ otherwise. If $\alpha(z,w) = 0$, then $\chi_{z,w}$ is identically zero, and in the sequel we intend $\chi_{z,w}(x,y)/\alpha(z,w) = 0$ whenever both terms vanish. Then we compute, for all $f \in \mathbb{C}^{|X|}$,

$$2\mathrm{Var}_\beta(f) = \sum_{x,y} (f(y) - f(x))^2 \mu_\beta(x)\mu_\beta(y) =$$

$$= \sum_{x,y} \left(\sum_{i=1}^{n(x,y)} f(p_i^{x,y}) - f(p_{i-1}^{x,y}) \right)^2 \mu_\beta(x)\mu_\beta(y) \leq$$

$$\leq \sum_{x,y} n(x,y) \sum_{i=1}^{n(x,y)} (f(p_i^{x,y}) - f(p_{i-1}^{x,y}))^2 \mu_\beta(x)\mu_\beta(y) \leq$$

$$\leq N \sum_{x,y,z,w} \chi_{z,w}(x,y)(f(z) - f(w))^2 \alpha(z,w) e^{-\beta(U(z)\vee U(w))} \frac{\mu_\beta(x)\mu_\beta(y)}{e^{-\beta(U(z)\vee U(w))}} \leq$$

$$\leq 2N \left(\max_{z,w} \sum_{x,y} \chi_{z,w}(x,y) \frac{\mu_\beta(x)\mu_\beta(y)}{e^{-\beta(U(z)\vee U(w))}} \right) Q_\beta(f,f). \tag{A.1.16}$$

Besides,

$$\chi_{z,w}(x,y) \frac{\mu_\beta(x)\mu_\beta(y)}{e^{-\beta(U(z)\vee U(w))}} = Z_\beta^{-1} \chi_{z,w}(x,y) \frac{\mu_0(x)\mu_0(y)}{\alpha(z,w} e^{-\beta(U(z)\vee U(w) - U(x) - U(y))} \leq$$

$$\leq e^{\beta m} \chi_{z,w}(x,y) \frac{\mu_0(x)\mu_0(y)}{\alpha(z,w) \sum_{v:U(v)=0} \mu_0(v)}$$

which yields the claim.

□

We now give a short sketch, following [44], of a non–commutative generalization of the above results, which we have already mentioned in section 5.3. Let X, U, μ_β, q_β and L_β be as above. Let $e_x : x \in X$ denote the canonical orthonormal basis of $\mathbb{C}^{|X|}$. Consider the von Neumann algebra $\mathcal{M} = M(|X|, \mathbb{C})$ which is generated, as a linear space, by the matrices e_{xy} defined by $e_{x,y}e_z = \delta_{yz}e_x$. Consider the state on $\mathcal{M}$ given by $\mu(f) = \langle \Phi, f\Phi \rangle$, where the scalar product is given by $\langle f, g \rangle = \mathrm{Tr}(f^* g)$, and $\Phi = N^{-1/2} I$. The corresponding GNS representation acts on the Hilbert space $\mathcal{K} = M(|X|, \mathbb{C})$ with the above trace scalar product.

We define a bijection D between $\mathbb{C}^{|X|}$ and the set of diagonal matrices by $D(f) = \sum_x f(x)e_{xx}$, and a Hamiltonian H in $\mathcal{M}$ as $H = D(U)$. The family of thermal states on $\mathbb{M}$ at inverse temperature β and corresponding to H is thus defined to be

$$\mu_\beta^q(a) = \mathrm{Tr}[\exp(-\beta H)A]/\mathrm{Tr}[\exp(-\beta H)], \quad a \in \mathcal{M}. \tag{A.1.17}$$

Now we construct a generator L_β^q satisfying detailed balance w.r.t. μ_β^q, and study some of its spectral properties. To this end, let $\mathcal{B}$ be the set of those $x, y \in X$ such that $U(x) < U(y)$ and $q_0(x,y) > 0$. For $j = (x,y) \in \mathcal{B}$ let

$$v_j = q_\beta(y,x)^{1/2} e_{xy}. \tag{A.1.18}$$

Last, define

$$L_\beta^q(a) = \sum_{j \in \mathcal{B}} \left((v_j^* a v_j - \frac{1}{2}[v_j^* v_j, a]_+) + \exp[-\beta w_j](v_j a v_j^* - \frac{1}{2}[v_j v_j^*, a]_+) \right), \quad a \in \mathcal{M}, \tag{A.1.19}$$

where $[a,b]_+ = ab + ba$, and $w_j = U(y) - U(x)$. Then it is possible to prove that L_β^q satisfies detailed balance with respect to μ_β^q for all β. Using the detailed balance condition for μ_β, one proves (cf. [15]) the following:

Lemma A.2. *For all $f \in \mathbb{C}^{|X|}$, one has*

$$L_\beta^q(D(f)) = D(L_\beta(f)). \tag{A.1.20}$$

To each L_β^q is associated a self–adjoint contraction S_β^q on $\mathcal{K}$; it is easy to prove that it is given by $S_\beta^q = S_{\beta,0}^q - S_{\beta,1}^q$, where

$$S_{\beta,0}^q f = \frac{1}{2} \sum_{x \neq y} q_\beta(x,y)[e_{xx}, f]_+; \tag{A.1.21}$$

$$S^q_{\beta,1} = \sum_{j=(x,y)\in B} (\exp[-\beta w_j/2] q_\beta(x,y)(e_{yx} f e_{xy} + e_{xy} f e_{yx})) \,. \tag{A.1.22}$$

Then S_β has a gap $\gamma^q(\beta)$ between $\lambda = 0$ and the rest of its spectrum. In order to estimate its behaviour for large β, recall that we have proved in Theorem A.1 that there exists constants c, C, m such that (A.1.11) holds. Let l be the maximum energy gap which can be attained in one step:

$$l = \max\{U(y) - U(x) \;:\; x, y \in X, q_0(x,y) > 0\}. \tag{A.1.23}$$

Define positive constants m_q, c_q by $m_q = \max\{m, l\}$, $c_q = \min\{c, \alpha\}$, where α is such that $q_0(x,y) > \alpha$ for any couple (x,y) with $q_0(x,y) > 0$. Then we have the following

Theorem A.3. *The spectral gap for S^q_β satisfies the bounds*

$$c_q e^{-\beta m_q} \le \gamma^q(\beta) \le C\, e^{-\beta m_q}. \tag{A.1.24}$$

Proof. Let $\mathcal{K}_1 \subset \mathcal{K}$ be the subspace of diagonal matrices, and P the orthogonal projection onto it. Then

$$S^q_{\beta,0} P = P S^q_{\beta,0}, \quad S^q_{\beta,1} P = S^q_{\beta,1} = P S^q_{\beta,1}. \tag{A.1.25}$$

This implies that $\mathcal{K}_1$ and its orthogonal complement $\mathcal{K}_2$ are invariant under $S^q_{\beta,0}, S^q_{\beta,1}$. Moreover, $S^q_\beta|_{\mathcal{K}_1} = S_\beta$, the self–adjoint contraction associated with $\mathcal{L}_\beta$, and $S^q_\beta|_{\mathcal{K}_2} = S^q_{\beta,0}$. In turn, $S^q_{\beta,0}$ satisfies the bound

$$S^q_{\beta,0} \ge \min_{x\in X} \left\{ \sum_{y\neq x} q_\beta(x,y) \right\} \ge$$

$$\ge \min_{x\in X} \left\{ e^{-\beta l} \sum_{y\neq x} \max\{q_\beta(x,y), q_\beta(y,x)\} \right) \ge e^{-\beta l}\alpha, \tag{A.1.26}$$

where the latter inequality follows from irreducibility, since for each x there must exist at least one y with $q_0(x,y) \neq 0$, and from the fact that $q_\beta(x,y)$ is monotonically nonincreasing as a function of β for any fixed couple (x,y).

□.

A.2. Langevin diffusions in $\mathbb{R}^d$

We now turn to proving some results similar to those proved in the previous section for the operator $L_\beta = \beta S_\beta$, $S_\beta = \beta^{-1}\Delta - \nabla U\cdot$ being the generator of a Langevin diffusion with drift $b = -\nabla U$ and diffusion matrix $a = \sqrt{2/\beta}\, I$. We follow the discussion in [16, 17], and note that in comparison to that paper our operator L_β is β times the operator L_ε considered there, upon defining $\beta = 2/\varepsilon^2$. The assumption in which the calculation below hold are even more general then those which were used in section 1. The notation will be the same of that section.

Assumption A.4. $U : \mathbb{R}^n \to [0,\infty)$ is a function of class C^2 with $U, |\nabla U| \to \infty$ as $|x| \to \infty$, and moreover such that $|\nabla U| - \Delta U$ is bounded below. Without loss of generality, we assume that the set of critical points of U is contained in a ball $B(0, R_0)$ on which $|\nabla U| \leq 1$.

It is well known that, under the above assumption, L_β is essentially self–adjoint on C_0^∞ as an operator acting in $\mathrm{L}^2(\mu_\beta)$. Moreover, $-L_\beta$ is nonnegative and there is a gap $\gamma(\beta)$ between $\lambda = 0$ and the rest of its L^2–spectrum.

We define a positive number m exactly in the same way in which the corresponding number was defined in section A.1, with curves of class C^1 taking the place of paths. We shall prove the following

Theorem A.5. *The following equality holds:*

$$\lim_{\beta\to\infty} \frac{\log(\gamma(\beta))}{\beta} = -m. \tag{A.2.1}$$

The proof of this theorem is rather involved, and we shall divide it into several step; some of the statements will not be proved in full detail, and in such cases we refer to [16, 17] for a complete discussion.

First, an upper bound on the quantity $\beta^{-1}\log(\gamma(\beta))$ can be given as in [12], whose arguments have been already sketched in Theorem A.1. In fact, the following Lemma holds:

Proposition A.6. *One has*

$$\limsup_{\beta\to\infty} \frac{\log(\gamma(\beta))}{\beta} \leq -m. \tag{A.2.2}$$

Proof (sketch). Choose x_0, y_0 (y_0 being an absolute minimum of U) such that $m = H(y_0, x_0) - U(x_0)$, the infimum being attained due to the fact that the absolute minima of U are contained in $B(0, R_0)$ by Assumption A.4. Let ε be sufficiently small (see below), and let $A_\varepsilon = \{x \in \mathbb{R}^d \mid H(y_0, x) < H(y_0, x_0) - 3\varepsilon\}$. Then, for large β, $B(y_0, 2\varepsilon) \subset A_\varepsilon$, $B(x_0, 2\varepsilon) \not\subset A_\varepsilon$. It is possible to construct a function ψ_ε such that $\psi_\varepsilon(z) = 1$ if $B(z, 1/\beta) \subset A_\beta$, and which vanishes if $B(z, 1/\beta) \subset A_\beta^c$. The function ψ_β

can be chosen so that $|\nabla\psi_\varepsilon|^2 \leq K\beta^{-2d-2}$, and it can be shown by computation that the quadratic form Q_β of L_β evaluated on ψ_ε satisfies

$$Q_\beta(\psi_\varepsilon,\psi_\varepsilon) \leq K\, Z_\beta^{-1}\varepsilon^{-2d-2}\exp\left[-\beta(H(y_0,x_0)-5\varepsilon\right],$$

whence

$$\mathrm{var}_\beta(\psi_\varepsilon) \geq Z_\beta^{-2}e^{-\beta U(x_0)+2\varepsilon)}\mu_{\mathrm{Leb}}(B(x_0,\varepsilon))\mu_{\mathrm{Leb}}(B(y_0,\varepsilon)).$$

The thesis follows by choosing $\varepsilon = 1/\beta$.

□

What follows will aim at proving a lower bound for $\beta^{-1}\log(\gamma(\beta))$; Assumption A.4 is assumed throughout. We say that U is strictly convex in a neighbourhood of infinity if there exists $R_1 > 0, C > 0$ such that $D^2U \geq C^{-1}I$ on $B(0,R_1)^c$.

Proposition A.7. *One has*

$$\liminf_{\beta\to\infty}\frac{\log(\gamma(\beta))}{\beta} \geq -m. \tag{A.2.3}$$

The above Proposition will be proved by showing that that there exists a polinomial P such that, at least for β large enough,

$$\mathrm{var}_\beta(f) \leq P(\beta)\exp[\beta m]Q_\beta(f,f), \tag{A.2.4}$$

for all $f \in H^1(\mathbb{R}^d,\mu_\beta)$. In fact, it suffice to prove (A.2.4) for $f \in \mathcal{D}(\mathbb{R}^d)$. Note that

$$2\mathrm{var}_\beta(f) = \int (f(x)-f(y))^2\,d\mu_\beta(x)d\mu_\beta(y). \tag{A.2.5}$$

We shall estimate the integral on the r.h.s. of (A.2.5) extended over $D_1 = [B(0,R_1+r_1)^c]^2$, $D_2 = [B(0,R_1+r_2)]^2$, $D_3 = B(0,R_1+r_1)\times B(0,R_1+r_2)^c \cup B(0,R_1+r_2)^c \times B(0,R_1+r_1)$, where $r_{1,2} > 0$ will be determined below. $I_j^\beta(f)$ will denote the integral on the r.h.s. of (A.2.5) extended to $D_j, j = 1,2,3$. We shall give sketchy proofs of the estimates concerning $I_{1,2}$, referring to [16, 17] for the (lengthy) estimation of I_3, whose result we shall quote below.

Lemma A.8. *There exists $r_1 > 0$ such that*

$$I_1^\beta(f) \leq C\beta^{-1}Q_\beta(f,f) + \beta^{-1}\,\mathrm{var}_\beta(f). \tag{A.2.6}$$

Proof. Choose $r_1 > 0$ large enough to be allowed to take a function κ of class C^1 with $0 \leq \kappa \leq 1$, $|\nabla k| < C^{-1}$ and $k = 1$ on $B(0,R_1+r_1)^c$, $k = 0$ on $B(0,R_1)$. Then κf is of class C^1 with support in $B(0,R_1)^c$, and we can suppose without loss of generality that its mean w.r.t. μ_β is zero. It is easy to realize that

$$I_1^\beta(f) \leq \int (\kappa f(x)-\kappa f(y))^2\,d\mu_\beta(x)d\mu_\beta(y) = I^\beta(\kappa f).$$

By an inequality due to Brascamp and Lieb, and by the convexity of U in a neighbourhood of infinity, one has

$$I^{\beta}(\kappa f) \leq (2\beta)^{-1} C \int |\nabla(\kappa f)|^2 d\mu_\beta. \tag{A.2.7}$$

Therefore the following inequality holds:

$$I^{\beta}(\kappa f) \leq C\beta^{-1}\left(\int f^2 |\nabla \kappa|^2 d\mu_\beta + \int |\nabla f|^2 \kappa^2 d\mu_\beta\right) \leq \beta^{-1}\mathrm{var}_\beta(f) + C\beta^{-1} Q_\beta(f,f).$$

□

In the previous Lemma, the value of m does not appear. It will indeed appear only in the estimation of I_2^β which we sketch below, and which of course follows the same lines of the proof of [12] for the case of compact manifolds.

Lemma A.9. *There exist positive constants K_1, K_2 such that*

$$I_2^\beta(f) \leq K_1 \beta^{3d}(1 + K_2 \beta^{d+1}) e^{\beta m} Q_\beta(f,f). \tag{A.2.8}$$

Proof. Since we have taken $R_1 > R_0$, we can suppose that m is calculated restricting to $B(0, R_1 + r_1)$. We assume without loss of generality that $|\nabla U| < 1$ on $B(0, R_1 + r_1)$. There is a covering of D_2 by $N_\beta = k_1 \beta^{2d}$ product of balls of radius $1/\beta$, k_1 independent of β. Set

$$\Lambda(\beta, f) = \sup\{\exp[-\beta(U(x) + U(y)] \times$$

$$\times \int_{B(x,1/\beta)\times B(y,1/\beta)} (f(v) - f(w))^2 \, dv dw \; ; \; (x,y) \in D_2\}. \tag{A.2.9}$$

Then $I_2^\beta(f) \leq Z_\beta^2 N_\beta \Lambda(\beta, f)$. Take, for each couple (x,y) and for all β, a curve of class C^1, say $\gamma_{\beta,x,y}$, which joins x and y and whose elevation is smaller than or equal to $m + U(x) + U(y) + 1/\beta$. We say that such a curve is "almost optimal" (a.o. in the sequel). Such a curve can be chosen so that, letting

$$L(\beta, x, y) = \inf\left\{\left(\int_0^1 |\dot\gamma(t)|^2 dt\right)^{1/2} : \gamma \in C^1 \text{ joining } x \text{ and } y, \gamma \text{ a.o.}\right\}, \tag{A.2.10}$$

and

$$L(\beta) = \sup\left\{L(\beta, x, y), (x,y) \in D_2\right\}, \tag{A.2.11}$$

one has

$$\sup\left\{|\dot\gamma_{\beta,x,y}(t)|, t \in [0,1]\right\} \leq L(\beta) + \frac{1}{2}. \tag{A.2.12}$$

Let $(\varsigma, \eta) \in [B(0, 1/\beta)]^2$, and consider the curve joining $x + \varsigma$ to $y + \eta$ defined by $\Gamma(t, \varsigma, \eta) = \gamma_{\beta,x,y}(t) + (1-t)\varsigma + t\eta$. Then the elevation of Γ is smaller than or equal to $m + U(x) + U(y) + 2\varepsilon$ and, moreover, $|\dot\Gamma| \leq L(\varepsilon) + 1$. Therefore

$$\int_{B(x,1/\beta)\times B(y,1/\beta)} (f(v) - f(w))^2 \, dv \, dw =$$

$$= \int_{[B(0,1/\beta)]^2} d\varsigma d\eta \int_0^1 |\nabla f(\Gamma(t,\varsigma,\eta))|^2 |\dot{\Gamma}(t,\varsigma,\eta)|^2 dt \leq$$
$$\leq (L(\beta)+1)^2 \exp\left[\beta(m+U(x)+U(y)+1/\beta)\right] \times$$
$$\times \int_0^1 dt \int_{(\varsigma\vee\eta)\leq 1/\beta} |\nabla f(\Gamma(t,\varsigma,\eta))|^2 \exp\left[-\beta U(\Gamma(t,\varsigma,\eta)\right] d\varsigma d\eta.$$

Finally a change of variables in the integral in the r.h.s. of the last inequality shows that

$$\Lambda(\beta,f) \leq 2^{-d} Z_\beta^{-1}(L(\varepsilon)+1)e^{\beta m} Q_\beta(f,f). \tag{A.2.13}$$

The claim follows from the estimate $L(\beta) \leq K\beta^{d+1}$, which follows as in [12] by constructing a piecewise C^1 curve γ joining x and y and almost optimal, with the property that $(\int_0^1 |\dot{\gamma}(t)|^2 dt)^{1/2} \leq K\beta^{d+1}$.

□

The proof of the following Lemma can be found in [17, pg 11–15].

Lemma A.10. *The following inequality holds:*

$$I_3^\beta(f) \leq \frac{2}{\beta C^{-1} - 2(d-1)} Q_\beta(f,f). \tag{A.2.14}$$

Corollary A.11. *If U is strictly convex in a neighbourhood of infinity, then*

$$\lim_{\beta\to\infty} \frac{\log(\gamma(\beta))}{\beta} = -m. \tag{A.2.15}$$

We want to sketch the proof that Corollary A.11 still holds even in the case in which U is not convex in a neighbourhood of infinity. The main technical argument is the following one:

Lemma A.12. *Under Assumption* A.4, *There exists $R_1 > R_0$ and $\hat{U} \in C^2(\mathbb{R}^d, \mathbb{R})$ such that $\hat{U} \geq U$ and $\hat{U}$ coincides with U in $B(0,R_0)$, and such that $D^2\hat{U} > C^{-1}I$ on $B(0,R_1)^c$. Moreover, $\hat{U}$ can be chosen so that its corresponding maximal elevation $\hat{m}$ equals m.*

Proof (sketch). Let $V_1(x) = V_1(|x|) = \sup\{U(x) : |x| \leq r\}$. Then V_1 is strictly convex and satisfies $V_1 \geq U$. Moreover, take $V = V_1 + c^{-1}|x|^2$.

To construct the function $\hat{U}$ we shall need some more definitions; let $K(s) = \{x : U(x) \leq s\}$, and choose s so that $B(0,R_0) \subset K(s)$. Take also $\varepsilon > 0$ sufficiently small, so that $|V(x) - V(y)| < m/2$ whenever $|x-y| < \varepsilon$ and $x, y \in K(s+\varepsilon)$ (this will be used in the part of the proof which we do not report).

Consider a function $\varrho \in C^2(\mathbb{R}^d)$ given by $\varrho(x) = \hat{\varrho}(U(x))$, where $\hat{\varrho} : \mathbb{R} \to \mathbb{R}$ is monotonically increasing, and choose $\hat{\varrho}$ so that $\varrho = 0$ on $K(s)$, and $\varrho = 1$ on $K(s+\varepsilon)^c$. We are now ready to set $\hat{U} = (1-\varrho)U + \rho V$. It is then possible to see that, x_0 being an absolute minimum and setting

$$m_1 = \sup\left\{H(x,x_0) - \hat{U}(x) \ : \ x \in K(s)^c \cap K(s+\varepsilon)\right\}, \tag{A.2.16}$$

one has $\hat{m} = m \vee m_1$. It is then a technical matter (cf. [16, pg. 17] to verify that $m_1 < m/2$, so that $\hat{m} = m$.

$\square$

The following step is an application of the Feynman–Kac formula, which yields a pointwise bound on the (normalized) eigenfunction $u_2(\beta)$ corresponding to the second eigenvalue of $-L_\beta$ (cf. [76, 35]).

Lemma A.13. *With the above notation,*

$$|u_2(\beta)(x)| \leq c_1 Z_\beta^{-1/2} \exp\left[\beta(U(x) - c_2|x|)\right] \tag{A.2.17}$$

for suitable positive constants c_1, c_2.

This fact can be used very easily to prove that, denoting by $\hat{\mathrm{var}}_\beta$ the variance corresponding to $\hat{U}$, one has

$$\hat{\mathrm{var}}_\beta(u_2(\beta)) \geq \frac{1}{4}. \tag{A.2.18}$$

Moreover it is also easy to prove that $Z_\beta/\hat{Z}_\beta \geq 1/2$, and that $H^1(\mu_\beta) \subset H^1(\hat{\mu}_\beta)$, the quantities with the overline " ˆ " referring to $\hat{U}$.

Theorem A.14. *Under Assumption* A.4,

$$\lim_{\beta\to\infty} \frac{\log\gamma(\beta))}{\beta} = -m. \tag{A.2.19}$$

Proof. We have to prove a lower bound of the form

$$\liminf_{\beta\to\infty} \frac{\log\gamma(\beta))}{\beta} \geq -m. \tag{A.2.20}$$

If (A.2.20) fails, there exist $v > 0$ and a sequence $\beta_n \to \infty$ such that $\beta_n^{-1}\log\gamma(\beta_n) < -(m+v)$. By the above computations for strictly convex potentials and the remarks preceding the statement of this Theorem, we compute

$$\gamma(\beta_n)Q_\beta(u_2^{\beta_n}, u_2^{\beta_n}) \geq \frac{Z_{\beta_n}}{\hat{Z}_{\beta_n}}\hat{Q}_{\beta_n}(u_2^{\beta_n}, u_2^{\beta_n}) \geq P(\beta_n)^{-1}e^{-\beta_n m}\hat{\mathrm{var}}_{\beta_n}(u_2(\beta_n)),$$

for some polinomial P. Hence, since $\gamma(\beta_n) < e^{-\beta_n(m+v)}$, we have:

$$\frac{1}{4}e^{-\beta_n m}P(\beta_n)^{-1} \leq e^{-\beta_n(m+v)},$$

which yields a contradiction in the limit as $n \to \infty$, since a careful look at the above proofs shows that P can be chosen so that $\lim_{\beta\to\infty} P(\beta) \neq 0$.

$\square$

References

[1] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, *Science* **220** : 621–680 (1983).

[2] S. Geman, D. Geman, Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.* **6** : 721–741 (1984).

[3] G. E. Hinton, T. J. Sejnowski, D. H. Ackley, Boltzmann machines: constraint satisfaction machines that learn, Department of Computer Science Technical Report CMU–CS–84–119, Carnegie-Mellon University, Pittsburgh, Pennsylvania, 1984.

[4] H. J. Kushner, Asymptotic global behaviour for stochastic approximations and diffusions with slowly decreasing noise effects: global minimization via Monte Carlo, Lefschetz Center for Dynamical Systems report 85–7, Div. Applied Mathematics, Brown University, Providence, RI, 1985.

[5] G. Parisi, Prolegomena to any further computer evaluation of the QCD mass spectrum, in: Progress in Gauge Field Theory, Cargese, 1983.

[6] R. A. Holley, D. W. Stroock, Simulated annealing via Sobolev inequalities, *Commun. Math. Phys.* **115** : 553–569 (1988).

[7] B. Gidas, Non-stationary Markov chains and convergence of the annealing algorithm, *J. Stat. Phys.* **39** : 73–131 (1985).

[8] B. Hajek, Cooling schedules for optimal annealing, *Math. Oper. Res.* **13** : 311–329 (1988).

[9] B. Gidas, Global optimization via the Langevin equation, in: Proceedings of the 24th conference on decision and control, 774–778 (1985).

[10] C. R. Hwang, Laplace's method revisited: weak convergence of probability measures, *Annals Prob.* **8** : 1177–1182 (1980).

[11] G. Royer, A remark on simulated annealing of diffusion process, *SIAM J. Control Optimization* **27** : 737–753 (1987).

[12] R. A. Holley, S. Kusuoka, D. W. Stroock, Asymptotics of the spectral gap with application to the theory of simulated annealing, *J. Funct. Anal.* **83** : 333–347 (1989).

[13] T. S. Chiang, C. R. Hwang, S. J. Sheu, Diffusions for global optimization in $\mathbb{R}^n$, *SIAM J. Control Optimization* **25** : 737–753 (1987).

[14] S. Geman, C. R. Hwang, Diffusions for global optimization, *SIAM J. Control Optimization* **24** : 1031–1043 (1986).

[15] C. R. Hwang, S. J. Sheu, Large time behaviours of perturbed diffusion Markov processes with applications (II): the asymptotic behaviour of the second eigenvalue of perturbed Fokker–Planck operators, preprint (1991).

[16] S. Jacquot, Comportement asymptotique de le seconde valeur propre des processus de Kolmogorov, *J. Multiv. Anal.* **40** : 335–347 (1991).

[17] S. Jacquot, Résultats d'ergodicité sur l'espace de Wiener et application au recuit simulé, Ph. D. Thesis, Université d'Orleans (1992).

[18] B. Helffer, "Semi-Classical Analysis for the Schrödinger operator and Applications", Lectures Notes in Mathematics **1336** (Springer Verlag, Berlin, 1988).

[19] B. Helffer, J. Sjostrand, Multiple wells in the semi–classical limit, I, *Comm. in P.D.E.* **9** : 337–408 (1984).

[20] B. Helffer, J. Sjostrand, Puits multiples en limit semi–classique, II, *Ann. I.H.P. (Physique Théorique)* **42** : 127–212 (1985).

[21] B. Helffer, J. Sjostrand, Multiple wells in the semi–classical limit, III, *Math. Nachtr.* **124** : 263–313 (1985).

[22] B. Helffer, J. Sjostrand, Puits multiples en limit semi–classique, IV, *Commun. in P.D.E.* **10** : 245–340 (1985).

[23] B. Helffer, J. Sjostrand, Puits multiples en limit semi–classique, V, in: Current Topics in Partial Differential Equations (Kinokuniya Company Lmt., Tokyo, 1985).

[24] B. Helffer, J. Sjostrand, Puits multiples en limit semi–classique, VI, *Ann. I.H.P (Physique Théorique)* **46** : 353–372 (1987).

[25] B. Simon, Semi–classical analysis of low–lying eigenvalues, I; non–degenerate minima, *Ann. I.H.P.* **38** : 295–307 (1983).

[26] B. Simon, Semi–classical analysis of low–lying eigenvalues, II; tunneling, *Ann. of Math.* **120** : 89-118 (1984).

[27] B. Simon, Instantons, double wells and large deviations, *Bull. A.M.S.* **7** : 323–326 (1982).

[28] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.* **97** : 1061–1083 (1976).

[29] L. Gross, C.I.M.E. session on "Dirichlet forms", Varenna, 1992.

[30] O. S. Rothaus, Logarithmic Sobolev inequalities and the spectrum of Schrödinger operators, *J. Funct. Anal.* **42** : 102–120 (1981).

[31] B. Zegarlinski, Log–Sobolev inequalities for infinite one–dimensional lattice system, *Commun. Math. Phys.* **133** : 147–162 (1990).

[32] R. A. Holley, D. W. Stroock, Logarithmic Sobolev inequalities and stochastic Ising models, *J. Stat. Phys.* **46** : 1159–1194 (1987).

[33] J. D. Deuschel, D. W. Stroock, "Large Deviations" (Academic Press, New York, 1989).

[34] S. R. S. Varadhan, "Large Deviations and Applications" CBMS–NSF, 1984).

[35] B. Simon, "Functional Integration and Quantum Physics" (Academic Press, New York, 1979).

[36] E. B. Davies, "Heat Kernels and Spectral Theory", (Cambridge University Press, Cambridge, 1989).

[37] J. L. Lions, E. Magenes, "Problémes aux Limites non Homogénes et Applications", vol. I–II (Dunod, Paris, 1968).

[38] P. Acquistapace, B. Terreni, A unified approach to abstract linear nonautonomous parabolic equations *Rend. Sem. Mat. Univ. Padova* **78** : 47–107 (1987).

[39] P. Acquistapace, Evolution operators and strong solutions of abstract linear parabolic equations, *Diff. Integral Eq.* **1** : 433–457 (1988).

[40] L. Younes, Estimation and Annealing for Gibbsian Fields. *Ann. Inst. H. Poincaré B* **24** : 269–294 (1988).

[41] O. Bratteli, D. W. Robinson, "Operator algebras and quantum statistical mechanics, vol. I–II (Springer–Verlag, Berlin, 1979).

[42] A. Kossakowski, A. Frigerio, V. Gorini, M. Verri, Quantum Detailed Balance and KMS Condition. *Commun. Math. Phys.* **57** : 97–110 (1977); erratum, *Commun. Math. Phys.* **60**, 96 (1978).

[43] V. Gorini, A. Frigerio, M. Verri, A. Kossakowski, E.C.G. Sudarshan, Properties of quantum markovian master equations. *Rep. Math. Phys.* **13** : 149–173 (1978).

[44] A. Frigerio, Simulated Annealing and Quantum Detailed Balance. *J. Stat. Phys.* **58**: 325–354 (1990).

[45] A. Verbeure, Detailed Balance and Critical Slowing Down. In: L. Accardi, W. von Waldenfels (eds.) Quantum Probability and Applications III. Proceedings, Oberwolfach 1987 (Lect. Notes Math., vol. 1303, pp. 354–362, Springer, 1988).

[46] A. Frigerio, G, Grillo, Time–dependent Schrödinger operators and simulated annealing, in: "Semigroup Theory and Evolution Equations, P. Clément, E. Mitidieri, B. de Pagter, Lecture Notes in Pure and Applied Mathematics Marcel Dekker, New York, pp. 165–179 (1991).

[47] G. Grillo, Logarithmic Sobolev inequalities and Langevin algorithms in $\mathbb{R}^n$, preprint (Sissa, Trieste, 1991).

[48] G. Grillo, Long–time behaviour of Langevin algorithms with time–dependent energy function, preprint (Università di Udine, 1992).

[49] A. Frigerio, G. Grillo, Simulated annealing with time–dependent energy function, to appear in *Mathematische Zeitschrift.*

[50] A. Frigerio, G. Grillo, Asymptotic behaviour of time–inhomogeneous evolutions on von Neumann algebras, preprint (Università di Milano, 1992).

[51] P. Baldi, "Equazioni differenziali stocastiche", Quad. Un. Mat. It. **28** (Pitagora 1984).

[52] A. Friedman, "Stochastic Differential Equations and Applications", vol.I (Academic Press 1975).

[53] R. Has'minskii, "Stochastic Stability of Differential Equations" (Sijthoff and Noordhöff, Alphen aan den Rijn, the Netherlands, 1980).

[54] H. Bateman, "Higher Transcendental Functions", vol. II (McGraw-Hill, 1953).

[55] S. Albeverio, J. Brasche, M. Röckner, Dirichlet forms and generalized Schrödinger operators, SFB 237–preprint n.49, Institut für Mathematik, Ruhr–Universität–Bochum, 1989.

[56] M. Reed, B. Simon, "Methods of Modern Mathematical Physics, II. Fourier Analysis, Self-Adjointness (Academic Press, New York, 1975).

[57] T. Kato, Remarks on holomorphic families of Schrödinger and Dirac operators, in "Differential Equations", I. Knowles, R. Lewis (eds.), North Holland, Amsterdam: 341–352 (1984).

[58] Y. G. Oh, Existence of semiclassical bound states of nonlinear Schrödinger equations with potential of class $(V)_\alpha$, *Comm. in P.D.E.* **13** : 1499–1519 (1988); erratum, *Comm. in P.D.E.* **14** 833–834 (1989).

[59] S. R. S. Varadhan, "Lectures on Diffusion Problems and Partial Differential Equations" (Springer-Verlag, Berlin, 1980).

[60] D. W. Stroock, S. R. S. Varadhan, "Multidimensional Diffusion Processes" (Springer-Verlag, Berlin, 1979).

[61] H. L. Cycon, R. G. Froese, W. Kirsch, B. Simon, "Schrödinger Operators" (Springer Verlag, Berlin, 1987).

[62] B. Simon, Schrödinger semigroups, *Bull. Am. Math. Soc.* **7** : 447–526 (1982).

[63] S. Albeverio, R. Høegh–Krohn, Frobenius theory for positive maps on von Neumann algebras, *Commun. Math. Phys.* **64** : 83–94 (1978).
[64] S. Watanabe, Ergodic theorems for W^*–dynamical semigroups, *Hokkaido Math. J.* **8**: 176–190 (1979).
[65] Y. Watanani, M. Enomoto, A Perron–Frobenius type theorem for positive linear maps on C^*–algebras, *Math. Japan* **24** : 53–63 (1979).
[66] A. Frigerio, Stationary states of quantum dynamical semigroups, *Commun. Math. Phys.* **63** : 269–276 (1987).
[67] A. Frigerio, M. Verri, Long–time asymptotic properties of dynamical semigroups on W^*–algebras, *Math. Z.* **180** : 275–286 (1982).
[68] D. W. Robinson, Strongly positive semigroups and faithful invariant states, *Commun. Math. Phys.* **85** : 129–142 (1982).
[69] H. Spohn, An algebraic condition for the approach to equilibrium of an open N–level system, *Lett. Math. Phys.* **2** : 33–38 (1977).
[70] S. Stratila, L. Zsido, "Lectures on von Neumann algebras" (Academie Bucuresti, 1979).
[71] H. Araki, Relative hamiltonian for faithful normal states of a von Neumann algebra, *Publ. Res. Inst. Math. Sci.* **9** : 165–209 (1973).
[72] M. J. Donald, Relative hamiltonians which are not bounded from above, *J. Funct. Anal.* **91** : 143–173 (1990).
[73] M. J.Donald, Continuity and relative hamiltonians, *Commun. Math. Phys.* **136** : 625–632 (1991).
[74] H. Kosaki, Lebesgue decomposition of states on a von Neumann algebra, *Amer. J. Math.* **107** : 697–735 (1985).
[75] Y. Amit, M. Piccioni, A nonhomogeneous Markov process for the estimation of Gaussian random fields with nonlinear observation, *Annals Prob.* **19** : 1664–1678 (1991).
[76] M. Reed, B. Simon, "Methods of Modern Mathematical Physics, IV. Analysis of Operators" (Academic Press, 1978).