HAL
open science

Annealing diffusions in a slowly growing potential

Pierre-André Zitt

To cite this version:

Pierre-André Zitt. Annealing diffusions in a slowly growing potential. Stochastic Processes and their Applications, 2008, 118 (1), pp.76-119. 10.1016/j.spa.2007.04.002 . hal-00084081

HAL Id: hal-00084081

https://hal.science/hal-00084081v1

Submitted on 5 Jul 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Annealing diffusions in a slowly growing potential

Pierre-André Zitt*

5th July 2006

Abstract

We consider a continuous analogue of the simulated annealing algorithm in $\mathbb{R}^d$, namely the solution of the SDE $dX_t = \sigma(t)dB_t - \nabla V(X_t)dt$, where V is a function called potential. We prove a convergence result, similar to the one in [Mic91], under weaker hypotheses on the potential function. In particular, we cover cases where the gradient of the potential goes to zero at infinity. The main idea is to replace the Poincaré and log-Sobolev inequalities used in [Mic91, HCS87] by weak Poincaré inequalities (introduced in [RW01]), and to estimate constants with measure-capacity criteria. We show that the convergence still holds for the "classical" schedule $\sigma(t) = c/\ln(t)$, where c is bigger than a constant related to V.

Keywords: simulated annealing, diffusion process, weak Poincaré inequality

MSC 2000: 90C59, 60J60, 60F99

Produit par le CCSd - 5 Jul 2006

Introduction

The goal of this article is to study a continuous analogue of a discrete optimization algorithm called *simulated annealing*. This algorithm was introduced in 1983 by Kirkpatrick, Gelatt and Vecchi, and aims at finding "good" (if not perfect) solutions to complex problems. The crucial idea is to perturb the standard gradient descent by a random noise; hopefully this noise will get the process out of traps (local minimas), and help it reach the global minimum. The noise is taken relatively big at the beginning, so that the process explores the space, and is gradually reduced thereafter.

The standard case is the discrete case (in time and space); here we consider a process on $\mathbb{R}^d$ in continuous time. Note that more complicated state spaces have been studied, see for example [Jac94, JR95, Jac96]; here we will stick to $\mathbb{R}^d$. This "annealing diffusion" process has already been studied by several authors. Hwang, Chiang and Sheu ([HCS87]) proved the convergence under quite strong assumptions, using comparisons with the associated (ordinary) differential equation and results on the trajectories (estimates of exit times from domains, etc.). The result was enhanced by Royer ([Roy89]). The approach we follow was developed by L. Miclo in [Mic92] (and in his doctoral dissertation [Mic91]), and reduces the problem to the convergence of a single quantity, the free energy. Since then, other questions have been asked: speed of convergence, choice of a better algorithm etc. (see *e.g.* the survey [Loc00]). Let us also note that the "functional inequalities" approach has also been used extensively for other (possibly discrete) models, and other closely related algorithms (see *e.g.* [DMM99] for a study of a generalized simulated annealing process).

A common feature of these works on global optimization on $\mathbb{R}^d$ is the quite strong assumptions they require on the growth of the potential. In particular, the norm of the gradient is supposed to go to infinity at infinity. These hypotheses are technically useful: they guarantee that, at any fixed temperature, the generator has a spectral gap, which in turn gives estimates on the rate of convergence. Let us note that the "cooling schedule" (*i.e.* the choice of the temperature as a function of time) for which the process converges is linked with the speed of explosion of the spectral gap, but that it can be read directly on the potential (see below the remarks on the constant $d^\star$).

A natural question arises: what happens when the gradient of the potential does not go to infinity, and when there is no spectral gap? Do we need to change the cooling schedule to reflect the slow-down of the diffusions at fixed temperature, or does the local structure of the potential dictate the optimal schedule?

Before we answer this question, let us be more precise and give our hypotheses.

We study the following optimization problem: how to find the minimum of a function V on the space $\mathbb{R}^d$. To solve this problem, we introduce the following stochastic differential equation:

$$\begin{cases} dX_t &= \sqrt{\sigma(t)}dB_t - \frac{1}{2}\nabla V(X_t)dt, \\ X_0 &\sim m_0. \end{cases}$$

*Équipe Modal'X (EA3454), Université Paris X, Nanterre

The function σ will be called *temperature*, and will be a (deterministic) function of time, decreasing to zero.

Intuitively, this process is similar to simulated annealing: we perturb a gradient descent by a stochastic term whose intensity decreases over time.

We would like to know if the process finds a point where the global minimum is reached; we will show that it does, in a weak sense.

1 Definition. *The annealing process starting from a law m_0 is said to converge if its law m_t at time t converges weakly to a measure supported by* $\operatorname{argmin} V$. *In particular, if the global minimum of V is reached in a single point x_0, the process converges if m_t goes to a Dirac mass at x_0.*

Let us now recall the result we would like to generalize: this is the main result of [HCS87, Roy89, Mic92], as it appears in [Mic92].

2 Theorem (L. Miclo). *If V satisfies some regularity assumptions, and the following conditions:*

- $V \xrightarrow[x\to\infty]{} \infty$,
- $|\nabla V| \xrightarrow[x\to\infty]{} \infty$,
- $|\nabla V| - \Delta V$ *is bounded from below,*

then there exists a constant $d^\star$ such that, for any $c > d^\star$, and for $\sigma(t) = c/\ln(t)$, the annealing process converges.

To understand the direction in which we generalize this result, let us note that this theorem applies for any potential V which is equal to $|x|^\alpha$ outside a compact set, whenever α is strictly bigger than 1. It is then a quite natural question to ask whether this still holds when α is strictly less than 1. Our hypotheses, which we now state, allow us to treat this case.

Hypothesis 1 (Global minimum). *The potential has a unique global minimum, located at the origin and $V(0) = 0$. Moreover, this minimum is non degenerate:* $\mathbf{Hess} V(0)$ *is positive definite.*

Hypothesis 2 (Growth at infinity). *The potential V goes to infinity at infinity faster than a logarithm:*

$$\exists m_V > 1, \exists C, \quad V(x) \geq \ln(|x|)^{m_V} - C.$$

Hypothesis 3 (Bounded gradient). *The potential V is continuously differentiable, and its gradient is bounded:*

$$\|\nabla V\|_\infty < \infty.$$

Hypothesis 4 (Concavity). *The Laplacian of V is negative at infinity: there exists a compact set K compact such that*

$$\forall x \notin K, \quad \Delta V(x) \leq 0.$$

One last hypothesis will be added in section 3, regarding the structure of local minima of V.

These hypotheses call for a few remarks.

The first one simplifies the problem at hand: there is only one goal to go after. If the weak limit of the equilibrium measures μ_σ (*cf.* infra) is known (some results in this direction may be found in [Mic92, Hwa80]), the arguments given here should work in the same way. The non-degeneracy hypothesis may be weakened too (see *e.g.* section 2 for a slight generalization in $d = 1$) However, this restriction allows for two simplifications: it gives an estimate of the partition function Z_σ, and avoids more intricate reasonings in the computation of the weak inequalities (section 3).

The growth hypothesis is not very restrictive. In particular, V may grow like $|x|^\alpha$ with $\alpha < 1$ (or even slower). These cases were not covered in the literature. Let us note that we do not know what happens in the limit case (when $m_V = 1$, *i.e.* the tails of the equilibrium measures are polynomial).

In the light of the previously known results, the bounded gradient assumption seems less stringent: in some sense, we already know what happens when the gradient is big. The hypothesis could probably be lifted if we allowed a polynomial growth, or a control by V, but we keep it for the sake of clarity.

Finally, the condition on ΔV seems more restrictive. It will only be used in the proof of the moment bound (section B). It could probably be replaced by a condition like $\Delta V \leq C|\nabla V|^2$. However, in the "natural example" where $V(x) = |x|^\alpha$ at infinity, the Laplacian is indeed negative if $\alpha < 1$, and this example

was one motivation for investigating the problem. Moreover, even this weakened hypothesis would not allow the existence of traps at infinity, however shallow they may be. It would be interesting to know what could happen if there were such traps: either they have no effect (in the sense that the same cooling schedule may be chosen), or they slow down the process too much and destroy the convergence.

Our principal result is the following.

3 Theorem. *If the potential V satisfies the hypotheses above, there exists a constant $d^\star$ such that, if we choose*

$$\sigma(t) = \frac{c}{\ln(t)},$$

with $c > d^\star$, the annealing process converges.

This result generalizes theorem 2 by allowing more general choices for the potential function. In particular, as we will see in the sequel, the equilibrium measures need not satisfy a Poincaré inequality. Nonetheless, the critical cooling schedule is the same, which contradicts the intuition that the speed was given by the Poincaré constants. In fact, what seems to prevail is the behavior of V in a compact set, and from a certain point of of view, that is precisely what the weak inequalities capture.

The remainder of the paper is organized in the following way. Firstly, we explain the analytic approach of L. Miclo and give the main line of the proof.

This proof, under our weakened hypotheses, uses *weak Poincaré inequalities*. We will need controls over their dependence on temperature: these are established in sections 2 and 3, respectively in the one- and multi-dimensional case. These three sections are the core of the proof of the convergence result.

The quite technical 4th section gathers definitions and results about Orlicz norms and weak inequalities. Finally, we postpone to the annexes a comparison between functions centered by their mean or by their median, a moment bound for the annealing process, and a brief proof of the estimation of the partition function.

1 The convergence of the process (main line of the proof)

1.1 A differential inequality for the free energy

Before we describe the main idea, we introduce some notation. Consider the SDE defining the annealing diffusion, but with a constant temperature σ. The process is then a classical diffusion with a gradient drift. The corresponding generator is given by:

$$L_\sigma : f \mapsto \frac{\sigma}{2}\Delta f - \frac{1}{2}\nabla V \nabla f.$$

The measure μ_σ defined by

$$d\mu_\sigma = \frac{1}{Z_\sigma}\exp\left(-\frac{V}{\sigma}\right)d\lambda,$$

is reversible for this process (Z_σ is a normalization constant). We will call μ_σ the *instantaneous equilibrium measure*.

It's easy to see that, as σ goes to zero, the measures μ_σ concentrate around the global minimum of the potential (which is found at the origin by hypothesis). In fact, we even have the following convergence.

4 Proposition. *The measures μ_σ converge weakly:*

$$\mu_\sigma \xrightarrow[\sigma\to 0]{} \delta_0.$$

Moreover, the normalization constant Z_σ behaves like $\sigma^{d/2}$.

The asymptotic behavior of Z_σ is proved in annex C.1.

In order to prove the convergence of the process, we follow the approach of L. Miclo ([Mic92]) and show that the relative entropy of the law of the process with respect to its instantaneous equilibrium measure goes to zero.

More precisely, let f_t be the density of $m_t = \mathcal{L}(X_t)$ with respect to the equilibrium measure μ_t. The relative entropy (also called free energy) is $I_t = \int f_t \log f_t d\mu_t$, which can be rewritten as $I_t = \mathbf{Ent}_{\mu_t}(\sqrt{f_t}^2)$. The finiteness of I_t is established in annex C.2. We would like to study the evolution of I_t; the natural idea is to differentiate it. One can justify the following formal computation:

5 Proposition (Differentiation of the free energy). *The derivative of the free energy is given by:*

$$\begin{aligned}\frac{dI_t}{dt} &= \frac{1}{\sigma(t)^2}\sigma'(t)\int V\times(1-f_t)d\mu_t - 2\sigma(t)\int|\nabla\sqrt{f_t}|^2\\ &= \frac{1}{\sigma(t)^2}\sigma'(t)\int V\times(1-f_t)d\mu_t - 2\sigma(t)\mathcal{E}_{\mu_t}(\sqrt{f_t}).\end{aligned}$$

Remark. *By $\mathcal{E}_{\mu t}(f)$ we denote $\int|\nabla f|^2 d\mu_t$. This is somewhat improper — strictly speaking, this is the energy associated with the generator $(1/2)\Delta-(1/(2\sigma))\nabla V\nabla$ (so we should multiply our energy by σ to get the "real" one). However, the classical criteria for functional inequalities are written for this form of the energy.*

The first term is set aside for the time being, we shall bound it later directly by a function of t.

Following the classical path leading from functional inequalities to semigroup estimates, we now try to control the energy term on the right hand side.

If the measures μ_t satisfied logarithmic Sobolev inequalities, everything would be fine: the energy of $\sqrt{f_t}$ could be controlled by its entropy with respect to μ_t, and we would get I_t back on the right hand side of the inequality. We would still have to know how the constants in the logarithmic Sobolev inequality depend on the small parameter σ, and get an upper bound for the first term, but we could get the convergence of I_t to zero.

Unfortunately, the scaling behavior of the constants in the logarithmic Sobolev inequality (*i.e.* the way they behave when σ goes to zero) is not clear. Moreover, this inequality need not hold, and in fact it won't under our hypotheses.

In Miclo's paper, the first difficulty is overcome thanks to a Poincaré inequality, weaker than the logarithmic Sobolev inequality, but for which the constants are well known. However, even this inequality won't be satisfied in our case, and we have to find another way.

Our idea is to consider a still weaker functional inequality, namely a *weak Poincaré inequality*, written with an Orlicz norm. Weak Poincaré inequalities were introduced by M. Röckner and F.-Y. Wang in [RW01], originally with an L^∞ norm and the mean of f instead of a median on the right hand side. We will give a brief account on weak inequalities and Orlicz norms in section 4, and explain the link between the original inequality and the one we use.

For now, let us just state this inequality. It reads:

$$\forall f,\forall r,\qquad \mathbf{Var}_{\mu_t}(f)\le\alpha_t(r)\mathcal{E}_{\mu_t}(f)+r\|f-m_f\|_\phi^2, \tag{1}$$

where m_f is a median of f under μ, and α_t, a decreasing function of r, is the compensating function. The Orlicz norm is not easily tractable, but we will see (*cf.* lemma 30) that it can be bounded by the entropy: there exists a C such that, for all positive f,

$$\|f-m_f\|_\phi^2\le C(\mu(f^2)+\mathbf{Ent}(f^2)).$$

At this point, the energy is bounded above by three terms: $\mu(f^2)$, the entropy of f and its variance. To get rid of the variance term, we would like to bound it by entropy-like quantities. To this end we introduce the following definition.

6 Definition. *For any probability measure μ and any positive f, we will call* pseudo-entropy *the quantity:*

$$\boldsymbol{Ps\text{-}Ent}(f)=\int f\log^2\left(e+\frac{f}{\|f\|_{1,\mu}}\right)d\mu.$$

With this definition in hand, we can state ([Mic92], lemma 4):

7 Lemma. *There exists a δ_0 such that, for all probability measure μ and all positive f with $\mu(f^2)=1$,*

$$\forall\delta<\delta_0,\frac{1}{\delta}\mathbf{Var}_\mu(f)+4\delta\boldsymbol{Ps\text{-}Ent}_\mu(f^2)\ge\mathbf{Ent}_\mu(f^2).$$

Let us put all these inequalities together: we get that for all probability measure μ, if μ satisfies the weak Poincaré inequality 1, then for all positive f with $\int f^2 d\mu=1$,

$$\delta\mathbf{Ent}_\mu(f^2)-4\delta^2\mathbf{Ps\text{-}Ent}(f^2)\le\mathbf{Var}_\mu(f)\le\alpha(r)\mathcal{E}_\mu(f,f)+Cr\mathbf{Ent}_\mu(f^2)+Cr.$$

This entails a lower bound on the energy:

$$\mathcal{E}_\mu(f,f) \geq -\frac{4}{\alpha(r)}\delta^2 \mathbf{Ps\text{-}Ent}(f^2) - C\frac{r}{\alpha(r)} + \frac{1}{\alpha(r)}(\delta - Cr)\mathbf{Ent}_\mu(f^2).$$

Let us get back into our special case, and take $\mu = \mu_t$, $f = \sqrt{f_t}$. The entropy $\mathbf{Ent}(f^2)$ just becomes I_t, and we can plug the inequality back in the differential equation for I_t:

$$\frac{dI_t}{dt} \leq \frac{1}{\sigma(t)^2}\sigma'(t)\int V \times (1-f_t)d\mu_t + 8\delta^2\frac{\sigma(t)}{\alpha_t(r)}\mathbf{Ps\text{-}Ent}(f_t) + 2C\sigma(t)\frac{r}{\alpha_t(r)} - 2(\delta - Cr)\frac{\sigma(t)}{\alpha_t(r)}I_t$$

Since σ is non-increasing in time, we may omit the 1 in $(1-f_t)$ in the first term, and since $f_t d\mu_t = dm_t$,

$$\frac{dI_t}{dt} \leq \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int V dm_t + 8\delta^2\frac{\sigma(t)}{\alpha_t(r)}\mathbf{Ps\text{-}Ent}(f_t) + 2C\sigma(t)\frac{r}{\alpha_t(r)} - 2(\delta - Cr)\frac{\sigma(t)}{\alpha_t(r)}I_t \tag{2}$$

Our goal is to obtain a differential inequality involving only I_t and explicit functions of t, so that we may deduce information on the evolution of I_t. Since σ is known, this leaves us with three questions. First, we have to obtain controls on $\int V dm_t$ and on the pseudo-entropy — we will get explicit bounds in t. Once this is done, we have to estimate the compensating function α_t. Finally we must choose r and δ depending on t in a suitable way, so that the inequality on I_t is good enough to prove the convergence to zero.

We now deal with the first problem.

1.2 Moment bounds and pseudo-entropy

The first inequality is a moment bound on the value of the potential at time t. The proof is postponed to the annexes.

8 Lemma. *Suppose that hypotheses 3 and 4 hold, and that the initial law m_0 satisfies: $\int V^p m_0(dx) < \infty$. Then there exists an M such that:*

$$\int V^p(x)m_t(dx) \leq M\sigma(t)^p \ln(t)^p (\ln\ln t)^{3p}.$$

The last result will be used directly, but it also helps us prove the following bound.

9 Lemma. *Suppose that $\int V^2 dm_0$ is finite, and that the cooling schedule has the form: $\sigma(t) = c/\ln(t)$, for a positive constant c. Then there exists an A such that, for all big enough t,*

$$\boldsymbol{Ps\text{-}Ent}(f_t) \leq A\ln(t)^2(\ln\ln(t))^6.$$

Proof. We differentiate the quantity under scrutiny, namely $J_t = \mathbf{Ps\text{-}Ent}_{\mu_t}(f_t)$. The following formal computation can be justified (*cf.* [Mic92]):

$$\frac{dJ_t}{dt} = -\frac{\sigma(t)}{2}\int F'(f_t)|\nabla f_t|^2 d\mu_t + 2\frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int \log(e+f_t)\frac{f_t}{e+f_t}\left(V - \int V(x)dm_t\right)dm_t,$$

where $F(x) = \frac{2x}{x+e}\log(x+e) + \log^2(x+e)$. Since F is non decreasing (in x), and σ is positive, the first term is bounded above by 0. Moreover, since V is positive and $1/\sigma$ increases, we may also forget the $\int V(x)dm_t$ in the second term. We get:

$$\begin{aligned}
\frac{dJ_t}{dt} &\leq 2\frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int \frac{f_t}{e+f_t}\log(e+f_t)V dm_t \\
&\leq 2\frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int \log(e+f_t)V dm_t \\
&\leq 2\frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\left(\int \log^2(e+f_t)dm_t\right)^{\frac{1}{2}}\left(\int V^2 dm_t\right)^{\frac{1}{2}} \\
&= 2\frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)J_t^{\frac{1}{2}}\left(\int V^2 dm_t\right)^{\frac{1}{2}}.
\end{aligned}$$

After dividing by $2J_t^{\frac{1}{2}}$, the left hand side becomes the derivative of $\sqrt{J_t}$. The right hand side may then be bounded (*cf.* previous lemma):

$$\begin{aligned}\frac{d\sqrt{J_t}}{dt} &\leq \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\left(\int V^2 dm_t\right)^{\frac{1}{2}} \\ &\leq \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\sqrt{M}\sigma(t)\ln(t))(\ln\ln(t))^3.\end{aligned}$$

The explicit value of σ allows us to simplify:

$$\frac{d\sqrt{J_t}}{dt} \leq \sqrt{M}\frac{1}{t}\left(\ln\ln(t)\right)^3.$$

An easy computation shows that the right hand side may be bounded by:

$$\sqrt{M}\frac{d}{dt}\left(\ln(t)(\ln\ln(t))^3\right).$$

To conclude the proof, we integrate this inequality between a (fixed and big enough) t_0 and the current time t. The constant A naturally depends on the initial law m_0 (through the value of M and through the pseudo-entropy at time t_0). □

1.3 From the differential inequality to the convergence of the entropy

It is now time to get back to our differential inequality and apply the bounds we just derived. We fix a logarithmic cooling schedule:

$$\sigma(t) = \frac{c}{\ln(t)}.$$

Recall that we showed (inequality 2):

$$\frac{dI_t}{dt} \leq \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)\int V dm_t + 8\delta^2\frac{\sigma(t)}{\alpha_t(r)}\mathbf{Ps\text{-}Ent}(f_t) + 2C\sigma(t)\frac{r}{\alpha_t(r)} - 2(\delta - Cr)\frac{\sigma(t)}{\alpha_t(r)}I_t$$

We use the moment bound (lemma 8) to deal with the first term, and lemma 9 to bound the second one.

$$\frac{dI_t}{dt} \leq \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)M\ln\ln(t)^3 + 8M\delta^2\frac{\sigma(t)}{\alpha_t(r)}(\ln(t))^2(\ln\ln(t))^6 + 2C\sigma(t)\frac{r}{\alpha_t(r)} - 2(\delta - Cr)\frac{\sigma(t)}{\alpha_t(r)}I_t$$

We number our four terms and define:

$$① = \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right)M\ln\ln(t)^3 \qquad ③ = 2C\sigma(t)\frac{r}{\alpha_t(r)}$$

$$② = 8M\delta^2\frac{\sigma(t)}{\alpha_t(r)}(\ln(t))^2(\ln\ln(t))^6 \qquad ④ = 2(\delta - Cr)\frac{\sigma(t)}{\alpha_t(r)}$$

The inequality becomes:

$$\frac{dI_t}{dt} = ① + ② + ③ - ④ I_t \tag{3}$$

This last inequality will allow us to prove that the free energy goes to zero. To this end, we use the same lemma as L. Miclo:

10 Lemma. *Let I be a positive function, and suppose:*

$$\frac{dI_t}{dt} \leq a(t) - b(t)I(t),$$

where a, b are positive functions and satisfy:

1. $\int^\infty b(t) = \infty$,

2. $\frac{a(t)}{b(t)} \xrightarrow{t\to\infty} 0$.

Then I goes to zero when t goes to infinity.

Our goal is now to use the inequality 3 to check the hypotheses of this lemma. We choose δ and r as follows.

$$\begin{cases} \delta_t &= \frac{1}{\ln(t)^2(\ln\ln(t))^7} \\ r_t &= \frac{1}{C\ln(t)^2(\ln\ln(t))^8} \end{cases} \tag{4}$$

This ensures:

$$\frac{③}{④} = \frac{Cr_t}{\delta_t - Cr_t} \sim \frac{C}{\ln\ln(t)} \to 0,$$
$$\frac{②}{④} = \frac{4M\delta_t^2}{\delta_t - Cr_t}\ln^2(t)(\ln\ln(t))^6 \sim 4M\delta_t \ln^2(t)(\ln\ln(t))^6 \to 0.$$

Two things remain to check:

$$\frac{①}{④} \to 0 \qquad \text{and} \qquad \int^\infty ④ = \infty.$$

This is where we need bounds on the weak Poincaré inequalities: we have to know how α_t behaves for our particular choice of r. This is the aim of the following sections, in one or many dimensions.

In both cases, we will get:

11 Lemma. *There exists a constant $d^\star$ such that, for all $D^\star > d^\star$,*

$$\exists C_\alpha \qquad \alpha_t(r_t) \le C_\alpha \exp\left(\frac{D^\star}{\sigma}\right).$$

For the cooling schedule $\sigma(t) = c/\ln(t)$, we get:

$$\alpha_t(r_t) \le C_\alpha t^{D^\star/c}.$$

In the one-dimensional case, this follows from theorem 12 below, and the choice of r_t. The multi-dimensional case is proved in theorem 17 and the discussion that follows it.

Remark. *The approach in the one- and multi-dimensional case will differ slightly. In the former, we prove a (full) weak Poincaré inequality,* i.e. *we estimate the whole function α_t, and then use this estimate at the point r_t. In the latter, we will only prove a bound on α_t at r_t and disregard the other points.*

We may know get back to our proof. Recall that we have assumed:

$$\sigma(t) = \frac{c}{\ln(t)}, \qquad c > d^\star,$$

so that we may always pick a $D^\star$ strictly less than C.

Let us check the two remaining points. First we must prove that ①/④ converges. Since σ is explicit and we know a bound on $\alpha(r)$, we see that:

$$\begin{aligned}\frac{①}{④} &= \frac{d}{dt}\left(\frac{1}{\sigma(t)}\right) M(\ln\ln(t))^3 \times \frac{\alpha_t(r_t)}{2(\delta - Cr)\sigma(t)} \\ &\le M'\frac{1}{t}\ln(t)^3\ln\ln(t))^{10}\alpha_t(r_t).\end{aligned}$$

where M, M' are constants.

Using the bound on α we just recalled (lemma 11), we get:

$$\frac{①}{④} \le M''\frac{t^{D^\star/c}}{t}\left((\ln t)^3(\ln\ln t)^{10}\right)$$

Since $c > D^\star$, ①/④ goes to zero, as was claimed.

Just in the same way, we have, for t big enough:

$$④ = 2(\delta_t - Cr_t)\frac{\sigma(t)}{\alpha_t(r_t)}$$
$$\sim M''(\ln t)^{-3}(\ln\ln t)^{-7}\frac{1}{t^{\frac{D^\star}{c}}}.$$

Once more, the condition $c > D^\star$ guarantees that the integral of this quantity diverges, which was expected.

This allows us to apply lemma 10, and prove that I_t converges to 0. Thanks to Pinsker's inequality, the total variation between m_t (law of the process) and μ_t (the instantaneous equilibrium) converges too. Since we already know that μ_t converges weakly to the Dirac mass δ_0, this concludes the proof.

1.4 Some remarks

Our theorem immediately raises a few questions. Some of these have already been asked when we discussed the hypotheses — equilibrium measures with polynomial tails are not covered, and we do not know what happens when there are traps at infinity.

It would also be interesting to know what happens if we cool faster than the "good" schedule. *A priori*, the process has no reason to converge to the global minimum; intuitively it should freeze in some local trap. One could ask if this trap is a good approximation of the global aim. Answering this question seems impossible in all generality, one should have to assume much more on the potential function, and on the starting point. The "analytic" approach may not be the best suited for this task.

2 The one-dimensional case

In this section we treat the case of a one-dimensional potential, for which we derive a weak Poincaré inequality (more precisely we prove lemma 11).

The major advantage of this case is that, in one dimension, explicit (Hardy-like) criteria are known for weak inequalities. Thus we are able to prove a quite general result (the de-coupling of the parameters s and σ in the weak inequality). This has a small price: we restrict ourselves to potentials that grow like a power of x, and do not cover the case $V(x) = \log(x)^\alpha$ at infinity (for some $\alpha > 1$). It should be noted that the multidimensional argument (*cf.* next section) may still be used in this logarithmic case.

Let us write down a few notations. The potential V is a real function, continuously differentiable. For any (small) σ, we denote by V_σ the function $\frac{1}{\sigma}V$, and by $Z_\sigma = \int e^{-V_\sigma(x)}dx$ the partition function. We normalize V_σ by defining Φ_σ: $\Phi_\sigma = V_\sigma + \log Z_\sigma$. The equilibrium measure μ_σ reads:

$$d\mu_\sigma = \frac{1}{Z_\sigma}\exp(-V_\sigma)d\lambda = \exp(-\Phi_\sigma)d\lambda.$$

We now state our hypotheses on V. We suppose there exists a compact set $[K_1, K_2]$ such that the following holds.

Hypothesis U 1 (Behavior near the minimum). *In $[K_1, K_2]$, V is bounded below by 0 and above $V(K_1) = V(K_2)$. It reaches its minimum only once, at x_1. Near this point, V behaves like:*

$$V(x) \sim (x - x_1)^b,$$

with $b > 1$. Finally, there exists δ such that V is bijective from $[x_1, x_1+\delta]$ onto its image, and from $[x_1-\delta, x_1]$ onto its image.

This generalizes a little the general assumptions on the minimum: if **Hess**V is positive definite at x_1, it satisfies this hypothesis with $b = 2$.

Hypothesis U 2 (Behavior outside the compact). *Outside the compact, V' and $|V''|/(V'^2)$ are bounded:*

$$\exists C_V \forall x \notin [K_1, K_2], \qquad \frac{|V''|}{V'^2} \leq C_V. \tag{5}$$

In particular, V' has no zero, V decreases before K_1 and increases after K_2.

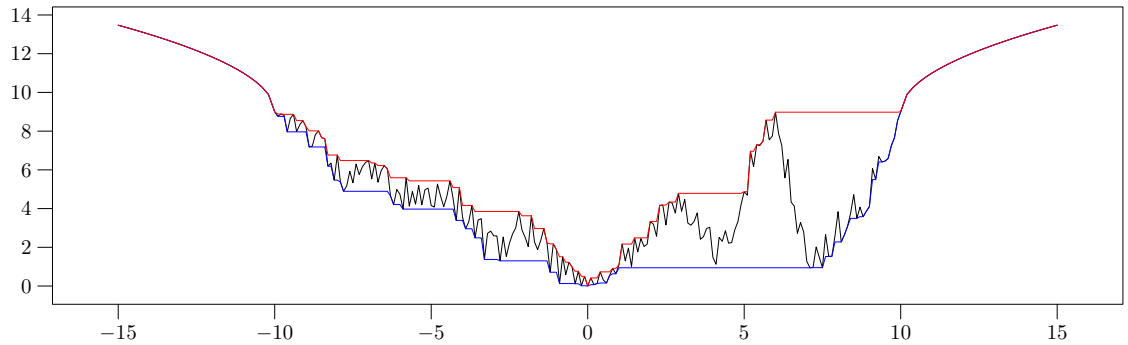


Figure 1: The potential V and the associated functions i and s. Here, $x_1 = 0$, $K_1 = -10$, $K_2 = 10$.

Hypothesis U 3 (The function β). *There exists a function β such that, for all x outside the compact,*

$$\beta\left(\frac{\exp(-V(x))}{V'(x)}\right) \geq \frac{1}{V'(x)^2}. \tag{6}$$

To apply the result to the annealing diffusion, we need an additional growth condition on β:

Hypothesis U 4 (Behavior of β near the origin). *There exist constants A, C such that, near 0, the following holds:*

$$\beta(s) \leq C\left(\log(\frac{1}{s})\right)^A.$$

Remark. *We shall note here that the last two hypotheses hold if $V(x) = x^\alpha$ outside a compact, with $\alpha \in (0,1]$, if we choose $\beta = C\left(\log(1/s)\right)^{\frac{4}{\alpha}-4}$ (cf. [RW01, BCR05]). If v grows like a logarithm to some power, this is not true (β behaves like a power of s). This explains the small loss of generality we spoke about above.*

We define, for all $x \geq x_1$, $i(x) = \inf\{V(y), y > x\}$ and $s(x) = \sup\{V(y), y \in [x_1, x]\}$. In the same way, $i(x) = \inf\{V(y), y < x\}$ and $s(x) = \sup\{V(y), y \in [x, x_1]\}$ for x less than x_1.

Outside $[K_1, K_2]$, we have $i = V = s$, so $s - i$ is continuous with compact support. We call $d^\star$ its maximum value.

The main result of this section may now be stated as follows.

12 Theorem. *The measure μ_σ satisfies a weak Poincaré inequality with the L^∞ norm, with a compensation function β_σ defined by:*

$$\beta_\sigma(s) = C\exp\left(\frac{d^\star}{\sigma}\right)\beta(s),$$

where β is given by the hypothesis. Similarly, μ_σ satisfies a weak inequality with an Orlicz norm and the modified function α_σ given by:

$$\alpha_\sigma(r) = C\beta_\sigma\left(C'\exp(-\frac{4}{r})\right) = C\exp\left(\frac{d^\star}{\sigma}\right)\beta\left(C'\exp\left(-\frac{4}{r}\right)\right).$$

Finally, there exists a constant A such that the following bound holds:

$$\alpha_\sigma(r) \leq C\exp\left(\frac{d^\star}{\sigma}\right)\frac{1}{r^A}.$$

To prove this, we will use a result from Barthe, Cattiaux and Roberto ([BCR05], theorem 3), which gives estimates on the compensating functions for the L^∞ norm. We will then use capacity-measure criteria to derive the result with the Orlicz norm. To state the result we need, we first give some additional notation.

Let m_σ will be a median of μ_σ, and for all x,

$$B_\sigma(x) = \frac{\int_{m_\sigma}^x e^{\Phi_\sigma(y)}dy \times \int_x^\infty e^{-\Phi_\sigma}dy}{\beta\left(\int_x^\infty e^{-\Phi_\sigma(y)}dy\right)}$$
$$B_\sigma = \sup_{x \geq m_\sigma} B_\sigma(x). \tag{7}$$

By symmetry, we also define $b_\sigma(x)$ and b_σ for $x \le m_\sigma$.

The result from [BCR05] reads:

13 Theorem. *Let $\beta : (0,1) \to \mathbb{R}^+$ be non increasing, and B_σ, b_σ be defined by (7). Then μ_σ satisfies the following weak Poincaré inequality :*

$$\mathbf{Var}_{\mu_\sigma}(f) \le C_\sigma \beta(s) \int |\nabla f|^2 d\mu_\sigma + d\,\mathrm{osc}(f)^2,$$

where $C \le 12 \max(b_\sigma, B_\sigma)$.

Note that their result is actually stronger, since it also gives a lower bound on the optimal constant C in terms of some quantities very similar to B_σ.

To use this result, we have to bound $B_\sigma(x)$, and this has to be done uniformly in x. We will split $\mathbb{R}$ into two domains, and show that, in some sense, our choice of β already deals with B_σ for large x, so that the crucial region is near the minimum x_1.

What happens for large x We study the $x \ge K_2$ by following the proof of corollary 4 in [BCR05].

14 Lemma. *For all σ, there exists a c_σ such that:*

$$\forall x \notin [K_1, K_2] \qquad \beta\left(\frac{2e^{-\Phi_\sigma(x)}}{\Phi'_\sigma(x)}\right) \ge \frac{c_\sigma}{\Phi'_\sigma(x)^2}. \tag{8}$$

One may choose $c_\sigma = \frac{1}{\sigma^2}$.

Proof. Recall that the same bound holds for V (*cf.* hypothesis U3); we try to carry it over to Φ_σ.

The behavior of V near its minimum allows us to get an equivalent for Z_σ using Laplace's method (*cf.* for example [Die68]); if $V \sim (x - x_1)^b$, we get

$$Z_\sigma \sim C\sigma^{1/b},$$

where C depends only on V. Let us bound the argument in the function β.

$$2\frac{\exp(-\Phi_\sigma)}{\Phi'_\sigma} = \frac{2\sigma}{Z_\sigma}\frac{\exp(-V/\sigma)}{V'} \le C'\sigma^{1-\frac{1}{b}}\frac{\exp(-V/\sigma)}{V'} \le C'\sigma^{1-\frac{1}{b}}\frac{\exp(-V)}{V'} \le \frac{\exp(-V)}{V'}$$

for σ small enough, because b is strictly greater than 1, so that $\sigma^{1-1/b}$ goes to zero. Since β decreases, we get, outside $[K_1, K_2]$:

$$\beta\left(2\frac{\exp(-\Phi_\sigma)}{\Phi'_\sigma}\right) \ge \beta\left(\exp(-V)/V'\right) \ge \frac{1}{V'^2} = \frac{1}{\sigma^2\Phi'^2_\sigma}.$$

□

15 Lemma. *For all $x \ge K_2$, we have the following inequalities:*

$$\int_{m_\sigma}^{x} e^{\Phi_\sigma} \le \int_{m_\sigma}^{K_2} e^{\Phi_\sigma(y)}dy + 2\frac{e^{\Phi_\sigma}}{\Phi'_\sigma}$$

$$\mu_\sigma([x,\infty)) \le 2\frac{e^{-\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \le 3\mu_\sigma([x,\infty)).$$

Proof. For all $x \ge K_2$ and σ small enough (less than $1/(2C_V)$), the hypothesis on V gives us:

$$\frac{|\Phi''_\sigma(x)|}{\Phi'_\sigma(x)^2} = \sigma\frac{|V''|}{V'^2} \le C_V\sigma \le \frac{1}{2}.$$

Therefore:

$$\left(\frac{\exp(\Phi_\sigma)}{\Phi'_\sigma}\right)' \ge \frac{1}{2}e^{\Phi_\sigma},$$

This gives the first result by integration. In a similar way,

$$\left(\frac{\exp(-\Phi_\sigma)}{\Phi'_\sigma}\right)' \in \left[\frac{1}{2}e^{-\Phi_\sigma}, \frac{3}{2}e^{-\Phi_\sigma}\right]$$

leads to the second claim. □

We are now in a position to bound $B_\sigma(x)$.

$$\begin{aligned}
B_\sigma(x) &= \mu_\sigma([x,\infty)) \times \frac{1}{\beta(\mu_\sigma([x,\infty)))} \times \int_{m_\sigma}^{x} e^{\Phi_\sigma(x)} \\
&\leq 2\frac{e^{-\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \times \frac{1}{\beta\left(2\frac{e^{-\Phi_\sigma}}{\Phi'_\sigma}\right)} \times \left(\int_{m_\sigma}^{K_2} e^{\Phi_\sigma(y)}dy + 2\frac{e^{\Phi_\sigma(x)}}{\Phi'_\sigma(x)}\right) && \text{(by lemma 15)} \\
&\leq 2\frac{e^{-\Phi_\sigma(x)}}{\Phi'_\sigma(x)} \times \frac{\Phi'_\sigma(x)^2}{c_\sigma} \times \left(\int_{m_\sigma}^{K_2} e^{\Phi_\sigma(y)}dy + 2\frac{e^{\Phi_\sigma(x)}}{\Phi'_\sigma(x)}\right) && \text{(by lemma 14)} \\
&\leq \frac{2}{c_\sigma}\Phi'_\sigma(x)\int_{m_\sigma}^{K_2} e^{V_\sigma(y)-V_\sigma(K_2)}dy + \frac{4}{c_\sigma}. && \text{(because } V(x) \geq V(K_2))
\end{aligned}$$

The hypotheses imply that $V(y) \geq V(K_2)$, whenever $|y| \leq K_2$. On the other hand, Φ'_σ is bounded above by C/σ (since V' is supposed to be bounded). Finally,

$$\forall x \geq K_2, \qquad B_\sigma(x) \leq \frac{C'}{c_\sigma \sigma},$$

where C' is independent of σ.

What happens in the well The general strategy here is to bound $B_\sigma(x)$ by studying only the numerator. The denominator can be (very) roughly bounded by $\beta(1/2)$ (which does not depend on σ). The partition function disappears, and we get:

$$B_\sigma(x) \leq C\int_{m_\sigma}^{x} e^{V(y)/\sigma}dy \times \int_{x}^{\infty} e^{-V(y)/\sigma}dy.$$

We need a bound on V near the median: under our hypotheses, since μ_σ converges weakly to δ_{x_1}, the continuity of V in x_1 yields (for σ small enough):

$$\forall x \in [(x_1, m_\sigma)], \qquad V(x) \leq d^\star/4.$$

Now we can bound the first integral in the following way:

$$\int_{m_\sigma}^{x} e^{V(y)/\sigma} \leq (K_2 - K_1)\exp\left(\frac{1}{\sigma}\max(s(x), d^\star/4)\right),$$

where $d^\star/4$ takes care of the case when m_σ is less than x_1.

We cut the second integral in two parts:

$$\int_{x}^{\infty} e^{-V_\sigma(y)}dy \leq \int_{x}^{K_2} e^{-V_\sigma(y)}dy + \int_{K_2}^{\infty} e^{-V_\sigma(y)dy}.$$

Since V is strictly increasing after K_2, we may apply Laplace's method to the second term. In the first one, we use a rough bound on V:

$$\int_{x}^{\infty} e^{-V_\sigma(y)}dy \leq (K_2 - K_1)\exp\left(-\frac{i(x)}{\sigma}\right) + C\exp\left(-\frac{V(K_2)}{\sigma}\right).$$

Since $i(x)$ is less than $V(K_2)$, the second term is less than the first one (up to a constant), and there exists C' such that:

$$\int_{x}^{\infty} e^{-V(y)}dy \leq C'\exp\left(-\frac{i(x)}{\sigma}\right).$$

Coming back to B_σ, we get:

$$\begin{aligned}
B_\sigma(x) &\leq C''\exp\left(\frac{1}{\sigma}\left(\max(s(x), d^\star/4) - i(x)\right)\right) \\
&\leq C'''\exp\left(\frac{d^\star}{\sigma}\right).
\end{aligned}$$

Conclusion : An upper bound on β Let us now gather the bounds on $B_\sigma(x)$ we derived in the preceding paragraphs.

16 Lemma. *There exists a C (independent of σ) such that, for all σ,*

$$B_\sigma = \sup_{x \geq m_\sigma} B_\sigma(x) \leq C \exp\left(\frac{d^\star}{\sigma}\right).$$

With this result in hand, we may apply Barthe, Cattiaux and Roberto's result (theorem 13): this proves the first claim of theorem 12.

The modified function α_σ is deduced from β_σ with the help of theorem 29 (see below, in section 4).

Finally, the growth hypothesis on β (U4) guarantees that, near 0, β is bounded by a power of $\ln(1/s)$; this immediately implies the last result, and concludes the proof.

3 The weak inequality in any dimension

We now turn to the proof of the weak inequality (the bound in lemma 11) in any dimension. We are going to need one more hypothesis on the structure of potential wells, to avoid "pathological" cases.

After that, we proceed in several steps. First we recall our aim and explain the main lines of the proof. During this proof, a certain "path" (in fact, an open set of $\mathbb{R}^d$) will appear. It will be used to derive a "capacity-measure" inequality. Eventually, we will go from this inequality to the one we seek, using a result from next section.

3.1 The last hypothesis on the potential

To write down the last hypothesis we shall make on V, we first need a few more notations.

For all $x \in \mathbb{R}^d$, we call Γ_x the set of paths from x to 0. For each such γ (γ is a continuous function from $[0,1]$ into $\mathbb{R}^d$), we call $h(\gamma)$ the "height" of γ, *i.e.* the highest value taken by V along γ:

$$h(\gamma) = \sup_{t \in [0,1]} V(\gamma(t)).$$

Now suppose we try to go from x to 0 while remaining as low as possible (*i.e.* we try to find a path where V is small). There is a minimum price to pay; whatever path we choose, we will necessarily go at least as high as:

$$h(x) = \inf_{\gamma \in \Gamma_x} h(\gamma).$$

We will call "good paths" the ones that stay below that minimal height:

$$\gamma \text{ is good} \Leftrightarrow h(\gamma) = h(x).$$

A priori, for a given x, a good path from x to 0 need not exist: it may well be the case that, if one tries to find γ such that $h(\gamma) \leq h(x) + 1/n$, one has to go farther and farther as n grows, and that no finite path achieves the infimum bound.

Finally, the height of the "potential barrier" between x and the global minimum will be called $d^\star(x)$:

$$d^\star(x) = h(x) - V(x),$$

and the height of the biggest barrier will be just $d^\star$:

$$d^\star = \sup_x d^\star(x).$$

Hypothesis 5. *The potential barriers have a bounded height:*

$$d^\star < \infty.$$

Moreover, each point can reach 0 by a "relatively short" good path. More precisely, there exists a function R (a maximal radius), from $\mathbb{R}^d$ to $\mathbb{R}$, which satisfies the following conditions:

- *For all x, the ball centered in zero and of radius $R(|x|)$ contains a good path for x:*

$$\forall x, \exists \gamma \in \Gamma_x, \begin{cases} \gamma([0,1]) & \subset \mathcal{B}_{R(|x|)} \\ \gamma & \textit{is good.} \end{cases}$$

- *The function R grows like a power of the distance to the origin:*

$$R(|x|) \leq c_R |x|^{d_R}.$$

3.2 The one-point weak inequality

As was said before, we will not prove in this section a full weak Poincaré inequality, *i.e.* we will not get (1) for all r. Instead, we just prove it for a specific value of r, namely the $r = r_t = (\ln t)^{-2}(\ln\ln t)^{-8}$ (*cf.* equation (4)). Since $\sigma(t) = c/(\ln t)$, we note that:

$$\begin{aligned} r_t &\geq C\frac{\sigma^2}{\ln(\sigma)^8} \\ &\geq C'\sigma^m, \end{aligned}$$

for some C, C' and σ small enough. Therefore, and since α_t decreases, we may prove an inequality with σ^m instead of r_t.

More precisely, we will get:

17 Theorem. *Let m be a real number, strictly smaller than $1 + m_V$, and let $D^\star$ be a constant, $D^\star > d^\star$. Then there exists a C_m such that, for all σ, the measure μ_σ satisfies the following one-point weak Poincaré inequality*

$$\forall f, \mathbf{Var}_{\mu_\sigma}(f) \leq C_m \exp\left(\frac{D^\star}{\sigma}\right) \int |\nabla g|^2 d\mu_\sigma + \sigma^m \|f - m_f\|_\phi^2.$$

where m_f is a median of f under μ_σ.

As was noted before, this entails

$$\alpha_t(r_t) \leq \alpha_t(\sigma^m) \leq C_m \exp\left(D^\star/\sigma\right),$$

which is the result of lemma 11.

The end of the section is devoted to the proof of the theorem. It can be sketched as follows.

The idea is to use a capacity-measure criterion restricted to certain sets (large enough sets). Intuitively, if a set A has a large μ_σ mass, it must contain points near the origin; and these points are the important ones, for measuring capacity as well as mass. For these sets, located near the origin, everything should behave as in the compact case, and the inequality should depend on σ in the same way as when a Poincaré inequality holds.

Let us fix $D^\star$, strictly bigger than $d^\star$. As was just said, we would like to compare the capacity and measure of large enough sets: let $\kappa > 0$ be the minimum mass we will consider (κ will depend on σ). Let A be a Borel set such that:

$$\mu_\sigma(A) \geq 2\kappa(\sigma).$$

Restricting ourselves to these large sets localizes the problem in some sense. To be more precise, we introduce two radii. The first one, r_σ, is such that:

$$\mu_\sigma(\mathcal{B}_{r_\sigma}) \geq 1 - \kappa.$$

The second one is deduced from it: it is a radius big enough to include good paths (*cf.* hypothesis 5) starting from any point in the small ball $\mathcal{B}_{r_\sigma}$.

$$R_\sigma = R(r_\sigma).$$

These two quantities depend on σ and κ; we will see that, for our choice of κ, r_σ and R_σ won't grow too fast as σ goes to zero.

Let $A = A' \cup A''$, where $A' = A \cap \mathcal{B}_{r_\sigma}$ and A'' is the complement set. Since $\mu_\sigma(A) \geq 2\kappa$ and $\mu_\sigma(A'') \leq \kappa$ (by definition of r_σ), $\mu_\sigma(A') \geq \kappa$, and:

$$\mu_\sigma(A) = \mu_\sigma(A') + \mu_\sigma(A'') \leq 2\mu_\sigma(A').$$

Intuitively, we need only consider the subset A', because it concentrates enough mass.

At this point, our set A' may still be very complicated. In particular, it could be scattered all over the ball $\mathcal{B}_{r_\sigma}$. To avoid this, we will once again restrict ourselves to a subset, trying to keep enough mass in the process.

This is done by cutting $\mathcal{B}_{r_\sigma}$ into small cubes. The bound on the gradient of V (hypothesis 3) helps us choose a good mesh, such that V does not vary too much inside a little cube.

18 Proposition. *For all η, there exists ϵ (depending only on V and η), such that, on each cube B with radius ϵ,*

$$\sup_B V - \inf_B V \leq \eta.$$

The parameter η will be chosen later.

So we cut $\mathcal{B}_{r_\sigma}$ into many little cubes of radius ϵ. This requires a certain number of cubes, which we call n_σ. We then have:

$$\mathcal{B}_{r_\sigma} = B_1 \cup B_2 \dots B_{n_\sigma}. \tag{9}$$

In the same way, N_σ will be the number of cubes necessary to cover $\mathcal{B}_{R_\sigma}$. We denote by A_i the intersection of A and B_i. We apply the pigeonhole principle to say that one of the A_i's must be large enough:

$$\exists i_0, \mu_\sigma(A_{i_0}) \geq \frac{1}{n_\sigma}\mu_\sigma(A').$$

To sum up our considerations on sets, for each A, we have found a subset A_{i_0} such that:

- A_{i_0} is a subset of a cube of radius ϵ,
- A_{i_0} is not too far from the origin ($A_{i_0} \subset \mathcal{B}_{r_\sigma}$)
- A_{i_0} is big enough compared to A : $\mu_\sigma(A_{i_0}) \geq \frac{1}{2n_\sigma}\mu_\sigma(A)$.

In some sense, we need only consider the case when A looks like a ball and is not too far from the origin. We are going to see how this can be used to build a certain path between A_{i_0} and 0, and from this path, deduce a capacity-measure inequality.

3.3 Building a path and straightening it out

Recall that our goal is to compare the capacity and the measure of sets, and more precisely to bound the capacity from below and the measure from above.

The capacity is defined by an infimum bound:

$$\mathrm{Cap}_\mu(A) = \inf\left\{\int |\nabla f|^2 d\mu, \mathbf{1}_A \leq f \leq 1, \mu(\mathrm{supp}\, f) \leq \frac{1}{2}\right\}. \tag{10}$$

Note that we only define capacities for sets whose measure is less than 1/2. This restriction explains why we use function recentered by their median when we deduce functional inequalities from capacity-measure criteria.

Since we seek a bound from below, we consider a function satisfying the conditions, and we try to bound:

$$\int |\nabla f|^2 d\mu.$$

The key idea is to find out a region of $\mathbb{R}^d$ which should contribute a lot to this integral. Since the function f equals 1 near A, and 0 near 0 (the measure of its support being less than 1/2), there must be a transition between A and 0: this is where the gradient of f appears. Still on the intuitive level, if the integral is to be small, we had better make this transition in a region where μ has less mass, *i.e.* in a zone where V is large. This is the reason why we introduced the good paths: to go from A to zero, a large contribution to the energy should appear along these good paths.

To put these ideas on a firm ground, we will build, starting from A (or more precisely from A_i), an open set $\mathcal{C}_A$ with good regularity properties, and then bound the capacity by integrals over this open set. This construction is depicted in figure 2.

Once this set is built, we proceed in two steps. First, for all function f satisfying the conditions of (10),

$$\int |\nabla f|^2 d\mu_\sigma \geq \int |\nabla f|^2 \mathbf{1}_{\mathcal{C}_A} d\mu_\sigma$$

On the path, we know by design that V is bounded above by $V(x_A^\star) + \eta$. Indeed, V is less than $V(x_A^\star)$ along γ, and the size ϵ of the cubes has been chosen so that on each cube, the oscillation of V is less than η. Therefore, we may compare our integral with an integral with respect to the Lebesgue measure.

$$\int |\nabla f|^2 d\mu_\sigma \geq \frac{1}{Z_\sigma}\exp\left(-V(x_A^\star) - \eta\right)\int |\nabla f|^2 \mathbf{1}_{\mathcal{C}_A} d\lambda. \tag{11}$$

The next step is to bound the latter integral on $\mathcal{C}_A$. Our only hypotheses is that f must be 1 on A, and 0 near zero. The idea is then to apply a Poincaré inequality to compare the energy to a variance. Unfortunately, though we know that a Poincaré inequality effectively holds under quite general assumptions for a bounded domain in $\mathbb{R}^d$ (this is proved in many textbooks on partial differential equations, see *e.g.* [Eva98],p. 275–276), the explicit constants and their behaviour when the domain changes is not well known. However, there is a case for which we have such explicit estimates, namely the case of convex domains.

19 Theorem (Poincaré inequality for convex domains). *Let L be a convex bounded domain in $\mathbb{R}^d$. Then the Lebesgue measure on L satisfies a Poincaré inequality, and the constant can be bounded above using only the diameter d_L of the domain:*

$$\mathbf{Var}_{\lambda_L}(f) = \int \left(f - \int f d\lambda_L \right) d\lambda_L \leq \frac{d_L^2}{\pi^2} \int |\nabla f|^2 d\lambda.$$

This theorem is proved *e.g.* by Payne and Weinberger, and Bebendorf in [PW60, Beb03]. Note that other bounds in more complicated cases have been derived (see [CL97] for star-shaped domains, or [Che90] for bounds depending on the geometry of the boundary).

Note that, by abuse of notation, we use **Var** for a non-normalized measure.

In order to use this result, we try to "straighten out" the set $\mathcal{C}_A$.

We will build a function ϕ sending $\mathcal{C}_A$ to a tube $\mathcal{L}_A$. This function will be defined piecewise, on each of the little cubes that $\mathcal{C}_A$ crosses. Let us denote these cubes as $C_0, \dots C_m$. It is easy to see that the intersection of $\mathcal{C}_A$ and one of these cubes can only take a finite number of shapes (up to a rotation and/or translation). In $d = 2$ for example, only two different shapes are possible (either a straight tube or a bended one, see figure 2). Each of these shapes may be "straightened out" into a tube by a diffeomorphism. We have to be a bit careful in choosing these diffeomorphisms ϕ_j (one for each shape). We will ask two things: they should behave like a rigid motion in the neighborhood of the edges (so we may "glue" two transformations together), and their Jacobian matrix should be sufficiently "nice" (the "niceness" needed will be made precise later). Such a choice is possible; see the figure 2 for an explanation of a possible way to find such good functions.

Once this is done, we only have to glue our pieces together. Let us denote the pieces $\mathcal{C}_A \cap C_i$ by T_i. We leave T_0 where it stands, and look at T_1. We have seen that it may be straightened into a tube, T_1': define ϕ on T_1 to be precisely this transformation. Now consider T_2: we can straighten it by one of our ϕ_j, and then use a rotation and/or a translation to put it next to T_1'. Since we have asked that the ϕ_j should be rigid motions near the edges, the two pieces of ϕ define a diffeomorphism from $T_1 \cup T_2$ to the straight tube $T_1' \cup T_2'$. We may iterate the process and eventually we get a diffeomorphism ϕ from $\mathcal{C}_A$ to $\mathcal{L}_A$. One can see on the figure that a little extra care is needed to deal with the end of the path $\mathcal{C}_A$ — however, adding just one ϕ_j to our set of transformations settles the question.

Remember that our goal is to use the Poincaré inequality on the convex set $\mathcal{L}_A$. For this to work, we need to control some quantities related to the map ϕ.

20 Proposition. *There exists a constant C_ϕ, which may depend on ϵ but not on σ, such that, at every point, the Jacobian matrix J_ϕ satisfies*

$$|\det(J_\phi)| \leq C_\phi,$$
$$\lambda_1({J_\phi}^t J_\phi) \geq C_\phi^{-1},$$

where $\lambda_1(M)$ is the smallest eigenvalue of the symmetric matrix M.

Proof. This holds by design of the map ϕ. At each point, ϕ is the composition of a rigid motion (which has no effect on the eigenvalues or the determinant of the Jacobian matrix), and of one of the ϕ_j. For a given ϕ_j, the properties hold: we have designed the ϕ_j as restrictions of diffeomorphisms on larger sets, so the bounds hold by compacity. Since there is a finite number of ϕ_j, we may choose bounds that do not depend on j. This proves that the bounds hold for ϕ. □

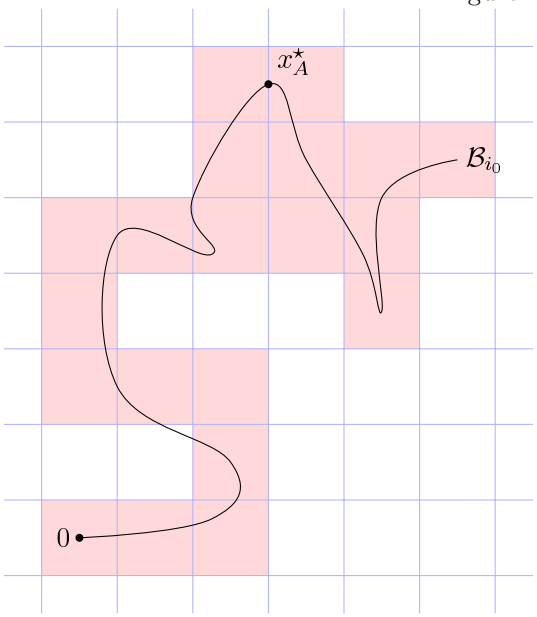
We may now give our "straightening" its rigorous form, namely a change of variables.

21 Proposition. *Let $\mathcal{U}$ and $\mathcal{V}$ be open sets, let ϕ be a diffeomorphism from $\mathcal{U}$ onto $\mathcal{V}$. If the inequalities in the preceding lemma hold with a constant C_ϕ, then for all continuously differentiable function f on $\mathcal{U}$, we have:*

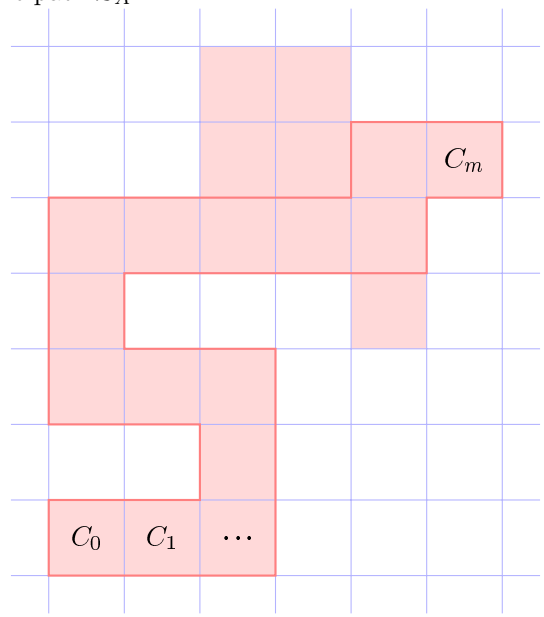
$$\int_{\mathcal{U}} |\nabla f|^2 d\lambda \geq C_\phi^{-2} \int_{\mathcal{V}} |\nabla g|^2 d\lambda.$$

where $g = f \circ \phi^{-1}$.

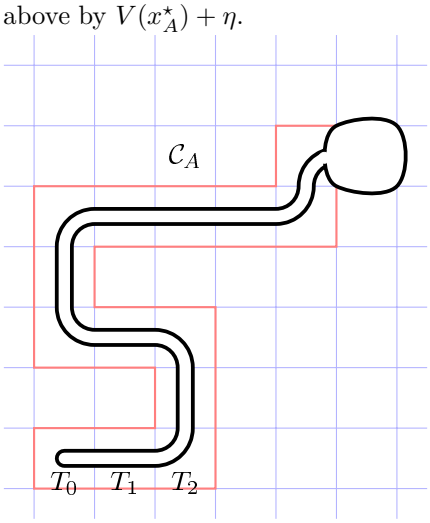
Figure 2: Building the path $\mathcal{L}_A$



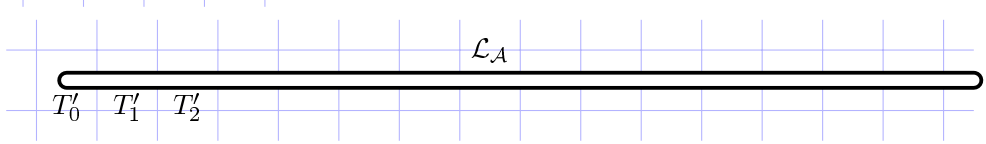
1. We consider a good path starting from the center of the cube $\mathcal{B}_{i_0}$, and going to the origin. On this path, V reaches its maximum at some $x_A^\star$, and on the colored region, V is bounded above by $V(x_A^\star) + \eta$.



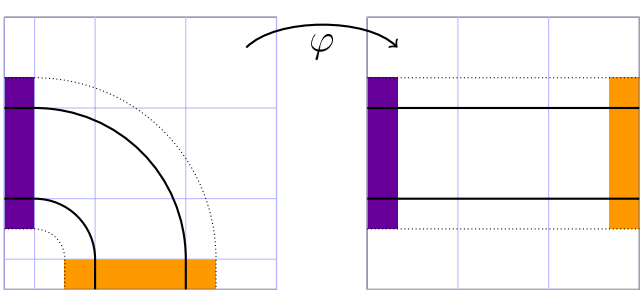
2. We pick a "path of cubes" from 0 to $\mathcal{B}_{i_0}$ which stays entirely within the colored region.



3. Within this path, we draw a smooth tube $\mathcal{C}_A$. The intersection of $\mathcal{C}_A$ and a given little cube may only take a finite numbre of shapes (up to a rigid motion); in this 2-dimensional drawing for example, we have either a straight tube (T_1) or a bended one (T_2). For technical reasons, we consider two more shapes at the end of the tube so that $\mathcal{B}_{i_0}$ lies entirely within $\mathcal{C}_A$.



4. Finally, the tube $\mathcal{C}_A$ is sent onto $\mathcal{L}_A$, a convex set for which we have an explicit Poincaré inequality.



This is how the bended tube on the left may be straightened. We consider a diffeomorphism which sends the regions between dotted lines on one another, and ask that it should be a rigid motion on the dark regions. Defining the transformation on a set (the region between dotted lines) larger than the tube (the region between plain lines) gives compacity bounds on the Jacobian.

Proof. It's a change of variables. Let us define F by $F(x) = |\nabla f|^2(x)$. Then:

$$\int_{\mathcal{U}} |\nabla f|^2 dx = \int_{\mathcal{U}} F(x)dx = \int_{\mathcal{V}} F \circ \phi^{-1} |\det J_\phi^{-1}| dy$$
$$\geq \frac{1}{C_\phi} \int_{\mathcal{V}} F \circ \phi^{-1} dy.$$

Since $f = g \circ \phi$, the gradients are given by:

$$(\nabla f)_x = ({}^t J_\phi)_x (\nabla g)_{\phi(x)}.$$

Taking norms, and using the lower bound on the first eigenvalue, we get:

$$(|\nabla f|^2)_x = {}^t\nabla g J_\phi {}^t J_\phi \nabla g \geq C_\phi^{-1} (|\nabla g|^2)_{\phi(x)}.$$

Rewriting this in y variables,

$$F \circ \phi^{-1}(y) = (|\nabla f|^2)_{\phi^{-1}(y)} \geq C_\phi^{-1} (|\nabla g|^2)_y.$$

Finally,

$$\int_{\mathcal{U}} |\nabla f|^2 \geq \frac{1}{C_\phi^2} \int_{\mathcal{V}} |\nabla g|^2.$$

□

Putting the last two propositions together, we can show:

22 Proposition. *There exists a C, depending only on ϵ, such that if f satisfies the following conditions:*

1. *f is continuously differentiable from $\mathcal{C}_A$ into $[0,1]$,*
2. *$\lambda(\{f = 0\}) \geq l_0$,*
3. *$\lambda(\{f = 1\}) \geq l_1$,*

then

$$\int |\nabla f|^2 d\lambda \geq \frac{C}{N_\sigma^2} \min(l_0, l_1).$$

We recall that N_σ is the number of balls of radius ϵ needed to cover the big ball $\mathcal{B}_{R_\sigma}$.

Proof. Suppose f satisfies the hypotheses. Define $g = f \circ \phi^{-1}$ as in the preceding proposition. The various bounds needed on the Jacobian matrix of ϕ are provided by proposition 20. These bounds also imply that g must vanish at least on a set of Lebesgue measure $C_\phi^{-1} l_0$, the same being true for the set where $g = 1$. The change of variables has shown:

$$\int |\nabla f|^2 d\lambda \geq C_\phi^{-2} \int |\nabla g|^2 d\lambda.$$

On the right hand side, we can now use the Poincaré inequality:

$$\int |\nabla g|^2 d\lambda \geq \frac{1}{C_P(\mathcal{L}_A)} \mathbf{Var}_\lambda(g).$$

The very purpose of our change of variables was to make the domain convex, so we could make use of theorem 19. The constant may therefore be bounded by the square of the diameter of $\mathcal{L}_A$. Since $\mathcal{L}_A$ results from gluing together at most N_σ little cubes of radius ϵ, the square of the diameter may be bounded by $\epsilon^2 N_\sigma^2$.

We now turn to the variance, and use the information on the sets where g is 0 or 1. We denote by l_0', l_1' the respective measures of these sets, and by m the mean of g ($m \in [0,1]$). Then:

$$\mathbf{Var}(g) = \int (g - m)^2 d\lambda \geq m^2 l_0' + (1-m)^2 l_1'.$$

The right hand side is easily shown to be greater than $(l_0' l_1' / (l_0' + l_1')$. The latter is bounded below by half the minimum of l_0' and l_1' (because the numerator is less than $2\max(l_0', l_1')$). Since $l_0' \geq C_\phi^{-1} l_0$, and the similar result holds for l_1,

$$\int |\nabla f|^2 d\lambda \geq \frac{C_\epsilon}{N_\sigma^2} \min(l_0, l_1).$$

□

We may now prove the measure-capacity inequality we are looking for. Indeed, recall that our aim is to bound the capacity of a set A from below by a function of its measure. The previous inequality is almost what we want: on the left hand side is (up to a factor, see (11) above) the quantity whose infimum gives the capacity (equation (10)), and on the right hand side l_0 and l_1 are measures of some sets. It remains to show that these measures may be compared to the measure of A.

3.4 The measure-capacity inequality

Let us put together the results from the previous section (equation (11) and proposition 22)

$$\begin{aligned}\int |\nabla f|^2 d\mu_\sigma &\geq \frac{1}{Z_\sigma} \exp\left(-V(x_A^\star) - \eta\right) \int |\nabla f|^2 d\lambda \\ &\geq \frac{C_\epsilon}{Z_\sigma N_\sigma^2} \exp\left(-V(x_A^\star) - \eta\right) \min(l_0, l_1), \end{aligned} \tag{12}$$

where l_0, l_1 are the Lebesgue measure of the following sets:

$$l_0 = \lambda(\{f = 0\} \cap \mathcal{C}_A) \qquad\qquad l_1 = \lambda(\{f = 1\} \cap \mathcal{C}_A).$$

To bound l_0, we use the fact that f vanishes on a sufficiently large set (as measured by μ_σ). Since μ_σ concentrates around 0, f should vanish near the origin. More precisely, for a fixed ϵ, we know that for σ small enough, the cube centered in 0 and of radius ϵ concentrates 3/4 of the measure. If this cube is labelled B_0, we have:

$$\mu_\sigma(\{f = 0\} \cap B_0) \geq \frac{1}{4}.$$

Since V is non negative, μ_σ and λ are easily compared.

$$\begin{aligned}\frac{1}{4} \leq \mu_\sigma(\{f = 0\} \cap B_0) &= \frac{1}{Z_\sigma} \int \mathbf{1}_{f=0} \mathbf{1}_{B_0} \exp(-\frac{V}{\sigma}) d\lambda \\ &\leq \frac{1}{Z_\sigma} \int \mathbf{1}_{f=0} \mathbf{1}_{B_0} d\lambda \end{aligned}$$

The integral on the right hand side is less than l_0, therefore:

$$l_0 \geq m_0 = \frac{Z_\sigma}{4}.$$

Let us derive a similar bound, m_1, for l_1. On the cube B_{i_0}, $V \geq V(x_A) - \eta$, so:

$$\begin{aligned}\mu_\sigma(A_{i_0}) &= \frac{1}{Z_\sigma} \int \mathbf{1}_{A_{i_0}} \exp(-\frac{V}{\sigma}) d\lambda \\ &\leq \frac{1}{Z_\sigma} \int \mathbf{1}_{A_{i_0}} \exp(-\frac{V(x_A)}{\sigma} + \frac{\eta}{\sigma}) d\lambda \\ &\leq \frac{1}{Z_\sigma} \exp\left(\frac{-V(x_A) + \eta}{\sigma}\right) \lambda(A_{i_0}). \end{aligned}$$

Therefore:

$$l_1 \geq \lambda(A_{i_0}) \geq m_1 = Z_\sigma \exp\left(\frac{V(x_A)}{\sigma} - \frac{\eta}{\sigma}\right) \mu_\sigma(A_{i_0}). \tag{13}$$

Since we would like to control $\min(l_0, l_1)$, we now have to compare the two bounds m_0 and m_1. This is possible thanks to the following inequality:

$$\mu_\sigma(A_{i_0}) \leq \frac{1}{Z_\sigma} \exp\left(\frac{-V(x_A) + \eta}{\sigma}\right) \epsilon^d.$$

If we gather almost all terms on the left hand side, we recognize m_1:

$$m_1 \leq \epsilon^d.$$

Since $m_0 = Z_\sigma/4$, it holds that $m_0 \geq Z_\sigma m_1 \epsilon^{-d}$, and since Z_σ goes to zero, it also holds that $m_1 \geq Z_\sigma m_1 \epsilon^{-d}$, so that both l_0 and l_1 may be bounded below by this quantity:

$$\min(l_0, l_1) \geq \frac{Z_\sigma^2}{\epsilon^d} \exp\left(\frac{V(x_A)}{\sigma} - \frac{\eta}{\sigma}\right) \mu_\sigma(A_{i_0}).$$

Going back to (12), we conclude:

$$\int |\nabla f|^2 d\mu_\sigma \geq \frac{C_\epsilon}{Z_\sigma N_\sigma^2} \exp\left(-V(x_A^\star) - \eta\right) \min(l_0, l_1)$$
$$\geq \frac{C'_\epsilon Z_\sigma}{N_\sigma^2} \exp\left(V(x_A) - V(x_A^\star) - 2\eta\right) \mu_\sigma(A_{i_0}).$$

By definition of $x_A^\star$, $V(x_A) - V(x_A^\star) - 2\eta \geq -d^\star - 2\eta \geq -D^\star$. On the other hand, A_{i_0} was chosen precisely because it contained enough of A's mass: $\mu_\sigma(A_{i_0}) \geq (2n_\sigma)^{-1}\mu_\sigma(A)$. Finally, every function f we can choose in the definition of capacity must satisfy:

$$\int |\nabla f|^2 d\mu_\sigma \geq \frac{C'_\epsilon \sigma^2}{N_\sigma^2 n_\sigma} \exp\left(-\frac{D^\star}{\sigma}\right) \mu_\sigma(A),$$

where $C'_\epsilon = CC_\epsilon$. Taking the infimum over all possible f finally yields the following result.

23 Proposition. *Let $\kappa(\sigma)$ be a positive number, less than $1/2$. Let n_σ, N_σ be defined as in the discussion near equation (9). Then the following bound holds:*

$$\forall A, \mu_\sigma(A) \geq \kappa(\sigma) \quad \Longrightarrow \quad \mu_\sigma(A) \leq \frac{N_\sigma^2 n_\sigma}{C'_\epsilon \sigma^2} \exp\left(\frac{D^\star}{\sigma}\right) \mathrm{Cap}_{\mu_\sigma}(A). \tag{14}$$

3.5 Conclusion

The bigger part of the proof has now been done; the last thing we need to check is that the number of balls n_σ and N_σ do not grow too fast as σ decreases. Then we will apply theorem 28 to deduce the one-point inequality of theorem 17 from our measure-capacity inequality.

Recall that we are given a real number m, strictly smaller than $1 + m_V$. Define $\kappa(\sigma) = \exp\left(-\frac{1}{\sigma^m}\right)$. We want to find an r_σ such that the mass of $\mathcal{B}_{r_\sigma}$ is greater than $1 - \kappa$. For any set A, we may write:

$$\mu_\sigma(A) = \frac{1}{Z_\sigma} \int \mathbf{1}_A \exp\left(-\frac{V}{\sigma}\right)$$
$$= \frac{Z_{2\sigma}}{Z_\sigma} \times \frac{1}{Z_{2\sigma}} \int \mathbf{1}_A \exp\left(-\frac{V}{2\sigma} - \frac{V}{2\sigma}\right)$$
$$= \frac{Z_{2\sigma}}{Z_\sigma} \times \int \mathbf{1}_A \exp\left(-\frac{V}{2\sigma}\right) d\mu_{2\sigma}.$$

If V takes large values on A, we can get a good bound:

$$\mu_\sigma(A) \leq \frac{Z_{2\sigma}}{Z_\sigma} \exp\left(-\frac{\inf_A V}{2\sigma}\right) \mu_{2\sigma}(A).$$

We get rid of the $\mu_{2\sigma}(A)$ by roughly bounding it by 1. Then we use the growth hypothesis on V (2), with $A = \mathcal{B}_{r_\sigma}^c$. In this case:

$$\inf_A V \geq \ln(r_\sigma)^{m_V}.$$

We fix an $m' \in]m, 1 + m_V[$, and choose:

$$r_\sigma = \exp\left(\left(\frac{1}{\sigma}\right)^{(m'-1)/m_V}\right),$$

which ensures:

$$\inf_A V \geq \left(\frac{1}{\sigma}\right)^{(m'-1)},$$
$$\mu_\sigma(\mathcal{B}_{r_\sigma}^c) \leq \frac{Z_{2\sigma}}{Z_\sigma} \exp\left(-\frac{1}{2\sigma^{m'}}\right)$$

The asymptotic behavior of Z_σ (*cf.* annex C.1) implies that $Z_{2\sigma}/Z_\sigma$ converges, and since $m' > m$,

$$\mu_\sigma(\mathcal{B}_{r_\sigma}^c) \leq \exp\left(-\frac{1}{\sigma^m}\right)$$

for σ small enough. This shows that r_σ satisfies the condition we wanted.

We may now end the proof of the theorem. Coming back to the measure-capacity inequality (14), we note that R_σ, n_σ and N_σ all behave like r_σ to a certain power (for R_σ we use hypothesis 5, and n_σ, N_σ are just a number of cubes of fixed radius in the big cubes of side length r_σ and R_σ). Therefore, there exists a C such that

$$\forall A, \mu_\sigma(A) \geq \kappa(\sigma) \quad \Longrightarrow \quad \mu_\sigma(A) \leq \frac{r_\sigma^C}{\sigma^2} \exp\left(\frac{D^\star}{\sigma}\right) \mathrm{Cap}_{\mu_\sigma}(A). \tag{15}$$

The value of r_σ and the fact that $m' - 1$ is strictly less than m_V makes $\exp(D^\star/\sigma)$ the biggest term, so that, up to a slight increase of $D^\star$,

$$\forall A, \mu_\sigma(A) \geq \kappa(\sigma) \quad \Longrightarrow \quad \mu_\sigma(A) \leq \exp\left(\frac{D^\star}{\sigma}\right) \mathrm{Cap}_{\mu_\sigma}(A).$$

This inequality, thanks to theorem 28 below, implies precisely the one-point weak Poincaré inequality we claimed in theorem 17.

4 A measure-capacity criterion for one-point weak Poincaré inequalities

4.1 Definitions

In this section we study the interplay between weak Poincaré inequalities and measure-capacity inequalities. Let us start by recalling exactly what a weak Poincaré inequality is.

24 Definition (M. Röckner and F.Y. Wang, [RW01]). *Let μ be a measure and $\mathcal{N}$ be a norm, stronger than the $L^2(\mu)$ norm. The measure μ is said to satisfy a* weak Poincaré inequality *for the norm $\mathcal{N}$ if there exists a decreasing positive function α, defined on $\mathbb{R}_+^\star$ such that:*

$$\forall f \in L^2(\mu), f \text{ such that } \mu f = 0, \forall r > 0, \mu(f^2) \leq \alpha(r)\mathcal{E}(f,f) + r\mathcal{N}(f)^2$$

If this holds, α will be called a compensating function.

Remark (on means and medians). *The original statement on weak Poincaré inequalities involves functions recentred by their mean value $\mu(f)$, and an L^∞ norm. However, the approach by measure-capacity inequalities developed in [BCR05, BCR] works with functions recentred by their median m_f. When the norm is the sup norm, it is easy to go from one to the other: the three quantities* $\mathrm{osc}(f)$, $\|f - m_f\|_\infty$ *and* $\|f - \mu(f)\|_\infty$ *are within (universal) bounds of each other.*

*Since we need to work with another norm, we will show that we can still go from $\mathcal{N}(f - m_f)$ to $\mathcal{N}(f - \mu f)$ (*cf. *equation* (19) *in annex A).*

This is equivalent to the slightly modified definition:

25 Proposition. *A weak Poincaré inequality holds if and only if:*

$$\forall r > 0, \exists c_r, \forall f \in L^2(\mu), \qquad \mu(f) = 0 \implies \mu(f^2) \leq c_r \mathcal{E}_\mu(f,f) + r\mathcal{N}(f)^2. \tag{16}$$

If the inequality holds for a given couple (r, c_r), we will say that μ satisfies a one-point weak Poincaré inequality.

Therefore the weak Poincaré inequality holds if and only if a one-point inequality holds for each point r.

Proof. The only thing to check is that we can deduce the inequality of the definition from (16). To each r, we associate c_r according to (16). Then we just define $\alpha(r) = \inf\{c_s; s \leq r\}$. The function α is decreasing. Now let f be a function in L^2 and $r > 0$. For any ϵ, we may find an $s \leq r$ such that:

$$c_s \leq \alpha(r) + \epsilon.$$

If we apply (16) with this s, we get (since $s \leq r$):

$$\begin{aligned} \mu(f^2) &\leq c_s \mathcal{E}(f) + s\mathcal{N}(f)^2 \\ &\leq \alpha(r)\mathcal{E}(f) + r\mathcal{N}(f)^2 + \epsilon\mathcal{E}(f). \end{aligned}$$

Since this is true for any ϵ, we may let it go to zero, and we have found a function α. □

We will be specifically interested in these inequalities for one special norm. We now define this norm and recall some of its properties, without proofs. For a short introduction (with the results we need here), see *e.g.* [Ale04]; for an extensive treatment we refer to [RR91].

Let ϕ, ψ be defined on $\mathbb{R}_+$ by $\psi(x) = x\log(1+x)$, $\phi(x) = \psi(x^2)$. For any measurable f, define the *Orlicz norm* (usually called the *Luxembourg norm*; there is another natural norm on the Orlicz space, but we won't need it here) of f to be:

$$\|f\|_\phi = \inf\left\{\lambda, \int \phi\left(\frac{|f|}{\lambda}\right) \leq 1\right\}.$$

Note that, with this definition, $\|1\|_\phi \neq 1$. The set of functions f for which this norm is finite is denoted L_ϕ, it is a vector space, and it is complete for the Orlicz norm. In the same way, if $\psi^\star, \phi^\star$ are the convex dual functions of ψ, ϕ, we may define the corresponding Orlicz spaces. It is easily seen that for every positive f, $\|f^2\|_\psi = \|f\|_\phi$. The dual functions allow us to state the following Hölder-like property.

26 Proposition (Hölder-Orlicz). *If f, g are two measurable functions, respectively in L_ψ and $L_{\psi^\star}$, then fg is in L^1, and*

$$\left|\int fg d\mu\right| \leq 2\|f\|_\psi \|g\|_{\psi^\star}.$$

The constant 2 is necessary because we work with Luxembourg norms. To conclude this account on Orlicz norm, we recall here the norm of an indicator function:

27 Proposition. *Let A be a measurable set. Then $\mathbf{1}_A$ is in the Orlicz space L_ψ and:*

$$\|\mathbf{1}_A\|_{\psi^\star} = \hat{\psi}(\mu(A)),$$

where $\hat{\psi}(x) = \frac{1}{(\psi^\star)^{-1}(1/x)}$. Moreover, for all x sufficiently small, we have the following bound:

$$\hat{\psi}(x) \leq \frac{2}{\log(1/x)}.$$

Proof. Once again we refer to [Ale04, RR91] for the first result. The explicit bound on $\hat{\psi}$ follows easily from the bound $\psi^\star \leq xe^x$ and the definition of $\hat{\psi}$. □

4.2 Measure-capacity inequalities for large sets and one-point inequalities

Here we show the result which was used in the preceding section: if we can compare the measure and the capacity of large sets, we can deduce a one-point weak inequality.

28 Theorem. *Suppose that there exists $\kappa < 1/2$, and a real constant C_κ such that, for every set A whose measure is larger than κ, we have:*

$$\mathrm{Cap}_\mu(A) \geq C_\kappa \mu(A). \tag{17}$$

Then μ satisfies the one-point weak Poincaré inequality:

$$\mathbf{Var}_\mu(g) \leq \frac{c}{C_\kappa}\int |\nabla g|^2 d\mu + \kappa \operatorname{osc}^2(g),$$

where c is universal. We may replace the L^∞ norm by an Orlicz norm, in which case the inequality reads:

$$\mathbf{Var}_\mu(g) \leq \frac{c}{C_\kappa}\int |\nabla g|^2 d\mu + \hat{\psi}(\kappa)\|g - m_g\|_\phi^2.$$

Remark. *Note that if* (17) *holds for* all *sets, regardless of their measure, then μ satisfies a (strong) Poincaré inequality (since we may take $\kappa = 0$). This is well-known,* cf. *[BCR] and references therein. This characterization of a functional inequality in terms of a relation between measures and capacities of sets is in fact more general, and provides a way to compare many functional inequalities. For a detailed account on these questions, and links with isoperimetric properties, we refer to [BCR] (especially section 5).*

Proof. We follow the proof of theorem 2 in [BCR05] (which deals with the (full) weak inequality).

Let f be a function and m a median for f. We cut the space in half, according to whether f is greater than m or not; we denote by Ω_+, Ω_- the two sets. The integral may be written as:

$$\mathbf{Var}_\mu(f) \leq \int (f-m)^2 d\mu = \int_{\Omega_+} (f-m)^2 d\mu + \int_{\Omega_-} (f-m)^2 d\mu.$$

We will show how to deal with the leftmost integral, the other one being similar.

$$c = \inf\{t \geq 0, \mu(g^2 > t) < \kappa\}.$$

If c is zero, then $\mu(g > 0)$ is less than κ, and:

$$\int_{\Omega_+} g^2 d\mu \leq \begin{cases} \kappa \sup g^2 \text{ in the } L^\infty \text{ case,} \\ \hat{\psi}(\kappa)\|f - m\|_\phi^2 \text{ in the Orlicz case,} \end{cases}$$

so the inequalities we are looking for hold in the half-space Ω_+.

Thus we need only consider the case where c is strictly positive. By a continuity argument (μ will always have a density), we can find a set Ω_0 such that $\mu(\Omega_0) = \kappa$ and $\{g^2 > c\} \subset \Omega_0 \subset \{g^2 \geq c\}$. We fix a $\rho > 1$, and introduce the level sets $\Omega_k = \{g^2 \geq \frac{c}{\rho^k}\}$. We decompose the integral over these sets:

$$\begin{aligned} \int_{\Omega_+} g^2 &= \int_{\Omega_0} g^2 d\mu + \sum_{k\geq 1} \int_{\Omega_k \setminus \Omega_{k-1}} g^2 d\mu \\ &\leq \int_{\Omega_0} g^2 d\mu + \sum_{k\geq 1} \frac{c}{\rho^{k-1}} \left(\mu(\Omega_k) - \mu(\Omega_{k-1})\right) \end{aligned}$$

The sum is dealt with thanks to an Abel transform:

$$\begin{aligned} &\sum_{k\geq 1} \frac{1}{\rho^{k-1}} (\mu_k - \mu_{k-1}) \\ &\qquad = \sum_{k\geq 1} \frac{\mu_k}{\rho^{k-1}} - \sum_{k\geq 0} \frac{\mu_k}{\rho^k} \\ &\qquad = \sum_{k\geq 1} \mu_k \left(\frac{1}{\rho^{k-1}} - \frac{1}{\rho^k}\right) - \mu_0. \end{aligned}$$

This is where we do not follow [BCR05]: since we simply suppose an inequality between capacity and measure, we can get rid of the μ_0 and write

$$\sum_{k\geq 1} \frac{1}{\rho^{k-1}} (\mu_k - \mu_{k-1}) \leq (\rho - 1) \sum_{k\geq 1} \frac{\mu_k}{\rho^k}.$$

The rest of the proof follows the same line as in [BCR05] — at this point, we use the measure-capacity inequality on each set Ω_k. They are designed to have their measure bigger than κ, so that we may apply our hypothesis:

$$\mu_k \leq \frac{1}{C_\kappa} \operatorname{Cap}(\Omega_k).$$

Now, to bound the capacity from above, we apply the definition with well-chosen functions g_k:

$$g_k = \min\left(1, \left(\frac{g - \sqrt{c\rho^{-k-1}}}{\sqrt{c\rho^{-k}} - \sqrt{c\rho^{-k-1}}}\right)_+\right)$$

This entails:

$$\begin{aligned} \mu_k &\leq \frac{1}{C_\kappa} \int |\nabla g_k|^2 d\mu \\ &\leq \frac{\rho^{k+1}}{C_\kappa c(\sqrt{\rho} - 1)^2} \int_{\Omega_k \setminus \Omega_{k-1}} |\nabla g|^2 d\mu. \end{aligned}$$

Summing over k, we get:

$$\int_{\Omega_+} g^2 d\mu \leq \int_{\Omega_0} g^2 d\mu + \frac{\rho(\rho - 1)}{C_\kappa(\sqrt{\rho} - 1)^2} \int |\nabla g|^2 d\mu.$$

We may now choose ρ; the (non optimal) choice $\rho = 4$ gives:

$$\int_{\Omega_+} g^2 d\mu \leq \int_{\Omega_0} g^2 d\mu + \frac{12}{C_\kappa} \int |\nabla g|^2 d\mu.$$

The only thing left to do is to take care of the integral on Ω_0. This is done with an Hölder-like inequality. In the Orlicz norm case, for example, we write:

$$\int_{\Omega_0} g^2 d\mu \leq 2\|g^2\|_\psi \|\mathbf{1}_{\Omega_0}\|_{\psi^\star}$$
$$\leq 2\|(f-m)^2\|_\psi \hat{\psi}(\kappa) \leq 2\hat{\psi}(\kappa)\|f-m\|_\phi^2.$$

thanks to the Hölder-Orlicz inequality and the relation between ϕ and ψ (see the beginning of this section). □

4.3 Weak inequalities for different norms

To conclude this section, let us state a corollary to the previous result, and prove that weak Poincaré inequalities for many different norms are in fact equivalent. Moreover, if a compensating function is known for one norm, we can immediately deduce a function for another norm; this result was used in the one dimensional case (section 2) where the explicit Hardy-like criteria were known for the L^∞ norm.

29 Theorem. *Let ϕ, ψ be two Young functions, with $\phi(x) = \psi(x^2)$. A measure μ satisfies a weak Poincaré inequality with the L^∞ norm if and only if it satisfies one with the Orlicz norm $\|\cdot\|_\phi$.*

Moreover, if β is a compensating function for the L^∞ norm, then the following function may be chosen for the Orlicz norm:

$$\alpha(s) = \frac{c}{4}\beta\left(\frac{1}{4}\hat{\psi}^{-1}\left(\frac{s}{2}\right)\right),$$

where c is universal (and the same as in the preceding result).

Proof. First, let us introduce a few notations. We will denote by $\mathbf{M\text{-}C}(\kappa, C(\kappa))$ the following comparison between measure and capacity:

$$\forall A, \mu(A) > \kappa \implies \mathrm{Cap}(A) \geq C(\kappa)\mu(A).$$

Similarly, $\mathbf{PWP}(r, C(r), \mathcal{N})$ will denote the one-point weak Poincaré inequality for a norm $\mathcal{N}$ with constants $(r, C(r)$, and $\mathbf{WP}(\alpha, \mathcal{N})$ will be the (full) weak inequality, with a norm $\mathcal{N}$ and a compensating function α. In the previous section, we showed:

$$\mathbf{M\text{-}C}(\kappa, C(\kappa)) \implies \mathbf{PWP}\left(\kappa, \frac{c}{C_\kappa}, \|\cdot\|_\infty\right),$$
$$\mathbf{M\text{-}C}(\kappa, C(\kappa)) \implies \mathbf{PWP}\left(2\hat{\psi}(\kappa), \frac{c}{C_\kappa}, \|\cdot\|_\phi\right).$$

Going the other way around is easy. Indeed, suppose that $\mathbf{PWP}(r, C(r), \|\cdot\|_\infty)$ holds. Let A be a set whose measure is less than $1/2$, but greater then $4r$. Let g be any function which may appear in the definition of the capacity of A (*cf.* (10)), and let m_g be a median of g. Then:

$$\mathbf{Var}_\mu g \leq C_r \int |\nabla g|^2 d\mu + r\|g - m_g\|_\infty.$$

Without loss of generality, we suppose that $0 \leq g \leq 1$, so that the L^∞ norm is bounded by 1. Moreover, r is less than $\mu(A)/4$, and the variance on the left hand side is bounded below by $(1/2)\min(\mu(A), 1/2) \geq (\mu(A)/2)$ (by the same argument used previously, during the proof of proposition 22). This entails:

$$\frac{\mu(A)}{2} \leq C_r \int |\nabla g|^2 d\mu + \frac{\mu(A)}{4}.$$

This immediately implies the measure capacity inequality $\mathbf{M\text{-}C}(4r, 4/C_r)$.

If we now try to derive an inequality with an Orlicz norm starting from one with an L^∞ norm, we just translate them in terms of measure and capacity:

$$\mathbf{PWP}(r, C_r, \|\cdot\|_\infty) \implies \mathbf{M\text{-}C}(4r, 4/C_r)$$
$$\implies \mathbf{PWP}(2\hat{\psi}(4r), \frac{cC_r}{4}).$$

If we are looking for a full weak Poincaré inequality, we fix an s, and define $r = (1/4)\hat{\psi}^{-1}(s/2)$. We may then apply $\mathbf{PWP}(r, \beta(r))$ to obtain:

$$\mathbf{PWP}(s, c\beta(r)/4, \|\cdot\|_\infty).$$

Since s is arbitrary, this concludes the proof. □

A Orlicz norms, entropy and centering

The proof of weak Poincaré inequalities starting from measure-capacity comparisons for an Orlicz norm leads us to consider norms of functions recentered by their median. In fact, what one obtains when applying these criteria is of the form:

$$\mathbf{Var}_\mu f^2 \leq \beta(s)\mathcal{E}(f) + s\|f - m_f\|_\phi^2,$$

where m_f is a median for f. The aim of this section is to bound this term by more tractable quantities (we will use an entropy and a moment).

More precisely we prove the following result:

30 Lemma. *There exists a C such that, for any positive f and any probability measure μ, the following holds:*

$$\|f - m_f\|_\phi^2 \leq C\left(\mathbf{Ent}_\mu(f^2) + 3\mathbf{E}_\mu(f^2)\right).$$

The proof is done in several steps, and borrows several arguments from [BG99]. First of all, we get rid of the median and replace it by a mean value.

$$\begin{aligned}\|f - m_f\|_\phi &\leq \|f - \mu f\|_\phi + \|\mu f - m_f\|_\phi \\ &\leq \|f - \mu f\|_\phi + |\mu f - m_f|. \end{aligned} \tag{18}$$

Let us consider the last term.

$$\mu f - m_f = \int f(x)d\mu - m_f = \int (f - m_f)_+ d\mu - \int (f - m_f)_- d\mu,$$

where the integrals are both positive. The absolute value of the left hand side may then be bounded above:

$$|\mu f - m_f| \leq \max\left(\int (f - m_f)_+ d\mu, \int (f - m_f)_- d\mu\right)$$

Each of the arguments in the max can be controlled by Hölder's inequality.

$$\begin{aligned}\int (f - m_f)_+ d\mu &= \int (f - m_f)\mathbf{1}_{f > m_f} d\mu \leq \|f - m_f\|_2 \|\mathbf{1}_{f > m_f}\|_2 \\ &\leq \frac{1}{\sqrt{2}}\|f - m_f\|_2 && \text{(since } \mu(f > m_f) < 1/2) \\ &\leq \frac{1}{\sqrt{2}}\frac{\sqrt{5}}{2}\|f - m_f\|_\phi && (\textit{cf.}\ \text{[BG99], lemma 4.3})\end{aligned}$$

Coming back to (18), we get:

$$\|f - m_f\|_\phi \leq \|f - \mu_f\|_\phi + |\mu f - m_f| \leq \|f - \mu f\|_\phi + \sqrt{\frac{5}{8}}\|f - m_f\|_\phi.$$

Since $\sqrt{\frac{5}{8}} \leq 1$, we may put it on the other side to get:

$$\|f - m_f\|_\phi \leq C\|f - \mu f\|_\phi \tag{19}$$

where $C = (1 - \sqrt{\frac{5}{8}})^{-1}$ is universal.

The next step is to bound the Orlicz norm by an entropy. Once again, we use a result from Bobkov and Götze ([BG99]):

$$\|f - \mu f\|_\phi^2 \leq \frac{3}{2}\sup_{a \in \mathbb{R}} \mathbf{Ent}_\mu((f + a)^2).$$

Since we would like to deal only with the entropy of f^2, we try to compare the entropies of translated functions. Rothaus' lemma tells us:

$$\mathbf{Ent}_\mu((f + a)^2) \leq \mathbf{Ent}_\mu(\tilde{f}^2) + 2\mathbf{Var}_\mu(f),$$

where $\tilde{f}$ is the centered function $f - \mu f$. The only thing left to do is to bound the entropy of the square of this centered function. This is done in the following lemma.

31 Lemma. *Let f be a positive function, and $\tilde{f} = f - \mu f$. Then the following holds:*

$$\mathbf{Ent}_\mu(\tilde{f}^2) \leq \mathbf{Ent}_\mu(f^2) + \int f^2 d\mu.$$

Proof. Both sides of the equation are homogeneous (of order two), so we may as well suppose $\int f^2 d\mu = 1$. We rewrite the left hand side.

$$\begin{aligned}\mathbf{Ent}_\mu(\tilde{f}^2) &= \int \tilde{f}^2 \log(\tilde{f}^2) d\mu - \mathbf{E}_\mu(\tilde{f}^2)\log(\mathbf{E}_\mu(\tilde{f}^2)) \\ &= \int \tilde{f}^2 \log(\tilde{f}^2) d\mu - \mathbf{Var}_\mu(f)\log(\mathbf{Var}_\mu(f)).\end{aligned}$$

The second term is easily dealt with. Indeed, since $\int f^2 = 1$, $\mathbf{Var}_\mu f$ must be between 0 and 1. Since $x \mapsto |x\log(x)|$ is bounded by $1/e$ on this interval, one can write:

$$\mathbf{Ent}_\mu(\tilde{f}^2) \leq \int (\tilde{f}^2 \log(\tilde{f}^2) d\mu + \frac{1}{e}.$$

We decompose the integral in two parts, according to whether f is less than 1 or not.

$$\begin{aligned}\mathbf{Ent}_\mu(\tilde{f}^2) &\leq \int \tilde{f}^2 \log(\tilde{f}^2)\mathbf{1}_{|\tilde{f}|\leq 1} d\mu + \int \tilde{f}^2 \log(\tilde{f}^2)\mathbf{1}_{|\tilde{f}|>} 1 d\mu + \frac{1}{e} \\ &\leq \int \tilde{f}^2 \log(\tilde{f}^2)\mathbf{1}_{|\tilde{f}|>1} d\mu + \frac{1}{e},\end{aligned}$$

since the first term is less than 0. Now, on the set where $|\tilde{f}|$ exceeds one, f must be above its mean: f is indeed positive, and since $\int f^2 d\mu = 1$, μf must be in $[0,1]$. So $|f - \mu f|$ may be greater than 1 only when f itself is greater than 1. This shows that, on $\{|\tilde{f}| > 1\}$,

$$1 \leq \tilde{f} = f - \mu f \leq f.$$

Since $x \mapsto x\log(x)$ increases on $[1, \infty)$, we have:

$$\begin{aligned}\mathbf{Ent}_\mu(\tilde{f}^2) &\leq \int \tilde{f}^2 \log(\tilde{f}^2)\mathbf{1}_{|\tilde{f}>1|} d\mu + \frac{1}{e} \\ &\leq \int f^2 \log(f^2)\mathbf{1}_{|\tilde{f}>1|} d\mu + \frac{1}{e}.\end{aligned}$$

At this point, remark that on $\{f > 1\}$, $f^2\log(f^2)$ is positive, and since $\mathbf{1}_{|\tilde{f}|>1} \leq \mathbf{1}_{f>1}$,

$$\begin{aligned}\mathbf{Ent}_\mu(\tilde{f}^2) &\leq \int f^2 \log(f^2)\mathbf{1}_{f>1} d\mu + \frac{1}{e} \\ &\leq \mathbf{Ent}(f^2) - \int f^2 \log(f^2)\mathbf{1}_{f<1} d\mu + \frac{1}{e} \\ &\leq \mathbf{Ent}(f^2) + \frac{2}{e}.\end{aligned}$$

Since $\frac{2}{e} \leq 1$, the proof is complete. □

Gathering our results, we have shown that:

$$\begin{aligned}\|f - m_f\|_\phi^2 &\leq C\|f - \mu_f\|_\phi^2 && \text{(inequality (19))} \\ &\leq \frac{3C}{2}\sup_{a\in\mathbb{R}} \mathbf{Ent}((f+a)^2) && \text{(Bobkov and Götze's lemma)} \\ &\leq \frac{3C}{2}\left(\mathbf{Ent}(\tilde{f}^2) + 2\mathbf{Var}_\mu(f)\right) && \text{(Rothaus' lemma)} \\ &\leq \frac{3C}{2}\left(\mathbf{Ent}(f^2) + 3\mathbf{E}_\mu(f^2)\right). && \text{(lemma 31)}\end{aligned}$$

The last line is precisely the result we claimed in lemma 30.

B A moment bound

In this annex we prove lemma 8. The proof mainly follows the one in Miclo's doctoral dissertation, with a few changes to accomodate our hypotheses.

B.1 Outline of the proof

We need to introduce some notation.

For $\epsilon > 0$, we denote by $\mathcal{L}_\epsilon$ the generator of the diffusion at fixed temperature ϵ:

$$\mathcal{L}_\epsilon = \frac{\epsilon}{2}\Delta - \frac{1}{2}\nabla V \nabla \cdot .$$

We will need a smooth version of a step function; we call it f and suppose that it satisfies:

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ \exp\left(-\exp\left(\frac{1}{x}\right)\right) & \text{on } [0,1], \\ 1 & \text{on } [2,\infty[. \end{cases}$$

We recall the hypotheses on V:

- It goes to infinity at infinity,
- its gradient ∇V is bounded, and
- its Laplacian ΔV is negative for large x.

Note that, since V is continuous, there must be an R such that ΔV is negative whenever $V(x) \geq R$.

Finally, let g be an increasing function, going to zero at zero.

The idea of the proof is that, as time goes by, the value of V at X_t has a typical scale, namely $\frac{1}{g(\sigma(t))}$, for a function g to be made precise later, so that when we try to estimate $\mathbf{E}(V^p(X_t))$, we only have to take into account the small values of V.

More precisely, let $\rho_\epsilon(\cdot) = f\left(g(\epsilon)V(\cdot) - (R+1)\right)$. This is a smooth approximation of $\mathbf{1}_{V \geq \frac{R}{g(\epsilon)}}$. We may bound the expectation of $V^p(X_t)$:

$$\begin{aligned} \mathbf{E}[V^p(X_t)] &= \mathbf{E}[V^p \rho_{\sigma(t)}(X_t)] + \mathbf{E}[V^p(1-\rho_{\sigma(t)}(X_t))] \\ &\leq \mathbf{E}[V^p \rho_{\sigma(t)}(X_t)] + \left(\frac{R+3}{g(\sigma(t))}\right)^p. \end{aligned} \tag{20}$$

To bound the first term, we use the explicit expression of the generator. Intuitively, we write, for $h_t = V^p \rho_{\sigma(t)}$:

$$\frac{d}{dt}(P_t h_t) = P_t \mathcal{L}_{\sigma(t)} h_t + P_t(\frac{d}{dt} h_t),$$

and integrate between two times t and t'. To ensure that everything exists, we use the stopping time $T_k = \inf\{t, V(X_t) \geq k\}$. We get:

$$\begin{aligned} \mathbf{E}[h_{t\wedge T_k}(X_{t\wedge T_k})] &= \mathbf{E}[h_{t'\wedge T_k}(X_{t'\wedge T_k})] \\ &+ \mathbf{E}[\int_{t'\wedge T_k}^{t\wedge T_k} \mathcal{L}_{\sigma(s)}(h_s)(X_s)ds \\ &+ \mathbf{E}[\int_{t'\wedge T_k}^{t\wedge T_k} \sigma'(s)g'(\sigma(s))f'\left(g(\sigma(s))V(X_s) - (R+1)\right)V^{p+1}(X_s)ds. \end{aligned}$$

Since V is positive, f and g increasing and σ decreases, the whole last term is negative. We try to estimate the second one, and study $\mathcal{L}_{\sigma(s)}h_s(X_s)$.

32 Lemma. *Let us define $\varphi : x \mapsto x\log^2(x)$. There exists an M and a time t' (which may depend on p and on the initial law) such that:*

$$\forall t \geq t', \forall x, \quad \mathcal{L}_{\sigma(t)}(h_t)(X_t) \leq \exp\left(-\frac{M}{\varphi\left(\sigma(t)g(\sigma(t))\right)}\right).$$

We postpone the proof and finish the argument. The inequality dictates the choice of g: $g = \ln(1/\cdot)^{-3}$ guarantees

$$\sigma(t)g(\sigma(t)) = \frac{1}{\ln(t)(\ln\ln(t))^3},$$
$$\varphi(\sigma(t)g(\sigma(t))) = \frac{\ln^2\left(1/\ln(t)(\ln\ln(t))^3\right)}{\ln(t)(\ln\ln(t))^3} = \frac{\ln^2\left(\ln(t)(\ln\ln(t))^3\right)}{\ln(t)(\ln\ln(t))^3}.$$

Indeed, the upper bound on the generator then becomes

$$\begin{aligned}\mathcal{L}_{\sigma(t)}(h_t)(X_t) &\leq \exp\left(-\frac{M}{\varphi(\sigma(t)g(\sigma(t))}\right)\\ &\leq \exp\left(-M\ln(t)\times\frac{(\ln\ln(t))^3}{\ln^2\left(\ln(t)(\ln\ln(t))^3\right)}\right).\end{aligned}$$

Since the ratio $(\ln\ln(t))^3/(\ln^2(\ln(t)\ln\ln(t)^3))$ goes to infinity, it eventually exceeds $2/M$, so that for t big enough,

$$\mathcal{L}_{\sigma(t)}(h_t)(X_t) \leq \exp\left(-2\ln(t)\right).$$

Going back to the bound on the expected value we were looking for, the two previous arguments imply:

$$\mathbf{E}[h_{t\wedge T_k}(X_{t\wedge T_k})] \leq \mathbf{E}[h_{t'\wedge T_k}(X_{t'\wedge T_k})] + \int_{t'}^{\infty}\exp\left(-2\ln(t)\right).$$

We succeeded in making the last integral finite. We can then let K go to infinity, and since t' is fixed, we get the existence of a constant M_p (which depends on p and on the initial law) such that:

$$\mathbf{E}[h_t(X_t)] \leq M.$$

Plugging this back into inequality (20) yields:

$$\mathbf{E}[V^p(X_t)] \leq M + \left(\frac{R+3}{g(\sigma(t))}\right)^p.$$

The expression of g shows that, for a new constant M:

$$\mathbf{E}[V^p(X_t)] \leq M(\sigma(t)\ln(t)(\ln\ln(t))^3)^p,$$

and the result is proved.

B.2 An estimate on the generator

We now turn to the proof of lemma 32. We have to bound $\mathcal{L}_\epsilon(\rho_\epsilon V^p)(x)$, and our first step will be to give a more explicit expression of this quantity. We will need the derivatives of $\rho_\epsilon(x)$. To alleviate notations, we will write $y = y(x,\epsilon) = g(\epsilon)V(x) - (R+1)$.

$$\begin{aligned}\rho_\epsilon(x) &= f(g(\epsilon)V(x) - (R+1)) = f(y),\\ \nabla\rho_\epsilon(x) &= g(\epsilon)f'(y)\nabla V(x),\\ \Delta\rho_\epsilon(x) &= g(\epsilon)^2 f''(y)|\nabla V|^2 + g(\epsilon)f'(y)\Delta V.\end{aligned}$$

The quantity we would like to estimate is

$$\mathcal{L}_\epsilon(\rho_\epsilon V^p)(x) = \rho_\epsilon\mathcal{L}_\epsilon V^p(x) + \epsilon\langle\nabla\rho_\epsilon, \nabla V^p\rangle(x) + V^p\mathcal{L}_\epsilon\rho_\epsilon(x)$$

We consider three cases, according to the value of $V(x)g(\epsilon)$.

V is small: $V(x)g(\epsilon) \in [0, R+1]$ On this interval, ρ_ϵ vanishes, so $\mathcal{L}_\epsilon(\rho_\epsilon)$ is zero.

V **is large.** Let λ be a strictly positive real, to be fixed later on. We consider the case where $V(x)g(\epsilon) \in [R+1+\lambda, \infty)$, which may be rewritten as: $y \in [\lambda, \infty)$. We develop the expression of $\mathcal{L}_\epsilon(\rho_\epsilon V^p)$.

$$\mathcal{L}_\epsilon(\rho_\epsilon V^p)(x) = \rho_\epsilon \mathcal{L}_\epsilon V^p(x) + \epsilon g(\epsilon) f'(y) \times pV^{p-1}|\nabla V|^2 + V^p \left(\frac{1}{2}\epsilon \Delta \rho_\epsilon - \frac{1}{2}\langle \nabla \rho_\epsilon, \nabla V\rangle\right).$$

We compute the derivatives of ρ_ϵ and put together the terms involving $|\nabla V|^2$.

$$\begin{aligned}\mathcal{L}_\epsilon(\rho_\epsilon V^p)(x) &= \rho_\epsilon \mathcal{L}_\epsilon V^p(x) + \left(\epsilon g(\epsilon) f'(y) p V^{p-1} + V^p \left(\frac{1}{2}\epsilon g(\epsilon)^2 f''(y) - \frac{1}{2} g(\epsilon) f'(y)\right)\right)|\nabla V|^2 \\ &\quad + \frac{1}{2}\epsilon g(\epsilon) f'(y) \Delta V \\ &= A + B + C.\end{aligned}$$

Since $V \times g(\epsilon) \geq R$, $V \geq R$. We already noted that R may be chosen so that, if V is bigger than R, ΔV is less than zero, and this makes the third term C negative. The term B can be rewritten as:

$$\begin{aligned}B &= \left(\epsilon g(\epsilon) f'(y) p V^{p-1} + V^p \left(\frac{1}{2}\epsilon g(\epsilon)^2 f''(y) - \frac{1}{2} g(\epsilon) f'(y)\right)\right)|\nabla V|^2 \\ &= V^p g(\epsilon)\left(\left(\frac{p\epsilon}{V} - \frac{1}{2}\right) f'(y) + \frac{1}{2}\epsilon g(\epsilon) f''(y)\right)|\nabla V|^2. \qquad (21)\end{aligned}$$

We add another condition on f: it should be concave when y is near 2 (*e.g.* on $[\frac{3}{2}, 2]$). On $[\lambda, 3\lambda/2]$, f''/f' is bounded — let M be a bound. This entails:

$$\forall y \geq \lambda, \qquad f''(y) \leq M f' y).$$

Coming back to B, we deduce:

$$B \leq \left(\frac{p\epsilon}{V} + \frac{M\epsilon g(\epsilon)}{2} - \frac{1}{2}\right) f'(y) g(\epsilon) V^p |\nabla V|^2.$$

The term between brackets is negative, uniformly in x as soon as ϵ is small enough.

Finally, the first term $A = \mathcal{L}_\epsilon V^p$ is also negative:

$$\begin{aligned}A &= \frac{\epsilon}{2}\Delta(V^p) - \frac{1}{2}\langle \nabla V, \nabla(V^p)\rangle \\ &= \frac{\epsilon}{2}\left(p(p-1)V^{p-1}|\nabla V|^2 + pV^{p-1}\Delta V\right) - \frac{1}{2}V^{p-1}|\nabla V|^2 \\ &\leq \left(\frac{p(p-1)\epsilon}{2} - \frac{1}{2}\right)|\nabla V|^2.\end{aligned}$$

Once more, the term between brackets is negative when ϵ is small. To conclude, for any λ, there exists an ϵ_0 such that:

$$\forall \epsilon < \epsilon_0, \forall x, \qquad V(x)g(\epsilon) \geq R + 1 + \lambda \implies \mathcal{L}_\epsilon(\rho_\epsilon V^p) \leq 0.$$

V **is of the order of** $R/g(\epsilon)$**.** This last case is that where $g(\epsilon)V(x) \in [R+1, R+1+\lambda]$. Let us reuse the decomposition $\mathcal{L}_\epsilon(\rho_\epsilon V^p) = A + B + C$ from the previous paragraph. The same reasoning applies for A and C, and they are both negative, so it suffices to get a bound on B. From (21):

$$B = \left(\left(\frac{p\epsilon}{V} - \frac{1}{2}\right) f'(y) + \frac{1}{2}\epsilon g(\epsilon) f''(y)\right) g(\epsilon) V^p |\nabla V|^2.$$

If we choose R sufficiently big and ϵ small enough, the quantity between brackets in front of $f'(y)$ is less than $1/4$.

$$B \leq \left(-\frac{1}{4} f'(y) + \frac{1}{2}\epsilon g(\epsilon) f''(y)\right) g(\epsilon) V^p |\nabla V|^2.$$

Recall that $f = \exp(-\tau)$, where $\tau(y) = \exp(1/y)$. This implies:

$$\begin{aligned}B &\leq \left(\frac{1}{4}\tau' f + \frac{1}{2}\epsilon g(\epsilon)(-\tau'' f + (\tau')^2 f)\right) g(\epsilon) V^p |\nabla V|^2 \\ &\leq \frac{1}{2}\left(\frac{1}{2}\tau' f + \epsilon g(\epsilon)(\tau'(y))^2 f(y)\right) g(\epsilon) V^p |\nabla V|^2\end{aligned}$$

Define $h_\epsilon = \frac{1}{2}\tau' f + \epsilon g(\epsilon)\tau'^2 f$. We study it by differentiating:

$$h'_\epsilon = \left(\frac{1}{2}\tau'' - \frac{1}{2}\tau'^2 + 2\epsilon g(\epsilon)\tau'\tau'' - \epsilon g(\epsilon)\tau'^3\right) f.$$

The explicit expression of τ ensures:

$$\exists \lambda \forall y \in [0,\lambda] \qquad 0 \leq \tau''(y) \leq \frac{1}{4}\tau'^2(y).$$

This λ does not depend on ϵ. This can be used to bound h'_ϵ from below:

$$\begin{aligned} h'_\epsilon(y) &\geq \left(-\frac{1}{2}\tau'^2(y) + \frac{1}{2}\epsilon g(\epsilon)\tau'(y)^3 - \epsilon g(\epsilon)\tau'(y)^3\right) f(y) \\ &\geq \left(-\frac{1}{2} - \frac{1}{2}\epsilon g(\epsilon)\tau'(y)\right)\tau'(y)^2 f(y). \end{aligned}$$

Let $y_{1,\epsilon}$ be the solution of the equation: $-1 - \epsilon g(\epsilon)\tau'(y) = 0$. When ϵ is small, $y_{1,\epsilon}$ will be less than λ, and the monotonicity of τ' will give:

$$\forall y \leq y_{1,\epsilon}, \quad h'_\epsilon(y) \geq 0.$$

Similarly, h'_ϵ can be bounded above:

$$\begin{aligned} h'_\epsilon &\leq \left(\frac{1}{8}\tau'^2(y) - \frac{1}{2}\tau'^2(y) - \epsilon g(\epsilon)\tau'^3(y)\right) f(y) \\ &\leq \left(-\frac{3}{8} - \epsilon g(\epsilon)\tau'(y)\right)\tau'(y)^2 f(y). \end{aligned}$$

Now, let $y_{2,\epsilon}$ be the root of $-\frac{3}{8} - \epsilon g(\epsilon)\tau'(y) = 0$. Once more, when ϵ is small, $y_{2,\epsilon}$ falls within $[0,\lambda]$. We deduce:

$$\forall y \in [y_{2,\epsilon}, \lambda], h'_\epsilon(y) \leq 0.$$

We now know the h_ϵ increases on $[0, y_{1,\epsilon}]$, and decreases on $[y_{2,\epsilon}, \lambda]$, so that its maximum must be reached somewhere between these two points. More precisely, whenever ϵ is less than some ϵ_0, it holds that

$$\exists y_\epsilon \in [y_{1,\epsilon}, y_{2,\epsilon}], \forall y \in [0,\lambda], \quad h_\epsilon(y) \leq h_\epsilon(y_\epsilon).$$

The bounds on y_ϵ, the fact that τ decreases and the equations defining $y_{1,\epsilon}, y_{2,\epsilon}$ allow us to conclude:

$$\begin{aligned} \forall y \leq \lambda, \qquad h_\epsilon(y) &\leq \left(\frac{1}{2}\tau'(y_\epsilon) + \epsilon g(\epsilon)\tau'(y_\epsilon)^2\right) f(y_\epsilon) \\ &\leq \left(\frac{1}{2}\tau'(y_{2,\epsilon}) + \epsilon g(\epsilon)\tau'(y_{1,\epsilon})^2\right) f(y_{2,\epsilon}) \\ &\leq \left(-\frac{3}{16\epsilon g(\epsilon)} + \frac{1}{\epsilon g(\epsilon)}\right) f(y_{2,\epsilon}) \\ &\leq \frac{1}{\epsilon g(\epsilon)} f(y_{2,\epsilon}). \end{aligned}$$

It remains to estimate $f(y_{2,\epsilon}) = \exp(-\tau(y_{2,\epsilon}))$. Since $y_{2,\epsilon}$ is defined as a solution of an equation involving τ', we would like to compare τ and τ'. The explicit expression of τ easily implies:

$$\ln(|\tau'(y)|) = \ln(y^{-2}) + \frac{1}{y} \geq \frac{1}{y},$$

therefore:

$$\tau(y) = y^2|\tau'(y)| \geq \frac{|\tau'(y)|}{\ln^2(|\tau'(y)|)}$$

Applying this for $y = y_{2,\epsilon}$, for which $|\tau'(y)| = 3/(8\epsilon g(\epsilon))$, entails:

$$\begin{aligned} \tau(y_{2,\epsilon}) &\geq \frac{3}{8\epsilon g(\epsilon)\ln^2(8\epsilon g(\epsilon)/3)} \\ &\geq \frac{3}{8\epsilon g(\epsilon)\ln^2(\epsilon g(\epsilon))}. \end{aligned}$$

Turning back to f, and defining $\varphi : x \mapsto x \ln^2(x)$, and $M = 3/8$, we have:

$$f(y_{2,\epsilon}) = \exp(-\tau(y_{2,\epsilon})) \leq \exp\left(-\frac{M}{\varphi(\epsilon g(\epsilon))}\right).$$

We now come back to the upper bound on B, and plug in the last equation.

$$B \leq \frac{1}{2} \times \frac{1}{\epsilon g(\epsilon)} \exp\left(-\frac{M}{\varphi(\epsilon g(\epsilon))}\right) g(\epsilon) V^p |\nabla V|^2.$$

Since we suppose that $V(x)g(\epsilon)$ belongs to $[R+1, R+2]$, we may bound V^p by $g(\epsilon)^{-p}$. We also supposed that ∇V is bounded, so that there exists an M' such that:

$$B \leq \frac{M'}{\epsilon g(\epsilon)^p} \exp\left(-\frac{M}{\varphi(\epsilon g(\epsilon))}\right).$$

Up to a slight change of the constant M in the exponential, we may neglect the pre-exponential term and write:

$$B \leq M'' \exp\left(-\frac{M}{\varphi(\epsilon g(\epsilon))}\right)$$

This concludes the proof.

C Regularity results and estimates on the process

C.1 An equivalent of the partition function

We recall here Laplace's method, which enable us to study the asymptotic behaviour of the partition function, *i.e.* the constant $Z_\sigma = \int \exp(-V/\sigma)dx$.

33 Theorem. *Let V be a function from $\mathbb{R}^d$ to $\mathbb{R}$, satisfying hypotheses 1 and 2 (V has a unique, well behaved, global minimum, and V goes to infinity at infinity rapidly enough). Then Z_σ exists, and the following holds:*

$$Z_\sigma \underset{\sigma \to 0}{\sim} \frac{1}{\sqrt{\det \mathbf{Hess} V}} \left(\frac{\sigma}{2\pi}\right)^{d/2}.$$

To prove this classical result, we cut the integral in two parts, the main one (near the origin) and a remainder. Before we proceed, let us remark that, up to a change of coordinates, we may as well suppose that $\mathbf{Hess}(V)_0$ is a diagonal matrix, and we have Taylor's formula:

$$V(x) = \frac{1}{2} \sum_i \lambda_i x_i^2 + \epsilon(x) \sum_i x_i^2,$$

where $\epsilon(x)$ goes to zero at 0. We choose an r such that, on $B = [-r, r]^d$, $\epsilon(x) \leq \frac{1}{4}(\inf \lambda_i) \sum x_i^2$.

Let us begin by the negligible part, outside of B. Since V goes to infinity, and 0 is the unique global minimum, there exists an $\eta > 0$ such that $V(x) \geq \eta$ outside B. We introduce an $\exp(-V)$ in the integral (the growth hypothesis makes it integrable), and use this bound:

$$\begin{aligned}
\int_{x \notin B} \exp(-V/\sigma)dx &= \int_{x \notin B} \exp(-V) \exp\left(-(1/\sigma - 1)V(x)\right) dx \\
&\leq \int_{x \notin B} \exp(-V)dx \exp\left(-(1/\sigma - 1)\eta\right) \\
&\leq Z_1 \exp\left(-(1/\sigma - 1)\eta\right).
\end{aligned}$$

Let us turn to the main term. We divide it by $\sigma^{d/2}$ (so that we only have to find a limit). We change variables and use $x = \phi_\sigma(y)$ defined by $x_i = y_i \sqrt{\sigma/\lambda_i}$.

$$\begin{aligned}
\sigma^{-d/2} \int_B \exp(-V/\sigma)dx_1 \cdots dx_n &= \sigma^{-d/2} \int \mathbf{1}_{x \in B} \exp\left(-\frac{1}{2} \sum_i \frac{\lambda_i}{\sigma} x_i^2 + \epsilon(x) \sum_i x_i^2\right) dx \\
&= \frac{1}{\sqrt{\lambda_1 \cdots \lambda_n}} \int \mathbf{1}_{\phi_\sigma(y) \in B} \exp\left(-\frac{1}{2} \sum y_i^2 + \epsilon(\phi_\sigma(y))\right) dy.
\end{aligned}$$

The function inside the integral converges pointwise to $\exp(-\sum y_i^2)$ when σ goes to zero (because $\phi_\sigma(y)$ goes to zero for a fixed y). It is bounded from above by the integrable function $\exp(-\frac{1}{4}\sum y_i^2$, and we may apply Lebesgue's dominated convergence:

$$\sigma^{-d/2}\int_B \exp(-V/\sigma)dx_1\cdots dx_n \xrightarrow[\sigma\to 0]{} \frac{(2\pi)^{d/2}}{\sqrt{\lambda_1\cdots\lambda_n}}$$

With the bound on the remainder, this gives the equivalent of Z_σ.

C.2 Finiteness of the entropy and regularity

We begin by proving that the relative entropy I_t is finite. To do this, we study directly the explicit density, which we know thanks to a Girsanov transform. We follow a proof from [Roy99], with a few minor changes to deal with the non-homogeneity in time.

Recall that the process X is defined by the following SDE:

$$dX_t = \sqrt{\sigma(t)}dB_t - \frac{1}{2}\nabla V(X_t)dt.$$

If we define a new reference martingale $M_t = \int_0^t \sqrt{\sigma(s)}dB_s$, we may define X as the solution to the SDE:

$$dX_t = dM_t - \frac{1}{2}\nabla V(X_t)dt.$$

Note that M_t is just a Brownian motion under a (deterministic) change of time — if we define $\tau(t) = \int_0^t \sigma(s)ds$, $M_{\tau^{-1}}(t)$ is a Brownian motion. To find the density of the law of X_t with respect to its equilibrium measure μ_t, we decompose it in three terms:

$$\frac{d\mathcal{L}(X_t)}{d\mu_t} = \frac{d\mathcal{L}(X_t)}{d\mathcal{L}(M_t)} \times \frac{d\mathcal{L}M_t}{d\lambda} \times \frac{d\lambda}{d\mu_t}.$$

To compute the first term, we use the (trajectorial) density of X with respect to M, which is given by Girsanov's theorem:

$$\begin{aligned} F &= \exp\left(-\frac{1}{2}\int \nabla V(M_s)dM_s - \frac{1}{2}\int_0^t \frac{|\nabla V|^2}{4}(M_s)d\langle M\rangle_s\right) \\ &= \exp\left(-\frac{1}{2}\int \nabla V(M_s)dM_s - \frac{1}{8}\int_0^t |\nabla V|^2(M_s)\sigma(s)ds\right). \end{aligned}$$

To get rid of the martingale term in the exponential, we apply Itô's formula to V and the martingale M:

$$V(M_t) = V(x) + \int_0^t \nabla V(M_s)dM_s + \frac{1}{2}\int_0^t \Delta V(M_s)d\langle M\rangle_s.$$

The functional F may thus be rewritten:

$$F = \exp\left(\frac{1}{2}V(x) - \frac{1}{2}V(M_t) + \int_0^t \left(\frac{1}{4}\Delta V(M_s) - \frac{1}{8}\Delta V(M_s)\right)\sigma(s)ds\right).$$

The three densities we are looking for are:

$$\begin{aligned} \frac{d\mathcal{L}(X_t)}{d\mathcal{L}(M_t)}(M_t) &= f(M_t) = \mathbf{E}[F|\mathcal{F}_{\{t\}}] \\ \frac{d\mathcal{L}M_t}{d\lambda}(y) &= \exp(-2v_t(y)) = (2\pi\tau(t))^{-d/2}\exp\left(-\frac{(x-y)^2}{2\tau(t)}\right) \\ \frac{d\lambda}{d\mu_t}(y) &= Z_{\sigma(t)}\exp\left(-\frac{V(y)}{\sigma(t)}\right). \end{aligned}$$

We take the product of these terms; the last two quantities may be put into the conditional expectation, so that the density we are looking for (say G) may be written as:

$$G(M_t) = Z_{\sigma(t)}\mathbf{E}\left[F\exp\left(\frac{V(M_t)}{\sigma(t)} - 2v_t(M_t)\right)|\mathcal{F}_{\{t\}}\right]$$

Let us now define $\gamma : x \mapsto x\log(x)$, and start to study I_t. By definition, $I_t = \int \gamma(G(y))d\mu_t(y)$. Since G is best expressed as a conditional expectation, we rewrite I_t:

$$\begin{aligned} I_t &= \mathbf{E}\left[\gamma(G(M_t))\frac{d\mu_t}{d\mathcal{L}(M_t)}\right] \\ &= \mathbf{E}\left[\gamma(G(M_t))\frac{1}{Z_{\sigma(t)}}\exp\left(-\frac{V(M_t)}{\sigma(t)} + 2v_t(M_t)\right)\right]. \end{aligned} \tag{22}$$

Since γ is convex, we may apply Jensen's conditional inequality to $\gamma(G(M_t))$, and develop γ:

$$\begin{aligned} \gamma(G(M_t)) &\leq \mathbf{E}\left[\gamma\left(Z_{\sigma(t)}F\exp\left(\frac{V}{\sigma} - 2v_t\right)\right)|\mathcal{F}_{\{t\}}\right] \\ &\leq \mathbf{E}\left[Z_\sigma F\exp\left(\frac{V}{\sigma} - 2v_t\right)\left(\log Z_\sigma + \log F + \frac{V}{\sigma} - 2v_t\right)|\mathcal{F}_{\{t\}}\right]. \end{aligned}$$

Multiply both sides by $\exp(-V/\sigma + 2v_t)$, and take the expected value; the left hand side becomes I_t (thanks to (22)), the conditioning disappears and we get:

$$I_t \leq \mathbf{E}\left[F\left(\log Z_\sigma + \log F + \frac{V(M_t)}{\sigma(t)} - 2v_t(M_t)\right)\right]$$

Recall that F is a density, so that $\mathbf{E}[F] = 1$, and we may take the constant Z_σ out of the expectation. We add and substract $(2/\sigma)\log(F)$ inside the integral — this will help us get rid of the term $V(M_t)/\sigma$:

$$I_t \leq \log(Z_\sigma) - (2/\sigma - 1)\mathbf{E}\left[F\log F\right] + \mathbf{E}\left[F\left(\frac{2}{\sigma}\log F + \frac{V(M_t)}{\sigma} - 2v_t(M_t)\right)\right].$$

Since $x\log x$ is bounded below, and $2/\sigma - 1$ is positive, the second term is bounded from above (for any finite time t). The same is true for the first term. The only thing to check is that the last term is finite; let us call this term A. Since F is given by an exponential, A is given by:

$$A = \mathbf{E}\left[F\left(\frac{1}{\sigma}V(x) + \frac{1}{4\sigma(t)}\int_0^t \left(2\Delta V - |\nabla V|^2\right)(M_s)\sigma(s)ds - 2v_t(M_t)\right)\right].$$

Let us consider the quantity between brackets. The first term is finite and does not depend on M_t. The integral is bounded above by something also independant of M_t (indeed, $2\Delta V - |\nabla V|^2$ is uniformly bounded from above, because ΔV is negative outside a compact set). The only thing left to check is that:

$$\mathbf{E}\left[F(-2v_t(M_t))\right] < \infty.$$

We have already seen the explicit value of v_t:

$$\exp\left(-2v_t(y)\right) = (2\pi\tau(t))^{-d/2}\exp\left(-\frac{(y-x)^2}{2\tau(t)}\right).$$

Taking logarithms, we see that:

$$-2v_t(y) = -\frac{d}{2}\log(2\pi\tau(t)) - \frac{(y-x)^2}{2\tau(t)}.$$

Since the last term is positive, this quantity is bounded from above by something which does not depend on y. Therefore, $\mathbf{E}[-F \times (2v_t(M_t))]$ is finite. This concludes the proof.

Contents

1 The convergence of the process (main line of the proof) 3

1.1 A differential inequality for the free energy . . . 3

1.2 Moment bounds and pseudo-entropy . . . 5

1.3 From the differential inequality to the convergence of the entropy . . . 6

1.4 Some remarks . . . 8

2 The one-dimensional case **8**

3 The weak inequality in any dimension **12**
3.1 The last hypothesis on the potential . . . 12
3.2 The one-point weak inequality . . . 13
3.3 Building a path and straightening it out . . . 14
3.4 The measure-capacity inequality . . . 18
3.5 Conclusion . . . 19

4 A measure-capacity criterion for one-point weak Poincaré inequalities **20**
4.1 Definitions . . . 20
4.2 Measure-capacity inequalities for large sets and one-point inequalities . . . 21
4.3 Weak inequalities for different norms . . . 23

A Orlicz norms, entropy and centering **24**

B A moment bound **26**
B.1 Outline of the proof . . . 26
B.2 An estimate on the generator . . . 27

C Regularity results and estimates on the process **30**
C.1 An equivalent of the partition function . . . 30
C.2 Finiteness of the entropy and regularity . . . 31

References

[Ale04] J. Alexopoulos – "A brief introduction to n-functions and Orlicz function spaces", `http://www.personal.kent.edu/~jalexopo/Lectures.pdf`, 2004.

[BCR] F. Barthe, P. Cattiaux et C. Roberto – "Interpolated inequalities between exponential and gaussian, orlicz hypercontractivity and application to isoperimetry", Revista Mat. Iberoamericana, To appear.

[BCR05] F. Barthe, P. Cattiaux et C. Roberto – "Concentration for independent random variables with heavy tails", *AMRX Appl. Math. Res. Express* (2005), no. 2, p. 39–60.

[Beb03] M. Bebendorf – "A note on the Poincaré inequality for convex domains", *Zeitschrift für Analysis und Ihre Anwendungen* **22** (2003), p. 751–756.

[BG99] S. G. Bobkov et F. Götze – "Exponential integrability and transportation cost related to logarithmic Sobolev inequalities", *Journal of Functional Analysis* (1999), no. 163, p. 1–28.

[Che90] R. Chen – "Neumann eigenvalue estimate on a compact Riemannian manifold", *Proc. Amer. Math. Soc.* **108** (1990), no. 4, p. 961–970.

[CL97] R. Chen et P. Li – "On Poincaré type inequalities", *Trans. Amer. Math. Soc.* **349** (1997), no. 4, p. 1561–1585.

[Die68] J. Dieudonné – *Calcul infinitésimal*, Hermann, Paris, 1968.

[DMM99] P. Del Moral et L. Miclo – "On the convergence and applications of generalized simulated annealing", *SIAM J. Control Optim.* **37** (1999), no. 4, p. 1222–1250 (electronic).

[Eva98] L. C. Evans – *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998.

[HCS87] C.-R. Hwang, T.-S. Chiang et S.-J. Sheu – "Diffusion for global optimization in $\mathbb{R}^n$", *Siam Journal on Control and Optimization* **25** (1987), p. 737–753.

[Hwa80] C.-R. Hwang – "Laplace's method revisited: weak convergence of probability measures", *Annals of Probability* **8** (1980), p. 1177–1182.

[Jac94] S. Jacquot – "Simulated annealing on Wiener space", *Stochastics Stochastics Rep.* **51** (1994), no. 3-4, p. 159–194.

[Jac96] — , "Simulated annealing for stochastic semilinear equations on Hilbert spaces", *Stochastic Process. Appl.* **64** (1996), no. 1, p. 73–91.

[JR95] S. Jacquot et G. Royer – "Ergodicity of stochastic plates", *Probab. Theory Related Fields* **102** (1995), no. 1, p. 19–44.

[Loc00] M. Locatelli – "Simulated annealing algorithms for continuous global optimization: Convergence conditions", *Journal of Optimization Theory and Applications* (2000), no. 104, p. 121–133.

[Mic91] L. Miclo – "Thèse de doctorat", Thèse, Université Paris VI, 1991.

[Mic92] — , "Recuit simulé sur $\mathbb{R}^n$. Étude de l'évolution de l'énergie libre", *Annales de l'Institut Henri Poincaré* **28** (1992), no. 2, p. 235–266.

[PW60] L. E. Payne et H. F. Weinberger – "An optimal Poincaré inequality for convex domains", *Archive for Rational Mechanics and Analysis* **5** (1960), p. 286–292.

[Roy89] G. Royer – "A remark on simulated annealing of diffusion processes", *SIAM Journal on Control and Optimization* **27** (1989), no. 6, p. 1403–1408.

[Roy99] — , *Une initiation aux inégalités de Sobolev logarithmiques*, Cours spécialisés, no. 5, SMF, 1999.

[RR91] M. M. Rao et Z. D. Ren – *Theory of Orlicz spaces*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 146, Marcel Dekker Inc., New York, 1991.

[RW01] M. Röckner et F.-Y. Wang – "Weak Poincaré inequalities and L^2 convergence rates of Markov semigroups", *Journal of Functional Analysis* **185** (2001), p. 564–603.