Monotonicity of the Extremal Functions for One-dimensional Inequalities of Logarithmic Sobolev Type

Laurent Miclo

Laboratoire d'Analyse, Topologie, Probabilités, UMR 6632, CNRS
39, rue F. Joliot-Curie, 13453 Marseille cedex 13, France
e-mail: miclo@latp.univ-mrs.fr

Summary. In various one-dimensional functional inequalities, the optimal constants can be found by considering only monotone functions. We study the discrete and continuous settings (and their relationships); we are interested in Poincaré or logarithmic Sobolev inequalities, and several variants obtained by modifying entropy and energy terms.

Keywords: Poincaré inequality, (modified) logarithmic Sobolev inequality, monotonicity of extremal functions, linear diffusions, birth and death process

MSC2000: 46E35, 46E39, 49R50, 26A48, 26D10, 60E15

1 Introduction and Result

On the Borel σ-field of $\mathbb{R}$, let μ be a probability and ν a positive measure. We are interested in the logarithmic Sobolev constant $C(\mu, \nu)$ defined (with the usual conventions $1/\infty = 0$, $1/0 = \infty$ and, most important, $0 \cdot \infty = 0$) by

$$C(\mu, \nu) \coloneqq \sup_{f \in \mathcal{C}} \frac{\operatorname{Ent}(f^2, \mu)}{\nu[(f')^2]} \quad \in \bar{\mathbb{R}}_+ \tag{1}$$

where $\mathcal{C}$ is the set of all absolutely continuous functions f on $\mathbb{R}$; f' denotes the weak derivative of f. Recall that in general the entropy of a positive, measurable function f with respect to a probability μ is defined as

$$\operatorname{Ent}(f, \mu) \coloneqq \begin{cases} \mu[f \ln(f)] - \mu[f] \ln(\mu[f]) & \text{if } f \ln(f) \text{ is } \mu\text{-integrable} \\ +\infty & \text{else} \end{cases}$$

and that this quantity belongs to $\bar{\mathbb{R}}_+$, as an immediate consequence of Jensen's inequality with the convex map $\mathbb{R}_+ \ni x \mapsto x \ln(x) \in \mathbb{R}$.

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLII*,
Lecture Notes in Mathematics 1979, DOI 10.1007/978-3-642-01763-6_2,

One of our aims is to show that the above definition of $C(\mu, \nu)$ is not modified when restricted to monotone functions:

Theorem 1. *Calling $\mathcal{D}$ the cone in $\mathcal{C}$ consisting of all functions f such that $f' \geqslant 0$ a.e., one has*

$$C(\mu, \nu) \coloneqq \sup_{f \in \mathcal{D}} \frac{\mathrm{Ent}(f^2, \mu)}{\nu[(f')^2]} \quad \in \bar{\mathbb{R}}_+.$$

This can be illustrated by the most famous case of the logarithmic Sobolev inequality (due to Gross [10]), where $\mu = \nu$ is a Gaussian (non degenerate) distribution; then the maximising functions are exactly the exponentials $\mathbb{R} \ni x \mapsto \exp(ax + b)$ with $a \in \mathbb{R}^*$ and $b \in \mathbb{R}$ (see Carlen's article [4]).

We shall also be interested in the following discrete version of the preceding result. For a given $N \in \mathbb{N}^*$, consider the discrete segment $E \coloneqq \{0, 1, ..., N\}$ as a linear non-oriented graph; call $A \coloneqq \big\{\{l, l+1\} \,:\, 0 \leqslant l < N\big\}$ the set of its edges. Denote by $\mathcal{C}$ the set of functions defined on E. If $f \in \mathcal{C}$, its discrete derivative f' is defined on A by

$$\forall\, 0 \leqslant l < N, \qquad f'\big(\{l, l+1\}\big) \coloneqq f(l+1) - f(l)$$

Let also be given a probability μ on E and a measure ν on A. These notations enable us to reinterpret (1) in this new setting, and, as above, our main concern will be to prove:

Theorem 2. *In this discrete framework, one has*

$$C(\mu, \nu) = \sup_{f \in \mathcal{D}} \frac{\mathrm{Ent}(f^2, \mu)}{\nu[(f')^2]} \quad \in \bar{\mathbb{R}}_+$$

where $\mathcal{D}$ is the cone in $\mathcal{C}$ consisting of those functions with positive derivative.

In fact, using interlinks between the continuous and discrete contexts, one can pass from one result to the other. So we shall start with the discrete situation, which is more immediate and better illustrates our itinerary; then similar properties in the continuous framework will derive from the discrete one. The discrete proof can also be directly translated, but precautions must be taken; more on this later.

These monotonicity properties will also be extended to some modified logarithmic Sobolev inequalities (discrete, as in Wu [18] or coutinuous in the sense of Gentil, Guillin and Miclo [9]).

More precisely, in the discrete case, one would like to replace the energy term $\nu[(f')^2]$ by the quantity $\mathcal{E}_\nu(f^2, \ln(f^2))$ defined for $f \in \mathcal{C}$ by

$$\sum_{\{l,l+1\} \in A} \nu(\{l, l+1\})[f^2(l+1) - f^2(l)][\ln(f^2(l+1)) - \ln(f^2(l))];$$

observe that this quantity is quadratically homogeneous. This will be done in

Theorem 3. *Consider the case that $E = \mathbb{Z}$, with the previous notations extended to this setting. One has*

$$\sup_{f\in\mathcal{C}} \frac{\mathrm{Ent}(f^2,\mu)}{\mathcal{E}_\nu(f^2,\ln(f^2))} = \sup_{f\in\mathcal{D}} \frac{\mathrm{Ent}(f^2,\mu)}{\mathcal{E}_\nu(f^2,\ln(f^2))}$$

In the continuous framework, let $H : \mathbb{R}_+ \to \mathbb{R}_+$ be a convex function such that $H(0) = 0$ and $H'(0) = 1$. We now wish to replace the energy term with the following quadratically homogeneous quantity:

$$\forall\, f \in \mathcal{C}, \qquad \mathcal{E}_{H,\nu}(f) := \int H\left(\left(\frac{f'}{f}\right)^2\right) f^2\, d\nu$$

where by convention the integrand equals $(f')^2$ on the set where f vanishes. As before, one then has

Theorem 4. *If μ is a probability on $\mathbb{R}$ and ν a measure on $\mathbb{R}$, one has*

$$\sup_{f\in\mathcal{C}} \frac{\mathrm{Ent}(f^2,\mu)}{\mathcal{E}_{H,\nu}(f)} = \sup_{f\in\mathcal{D}} \frac{\mathrm{Ent}(f^2,\mu)}{\mathcal{E}_{H,\nu}(f)}.$$

Similar results will be obtained when it is the entropy which is modified; for a precise statement, see sub-section 5.3.

But our main motivation comes from the modified logarithmic Sobolev inequalities in Theorems 3 and 4, because we hope that the monotonicity properties we have established eventually allow to apply Hardy inequalities. Indeed, the link between Hardy and modified logarithmic Sobolev inequalities is still poorly understood, whereas that between Hardy and Poincaré, or classical logarithmic Sobolev, inequalities is clear (see for instance Bobkov and Götze's article [3]).

Besides, let us mention that similar results for the Poincaré constant have already been obtained, in the discrete case by Chen (in the proof of Theorem 3.2 in [7]) and in the continuous case by Chen and Wang (Proposition 6.4 in [6], see also the end of the proof of Theorem 1.1 in Chen [8]), for diffusions which are regular enough. Their method partially rests on the equation satisfied by a maximising function (which then is an eigenvector associated to the spectral gap). But it does not clearly adapt to logarithmic Sobolev inequalities, nor even, in the case of the Poincaré constant, to the irregular situations considered above (see for instance the continuity hypothesis needed in the second part of Theorem 1.3 of Chen [8]); therefore we prefer another approach. In particular, we do not a priori deal with the problem of existence of a maximising function (which is crucial in the approach by Chen and Wang [6, 8]). Furthermore, it may be preferable to attack this existence question a posteriori, when discussion is restricted to increasing functions; for rather regular situations, see also the last remark in Section 4.

Still in the case of the Poincaré constant, observe that the equation giving the maximising functions (if they exist) is not easily exploited, for it already involves the Poincaré constant which is unknown in general. Moreover, if in this equation the constant is replaced by the inverse of an eigenvalue other than 0 and the spectral gap, the functions which satisfy this new equation are the corresponding eigenvectors, which are not monotone (under irreducibility hypotheses; see for instance [12]). Therefore we prefer to base our approach on Dirichlet forms rather than on the equation possibly satisfied by the maximising functions.

Let us add that, at least in the case of the Poincaré constant, some monotonicity properties can also be obtained when the underlying graph is a tree. See [12] for a description of the eigenspace associated to the spectral gap (in the discrete case).

The outline of the article is as follows: the next section deals with monotonicity properties for the spectral gap; they have to be considered first to treat the case when no extremal function exists in the above logarithmic Sobolev inequalities. The situations when it exists will then be studied in Section 3, still in the discrete setting. Then Section 4 will extend discussion to the continuous setting, by two different ways. The last section will be devoted to extensions with modified entropy or energy.

Last, I wish to thank the referee whose sugestions led to a better presentation.

2 Poincaré Inequality

In the discrete setting presented in the introduction, we consider the inverse of the spectral gap (also called Poincaré constant) associated to μ and ν, defined by

$$A(\mu,\nu) \coloneqq \sup_{f\in\mathcal{C}} \frac{\mathrm{Var}(f,\mu)}{\nu[(f')^2]} \quad \in \bar{\mathbb{R}}_+, \tag{2}$$

where we recall that the variance of a measurable function f with respect to a probability μ is defined by

$$\mathrm{Var}(f,\mu) = \int \big(f(y)-f(x)\big)^2 \,\mu(dx)\,\mu(dy) \quad \in \bar{\mathbb{R}}_+.$$

The interest for us of $A(\mu,\nu)$ comes from Theorem 2.2.3 in Saloff-Coste's course [17], where a result due to Rothaus [13, 14, 15] is adapted to the continuous case (in a more general framework than our one-dimensional one). It says that either $C(\mu,\nu) = 2A(\mu,\nu)$, or there exists a function $f \in \mathcal{C}$ such that $C(\mu,\nu) = \mathrm{Ent}(f^2,\mu)/\nu[(f')^2]$. This alternative is shown by considering a maximising sequence in (1). So, keeping in mind the aim presented in the introduction, it is useful and instructive to start with its analogue for the spectral gap:

Proposition 1. *Definition (2) is not changed when $\mathcal{C}$ is replaced with $\mathcal{D}$, that is, when only monotone functions are considered.*

But first observe that the supremum featuring in (2) is always achieved. To establish that, two situations will be distinguished.

a) The non-degenerate case where $\nu(a) > 0$ for every $a \in A$. As the expressions $\mathrm{Var}(f, \mu)$ and $\nu[(f')^2]$ are invariant when a constant is added to f and as they are quadratically homogeneous, in (2) one may consider only functions f such that $f(0) = 0$ and $\nu[(f')^2] = 1$. Let now $(f_n)_{n\in\mathbb{N}}$ be a maximising sequence for (2) which satisfies those two conditions. The hypothesis on ν clearly ensures boundedness in $\mathbb{R}^{1+N}$ of the sequence $(f_n)_{n\in\mathbb{N}}$. Hence a convergent subsequence can be extracted, with limit a function f. This limit also verfies $\nu[(f')^2] = 1$, wherefrom one easily deduces that $A(\mu, \nu) = \mathrm{Var}(f, \mu)/\nu[(f')^2]$, showing existence of an extremal function for (2).

b) If $\nu(\{i, i+1\}) = 0$ for some $\{i, i+1\} \in A$, two sub-cases can be considered:

b1) If $\mu(\{0, ..., i\}) > 0$ and $\mu(\{i+1, ..., N\}) > 0$, putting $f = \mathbb{1}_{\{i+1,...,N\}}$, one has $\mathrm{Var}(f, \mu) > 0$ and $\nu[(f')^2] = 0$, hence $C(\mu, \nu) = +\infty$ and f is extremal.

b2) Else, one among $\mu(\{0, ..., i\})$ and $\mu(\{i+1, ..., N\})$ vanishes, and the problem can be restricted to the segment $\{0, ..., i\}$ or $\{i+1, ..., N\}$, whichever has mass 1. By iteration, one is then back to one of the preceding cases.

Note that in the above case (b1), Proposition 1 is established; so we can henceforth assume that $\nu > 0$ on A. This observation is also valid for the logarithmic Sobolev constant, and it almost makes it possible to assume the irreducibility hypothesis of Theorem 2.2.3 of Saloff-Coste [17], except that μ was a priori not supposed to be strictly positive on E. Yet, one always can revert to this situation: call $0 \leqslant x_0 < x_2 < \cdots < x_n \leqslant N$ the elements of E with strictly positive μ-weight. Given some real numbers $y_0, y_1, ..., y_n$, consider the affine sub-space of $\mathcal{C}$ consisting of those functions f such that $f(x_i) = y_i$ for each $0 \leqslant i \leqslant n$, and try minimising $\nu[(f')^2]$ therein. For fixed $0 \leqslant i < n$, this leads to look for the functions g on $\{x_i, x_i+1, ..., x_{i+1}\}$ which minimise $\sum_{x_i \leqslant x < x_{i+1}} \nu(\{x, x+1\})(g'(\{x, x+1\}))^2$ under the constraints $g(x_i) = y_i$ and $g(x_{i+1}) = y_{i+1}$. By a simple application of the equality case in the Cauchy-Schwarz inequality, this optimisation problem admits the following unique solution:

$$\forall\ x_i \leqslant x \leqslant x_{i+1},$$
$$g(x) = y_i + \left(\sum_{x_i \leqslant y < x_{i+1}} \frac{1}{\nu\big(\{y, y+1\}\big)} \right)^{-1} \sum_{x_i \leqslant y < x} \frac{y_{i+1} - y_i}{\nu\big(\{y, y+1\}\big)}. \tag{3}$$

So, setting

$$\forall\ 0 \leqslant i \leqslant n, \qquad \widetilde{\mu}(i) \coloneqq \mu(x_i)$$
$$\forall\ 0 \leqslant i < n, \qquad \widetilde{\nu}(\{i, i+1\}) \coloneqq \left(\sum_{x_i \leqslant y < x_{i+1}} \frac{1}{\nu(\{y, y+1\})} \right)^{-1},$$

one would be reduced to a situation where the underlying probability is everywhere strictly positive; moreover, using (3), one easily switches back and forth between extremal functions for both problems. This would also fully justify the reminder before Proposition 1.

On the other hand, we shall also discard the trivial case when μ is a Dirac mass; this ensures that $A(\mu, \nu) > 0$.

We can now be a little more precise on the maximising functions in (2):

Lemma 1. *Let f be a function realizing the maximum in (2). Assuming that $\nu > 0$ on A and that μ is not a Dirac mass, every maximising function has the form $af + b\mathbb{1}$ where $a \in \mathbb{R}^*$, $b \in \mathbb{R}$ and $\mathbb{1}$ denotes the constant function with value 1.*

Proof. Clearly, if f is maximising and if $a \in \mathbb{R}^*$ and $b \in \mathbb{R}$, $af + b\mathbb{1}$ is also maximising in (2).

Conversely, let g be maximising in (2); by subtracting a constant, we may suppose that $\mu[g] = 0$. By variational calculus around g (i.e., by considering $g + \epsilon h$, with $\epsilon \in \mathbb{R}$ and any $h \in \mathcal{C}$, and taking a first order expansion when $\epsilon \to 0$ of the ratio $\mathrm{Var}(g + \epsilon h, \mu)/\nu[(g' + \epsilon h')^2]$), one easily sees that for each $i \in E$, g satisfies

$$A(\mu,\nu)\big[\nu\big(\{i,i{+}1\}\big)\big(g(i) - g(i{+}1)\big) + \nu\big(\{i{-}1,i\}\big)\big(g(i) - g(i{-}1)\big)\big] = \mu(i)g(i)$$

with the conventions $\nu\big(\{-1,0\}\big) = 0 = \nu\big(\{N, N+1\}\big)$.

Now, since $A(\mu,\nu) > 0$ and $\nu > 0$ on A, starting from $g(0)$ these equations inductively determine $g(1)$, $g(2)$, up to $g(N)$. Note that $g(0) \neq 0$, else we would end up with $g \equiv 0$, contradicting $A(\mu, \nu) > 0$. So there is at most one minimising function g for (2) which satisfies $\mu[g] = 0$ and $g(0) = 1$. This is exactly what the lemma asserts. □

Given a maximising f for (2), our strategy to show its monotonicity will be as follows: supposing on the contrary f not to be monotone, we shall decompose f as $\widetilde{f} + \widehat{f}$, with $\widetilde{f}$ (and hence also $\widehat{f}$) not belonging to the linear span $\mathrm{Vect}(\mathbb{1}, f)$, and with

$$\begin{aligned}\mathrm{Var}(f,\mu) &= \mathrm{Var}(\widetilde{f},\mu) + \mathrm{Var}(\widehat{f},\mu)\\ \nu[(f')^2] &\geqslant \nu[(\widetilde{f}')^2] + \nu[(\widehat{f}')^2].\end{aligned}$$

Clearly, these two relations imply that $\widetilde{f}$ and $\widehat{f}$ also are maximising for (2), a contradiction since $\widetilde{f}$ and $\widehat{f}$ do not have the form required by Lemma 1.

So let f be maximising for (2) but not monotone.

A point $i \in E$ will be called a local maximum of f if for each $j \in E$ verifying $f(j) > f(i)$, the segment $[\![i, j]\!]$ (the sub-segment of E with endpoints i and j) contains an element k such that $f(k) < f(i)$. By definition, a local minimum of f will be a local maximum of $-f$.

We shall now construct $\widetilde{f}$ by splitting f at a particular level. Replacing f by $-f$ if necessary, we may choose a local maximum i in $[\![1, N-1]\!]$ such that f has a local minimum in $[\![0,i]\!]$ and another one in $[\![i,N]\!]$. Among such local maxima i, choose one which minimises $f(i)$, and call it i_0. Denote by i_1 (respectively i_{-1}) the closest local minimum on the right (respectively on the left) of i_0. By possibly reversing the order of $[\![0,N]\!]$, one can suppose that $f(i_{-1}) \leqslant f(i_1)$. Also, set $i_2 \coloneqq \max\{y \geqslant i_1 \,:\, \forall\, i_1 \leqslant x \leqslant y,\ f(x) = f(i_1)\}$.

For $s \in [f(i_1), f(i_0)]$, let $S_s \coloneqq [\![a_s, b_s]\!]$ be the discrete segment whose ends are defined by

$$\begin{aligned} a_s &\coloneqq \min\{x \in [\![i_{-1}, i_0]\!] \,:\, f(x) \geqslant s\} \\ b_s &\coloneqq \min\{x \in [\![i_2, N]\!] \,:\, f(x) \geqslant s\} - 1 \end{aligned}$$

(with the convention that $b_s = N$ if the latter set is empty).

By those choices, particularly by minimality of i_0, one easily verifies that for any $s \in [f(i_1), f(i_0)]$, f is increasing (this is always understood in the wide sense) on $[\![a_s, i_0]\!]$, decreasing on $[\![i_0, i_2]\!]$ and increasing on $[\![i_2, b_s+1]\!]$ (the reader is urged to draw a picture).

Still for $s \in [f(i_1), f(i_0)]$, set for $x \in E$

$$\begin{aligned} \widetilde{f_s}(x) &= f(x)\mathbb{1}_{S_s^c}(x) + s\mathbb{1}_{S_s}(x) \\ \widehat{f_s}(x) &= (f(x) - s)\mathbb{1}_{S_s}(x). \end{aligned}$$

One has indeed $f_s = \widetilde{f_s} + \widehat{f_s}$, and the claimed decomposition will be obtained owing to the following two lemmas.

Lemma 2. *For any $s \in]f(i_1), f(i_0)[$, one has*

$$\nu[(f')^2] \geqslant \nu[(\widetilde{f'_s})^2] + \nu[(\widehat{f'_s})^2].$$

Proof. An immediate calculation first gives

$$\nu[(f')^2] = \nu[(\widetilde{f'_s} + \widehat{f'_s})^2] = \nu[(\widetilde{f'_s})^2] + \nu[(\widehat{f'_s})^2] + 2\nu[\widetilde{f'_s}\widehat{f'_s}]$$

and then

$$\begin{aligned} \nu[\widetilde{f'_s}\widehat{f'_s}] = \nu\big(\{a_s - 1, a_s\}\big)\big(s - f(a_s - 1)\big)\big(f(a_s) - s\big) \\ + \nu\big(\{b_s, b_s + 1\}\big)\big(f(b_s + 1) - s\big)\big(s - f(b_s)\big) \end{aligned}$$

(still with the convention that $\nu\big(\{N, N+1\}\big) = 0$). Now, from the fact that $s \in]f(i_1), f(i_0)[$, it appears that $f(i_{-1}) \leqslant f(a_s - 1) < s \leqslant f(a_s) \leqslant f(i_0)$ and $f(i_2) \leqslant f(b_s) < s \leqslant f(b_s + 1)$, which allows to notice that $\nu[\widetilde{f'_s}\widehat{f'_s}] \geqslant 0$, wherefrom the claimed inequality derives. □

Lemma 3. *There exists $s_0 \in]f(i_1), f(i_0)[$ such that*

$$\mathrm{Var}(f, \mu) = \mathrm{Var}(\widetilde{f_s}, \mu) + \mathrm{Var}(\widehat{f_s}, \mu).$$

Proof. The difference between the left and right hand sides is but twice the covariance of $\widetilde{f}_s$ and $\widehat{f}_s$ under μ, which equals

$$\begin{aligned}\mu\big[\big(\widetilde{f}_s - \mu[\widetilde{f}_s]\big)\big(\widehat{f}_s - \mu[\widehat{f}_s]\big)\big] &= \mu\big[\big(\widetilde{f}_s - \mu[\widetilde{f}_s]\big)\widehat{f}_s\big] \\ &= \big(s - \mu[\widetilde{f}_s]\big)\,\mu\big[(f-s)\mathbb{1}_{S_s}\big]. \end{aligned} \tag{4}$$

Hence, it suffices to find an $s \in \,]f(i_1), f(i_0)[$ such that $\mu[(f-s)\mathbb{1}_{S_s}] = 0$. Put $i_3 := b_{f(i_0)} + 1$; from the increasingness of f on $[\![i_{-1}, i_0]\!]$ and on $[\![i_2, i_3]\!]$, one is easily convinced that the map $\Psi : [f(i_1), f(i_0)] \ni s \mapsto \mu[(f-s)\mathbb{1}_{S_s}]$ is continuous. Now, the pattern of f on $[\![i_{-1}, i_3]\!]$ implies that $\Psi(f(i_1)) > 0$ and $\Psi(f(i_0)) < 0$, so there exists $s_0 \in \,]f(i_1), f(i_0)[$ such that $\Psi(s_0) = 0$. □

Notate $\widetilde{f} = \widetilde{f}_{s_0}$ and $\widehat{f} = \widehat{f}_{s_0}$, where s_0 is chosen as in the preceding lemma. To finalize the proof of Proposition 1, it remains to see that $\widetilde{f}$ is not in $\mathrm{Vect}(f, \mathbb{1})$. To this end, notice that i_1 is no longer a local minimum for $\widetilde{f}$ (this function may go down from i_1 to i_{-1}, and yet $\widetilde{f}(i_1) = s_0 > f(i_1) \geqslant f(i_{-1}) = \widetilde{f}(i_{-1})$), and consequently $\widetilde{f}$ cannot be written as $af + b\mathbb{1}$ with $a > 0$ and $b \in \mathbb{R}$. On the other hand, the inequalities $\widetilde{f}(i_{-1}) < \widetilde{f}(i_0)$ and $f(i_{-1}) < f(i_0)$ also show that $\widetilde{f}$ cannot be written as $af + b\mathbb{1}$ with $a \leqslant 0$ and $b \in \mathbb{R}$. Therefore the claimed result follows.

3 Splitting up the Entropy

Our aim here is to establish (2) in the discrete setting. According to the results from the preceding section, it suffices to consider the case when there exists a (non constant) maximising f for (1). For else, a maximising family for the logarithmic Sobolev inequality is $(1 + f/(n+1))_{n\in\mathbb{N}}$, where f is a maximising function for the corresponding Poincaré inequality (and hence f is monotone). Globally, the scheme of our proof will be similar to that of the previous section, most of whose notation will be kept in use.

First of all, observe that one may from now on suppose that $f \geqslant 0$, by possibly replacing f with $|f|$, since one has $\nu[(|f|')^2] \leqslant \nu[(f')^2]$. Assume now the hypothesis (to be refuted) that f is not monotone. Two possibilities arise: either f has a local maximum i in $[\![1, N-1]\!]$ such that there is a local minimum in $[\![0, i]\!]$ and one in $[\![i, N]\!]$, or the same holds for $-f$. We shall consider the first case only; the second one is very similar and left to the reader (one has to work with the negatively valued function $-f$).

As in section 2, i_{-1}, i_0, i_1, i_2 and i_3 are defined, then, for $s \in [f(i_1), f(i_0)]$, S_s, $\widetilde{f}_s$ and $\widehat{f}_s$. Our main task will consist in "splitting up" the entropy:

Lemma 4. *There exists* $s_1 \in \,]f(i_1), f(i_0)[$ *such that*

$$\mathrm{Ent}(f^2, \mu) = \mathrm{Ent}(\widetilde{f}^2_{s_1}, \mu) + \mathrm{Ent}\big((s + \widehat{f}_{s_1})^2, \mu\big).$$

Proof. First remark that for all $s \in [f(i_1), f(i_0)]$ and for all function $F : \mathbb{R}_+ \to \mathbb{R}$, one has

$$\mu\big[F(f)\big] = \mu\big[F(\widetilde{f}_s)\big] + \mu\big[F(s + \widehat{f}_s)\big] - F(s). \tag{5}$$

Indeed, by definition, one can perform the following expansion:

$$\begin{aligned}\mu\big[F(f)\big] &= \mu\big[\mathbb{1}_{S_s^c} F(\widetilde{f}_s)\big] + \mu\big[\mathbb{1}_{S_s} F(s + \widehat{f}_s)\big] \\ &= \mu\big[F(\widetilde{f}_s)\big] - \mu\big[\mathbb{1}_{S_s} F(s)\big] + \mu\big[F(s + \widehat{f}_s)\big] - \mu\big[\mathbb{1}_{S_s^c} F(s)\big] \\ &= \mu\big[F(\widetilde{f}_s)\big] + \mu\big[F(s + \widehat{f}_s)\big] - F(s).\end{aligned}$$

In particular, applying this to $F : \mathbb{R}_+ \ni u \mapsto u^2 \ln(u^2)$, it appears that

$$\mathrm{Ent}(f^2, \mu) - \mathrm{Ent}(\widetilde{f}^2_{s_1}, \mu) - \mathrm{Ent}\big((s + \widehat{f}_{s_1})^2, \mu\big) = \varphi(y'_s) + \varphi(x'_s) - \varphi(y) - \varphi(x_s)$$

with φ the convex map given by $\varphi : \mathbb{R}_+ \ni u \mapsto u \ln(u)$ and

$$\begin{aligned} y'_s &= \mu\big[\widetilde{f}^2_s\big] \qquad & y &= \mu[f^2] \\ x'_s &= \mu\big[(s + \widehat{f}_s)^2\big] \qquad & x_s &= s^2. \end{aligned}$$

Resorting again to (5), but with $F(s) = s^2$, it appears that $x_s + y = x'_s + y'_s$, which means that both segments $[x_s, y]$ and $[x'_s, y'_s]$ have the same midpoint. So, by convexity of φ, the inequality $\varphi(x_s) + \varphi(y) \geqslant \varphi(x'_s) + \varphi(y'_s)$ is equivalent to $|y - x_s| \geqslant |y'_s - x'_s|$. Or also, if some $s_1 \in]f(i_1), f(i_0)[$ happens to be such that $|y - x_s| = |y'_s - x'_s|$, then the equality in Lemma 3 holds (without even using the convexity of φ). Now one computes (still owing to (5) with $F(s) = s^2$) that

$$\begin{aligned} y'_s - x'_s &= \mu[\widetilde{f}^2_s] - \mu\big[(s + \widehat{f}_s)^2\big] = \mu[f^2] + s^2 - 2\mu\big[(s + \widehat{f}_s)^2\big] \\ &= \mu[f^2] - s^2 - 2\mu[\widehat{f}^2_s] - 4s\mu[\widehat{f}_s] = y - x_s - 2\mu\big[\widehat{f}_s(\widehat{f}_s + 2s)\big]. \end{aligned}$$

Hence it suffices to find an $s \in]f(i_1), f(i_0)[$ such that $\mu\big[\widehat{f}_s(\widehat{f}_s + 2s)\big] = 0$. But $\widehat{f}_s + 2s$ is a positive function, whereas $\widehat{f}_s$ is positive for $s = f(i_1)$ and negative for $s = f(i_0)$. The claim follows by continuity of the application $[f(i_1), f(i_0)] \ni s \mapsto \mu[\widehat{f}_s(\widehat{f}_s + 2s)]$, which is easily seen not to vanish at the endpoints. □

Besides, according to Lemma 2, one has for all $s \in]f(i_1), f(i_0)[$

$$\nu[(f')^2] \geqslant \nu[(\widetilde{f}'_s)^2] + \nu[(\widehat{f}'_s)^2] = \nu[(\widetilde{f}'_s)^2] + \nu[((s + \widehat{f}_s)')^2].$$

Using the notation and proof of that Lemma again, one can even say a little more: equality can hold only if for all edges $a \in A$ one has $\widetilde{f}'_s(a)\widehat{f}'_s(a) = 0$, which in particular entails that $f(a_s) = s$. So, for $s \in]f(i_1), f(i_0)[$, the discrete segment S_s contains at least three different points, a_s, i_0 and i_1.

Now, what we saw just before implies that $\widetilde{f}_{s_1}$ and $s_1 + \widehat{f}_{s_1}$ also are maximising functions for (1), and that necessarily

$$\nu\big[(f')^2\big] = \nu\big[(\widetilde{f}'_{s_1})^2\big] + \nu\big[\big((s_1 + \widehat{f}_{s_1})'\big)^2\big],$$

for else, one would have

$$\begin{aligned}\frac{\mathrm{Ent}(f^2,\mu)}{\nu\big[(f')^2\big]} &< \frac{\mathrm{Ent}(\widetilde{f}^2_{s_1},\mu) + \mathrm{Ent}\big((s+\widehat{f}_{s_1})^2,\mu\big)}{\nu\big[(\widetilde{f}'_{s_1})^2\big] + \nu\big[\big((s_1+\widehat{f}_{s_1})'\big)^2\big]}\\ &\leqslant \max\left(\frac{\mathrm{Ent}(\widetilde{f}^2_{s_1},\mu)}{\nu\big[(\widetilde{f}'_{s_1})^2\big]}, \frac{\mathrm{Ent}\big((s+\widehat{f}_{s_1})^2,\mu\big)}{\nu\big[\big((s_1+\widehat{f}_{s_1})'\big)^2\big]}\right)\end{aligned}$$

(the first inequality uses that by construction $\mathrm{Ent}(f^2,\mu) > 0$). Therefore there exist three successive points in S_{s_1} where $\widetilde{f}_{s_1}$ assumes the same value (namely, s_1) and we shall now verify that this is not possible, more precisely that this would imply constancy of $\widetilde{f}_{s_1}$, which does not hold (for $\widetilde{f}_{s_1}(i_{-1}) < \widetilde{f}_{s_1}(i_0)$). Indeed, by variational calculus around a maximising function f, one sees that f must verify for all $i \in E$ (with the usual conventions)

$$\begin{aligned}C(\mu,\nu)\,\big[\nu\big(\{i,i{+}1\}\big)\big(f(i)-f(i{+}1)\big) + \nu\big(\{i{-}1,i\}\big)\big(f(i)-f(i{-}1)\big)\big]\\ = \mu(i)f(i)\,\ln\left(\frac{f^2(i)}{\mu[f^2]}\right).\end{aligned}$$

Recall that discussion has been reduced to the situation that μ, ν and $C(\mu,\nu)$ are strictly positive (see before Lemma 1); so if f takes the same value v at three successive points $y-1$, y and $y+1$, with $0 < y < N$, then the preceding equation taken at $i = y$ forces $v\ln(v^2/\mu[f^2]) = 0$, that is to say, $v = 0$ or $v = \sqrt{\mu[f^2]}$. Applying then the equation at $i = y+1$ instead, one obtains $f(y+2) = f(y+1)$, at least if $y \leqslant N-2$. Similarly, for $i = y-1$, $f(y-2) = v$ if $y \geqslant 2$. So equality $f(i) = v$ propagates everywhere and f is constanty equal to v.

These arguments terminate the proof of (2) by replacing the recourse to Lemma 1. For even though the knowledge of $\mu[f^2]$ and of $f(0)$ determines a maximising function f for (1) owing to the linear structure of the graph E (still for fixed μ and ν verifying $C(\mu,\nu) > 0$ and $\nu > 0$ on A, as we were allowed to suppose in the preceding section), here this no longer implies Lemma 1 because the term $\mu(i)f(i)\ln(f^2(i)/\mu[f^2])$ above is not affine in $f(i)$. Besides, this lemma never holds in the context of logarithmic Sobolev inequalities. Indeed, let again f be a positive function which maximises (1). Perturbating f by a constant function and performing a variational computation, one obtains $\mu[f\ln(f/\mu[f^2])] = 0$. Set $F(t) = \mu[(f+t)\ln((f+t)/\mu[(f+t)^2])]$ for all $t \geqslant 0$. Differentiating twice this expression on $\mathbb{R}_+^*$, one obtains

$$F''(t) = 2\int \frac{1}{f+t}\,d\mu - 2\frac{\mu[f+t]}{\mu[(f+t)^2]}\left(2 - \frac{\mu[f+t]^2}{\mu[(f+t)^2]}\right).$$

Using Jensen's inequality $\mu[1/(f+t)] \geqslant 1/\mu[f+t]$ and the fact that the map $[0,1] \ni x \mapsto x(2-x)$ is bounded by 1, it appears that F'' is strictly positive on $\mathbb{R}_+^*$ if f is not μ-a.s. constant (consider the case when Jensen's inequality is an equality). So, there may exist at most two $t \geqslant 0$ such that $F(t) = 0$.

Remark 1. The inequality $\mu[\widehat{f}_{f(i_1)}(\widehat{f}_{f(i_1)} + 2f(i_1))] > 0$ does not allow to deduce that $\mathrm{Ent}(f^2, \mu) < \mathrm{Ent}(\widetilde{f}^2_{f(i_1)}, \mu) + \mathrm{Ent}((f(i_1) + \widehat{f}_{f(i_1)})^2, \mu)$; this is true only under additional conditions concerning the signs of $y'_{f(i_1)} - x'_{f(i_1)}$ and $y - x_{f(i_1)}$ (a similar observation holds at $s = f(i_0)$). The possibility for $y'_s - x'_s$ and $y - x_s$ to change sign when s ranges over$[f(i_1), f(i_0)]$ (the worst case is when such changes precisely occur where $\mu[\widehat{f}_s(\widehat{f}_s + 2s)]$ vanishes) is as much a nuisance as the the factor $s - \mu[\widetilde{f}_s]$ which appeared in (4). Therefore we are a priori not sure of the existence of some $s \in [f(i_1), f(i_0)]$ making one of the functions $\widetilde{f}_s$ and $s + \widehat{f}_s$ "strictly more maximising" than f. On the opposite, in the spectral gap case, this conclusion was nonetheless reachable, by using the extra fact that the map $[f(i_1), f(i_0)] \ni s \mapsto s - \mu[\widetilde{f}_s]$ is increasing (more precisely, a further analysis easily shows that $[f(i_1), f(i_0)] \ni s \mapsto s - \mu[\widetilde{f}_s]$ is increasing).

4 Continuous Situation

So we come back to the framework first considered in the introduction. We shall only deal with the case of the logarithmic Sobolev constant; the Poincaré constant can be treated in a very similar way. As already explained, the continuous situation will be reduced to the discrete one, thus giving the proof a slight probabilistic touch. We shall also consider the other possibility, to adapt the previous proofs, which leads to further analysing the (almost) minimising functions. But whichever way is chosen, the beginning of the proof appears to need some regularization as its first step.

For $M > 0$, let $\mathcal{C}_{[-M,M]}$ (respectively $\mathcal{D}_{[-M,M]}$) be the sub-set of $\mathcal{C}$ (respectively of $\mathcal{D}$) consisting of the absolutely continuous functions with weak derivative a.e. null on $]-\infty, -M] \cup [M, +\infty[$. Also, put

$$C_{[-M,M]}(\mu,\nu) := \sup_{f \in \mathcal{C}_{[-M,M]}} \frac{\mathrm{Ent}(f^2,\mu)}{\nu[(f')^2]}$$

$$D_{[-M,M]}(\mu,\nu) := \sup_{f \in \mathcal{D}_{[-M,M]}} \frac{\mathrm{Ent}(f^2,\mu)}{\nu[(f')^2]}.$$

One is easily convinced that these two quantities increase with $M > 0$ and that they respectively converge for large M to $C(\mu,\nu)$ and

$$D(\mu,\nu) := \sup_{f \in \mathcal{D}} \frac{\mathrm{Ent}(f^2,\mu)}{\nu[(f')^2]} \quad \in \bar{\mathbb{R}}_+.$$

Call $\nu_{[-M,M]}$ the restriction of ν to $[-M,M]$ (it vanishes outside this interval) and $\mu_{[-M,M]}$ the probability obtained by accumulating on the endpoints $-M$ and M the mass outside $[-M,M]$; i.e., $\mu_{[-M,M]}$ is defined by

$$\mu_{[-M,M]}(B) := \mu(B\cap\,]-M,M[) + \mu(]-\infty,M])\delta_{-M}(B) + \mu([M,+\infty[)\delta_M(B)$$

for B any Borel set in $\mathbb{R}$. The interest of these measures is that $C_{[-M,M]}(\mu,\nu) = C(\mu_{[-M,M]},\nu_{[-M,M]})$ and $D_{[-M,M]}(\mu,\nu) = D(\mu_{[-M,M]},\nu_{[-M,M]})$, so the convergences seen above allow restriction to the case that μ and ν are supported in the compact $[-M,M]$, where $M>0$ is fixed from now on. We shall also content ourselves with only considering functions defined on $[-M,M]$.

Denote by λ the restriction of the Lebesgue measure to $[-M,M]$ and, by abuse of language, still call ν the Radon-Nikodym derivative of ν with respect to λ (which exists without any restriction on ν, provided the value $+\infty$ is allowed; see for instance [11]). As weak derivatives are only a.e. defined, it is well known that $C(\mu,\nu)$ (or $D(\mu,\nu)$) is not modified when ν is replaced with the measure having ν as density with respect to λ, which we henceforth assume. One can also without loss suppose the function ν to be minorated by an a.e. strictly positive constant. Indeed, this derives from the fact that for any $f\in\mathcal{C}$, one has

$$\lim_{\eta\to 0_+}\frac{\mathrm{Ent}(f^2,\mu)}{\int (f')^2\,(\eta\wedge\nu)d\lambda} = \frac{\mathrm{Ent}(f^2,\mu)}{\nu[(f')^2]}$$

and that this convergence is monotone. So, by exchanging suprema, equality is preserved in the limit. Hence $\eta>0$ wil be fixed in the sequel, so that $\nu\geqslant\eta$ everywhere on $[-M,M]$, i.e., a suitable version of ν is chosen; but beware, ν may still assume the value $+\infty$ (remark that obtaining the corresponding majorization of ν would be more delicate).

The next procedure consists in modifying μ and is a little less immediate; a general preparation is needed:

Lemma 5. *On some measurable space, let μ be a probability and f and g two bounded, measurable functions. Suppose that $\|g-f\|_\infty\leqslant\epsilon\leqslant 1$ (uniform norm) and that the oscillation of f (i.e., $\mathrm{osc}(f) := \sup f - \inf f$) is majorized by a, where ϵ and a are positive real numbers. Then there exists a number $b(a)\geqslant 0$, depending only upon a, such that*

$$\left|\mathrm{Ent}(g^2,\mu) - \mathrm{Ent}(f^2,\mu)\right| \leqslant b(a)\,\epsilon.$$

Proof. Note that $|f|$ and $|g|$ fulfill the same hypotheses as f and g; so no generality is loss by further supposing f and g to be positive.

Two situations are then distinguished, according to $\mu[f]$ being "large" or "small". We shall start with the case when $\mu[f]\leqslant 2+2a$. This ensures that f is majorized by $2+3a$ and g by $3+3a$. Now, on the interval $[0,3+3a]$, the derivative of the map $t \mapsto t^2 \ln(t^2)$ is bounded by a finite quantity $b_1(a)$; this entails that

$$\begin{aligned}\big|\mu[g^2 \ln(g^2)] - \mu[f^2 \ln(f^2)]\big| &\leqslant \mu\big[\big|g^2 \ln(g^2) - f^2 \ln(f^2)\big|\big] \\ &\leqslant b_1(a)\, \mu[|g-f|] \leqslant b_1(a)\, \epsilon.\end{aligned}$$

Similarly, the norm inequality $\left|\sqrt{\mu[g^2]} - \sqrt{\mu[f^2]}\right| \leqslant \sqrt{\mu[(g-f)^2]}$ in $\mathbb{L}^2(\mu)$ yields

$$\big|\mu[g^2]\ln(\mu[g^2]) - \mu[f^2]\ln(\mu[f^2])\big| \leqslant b_1(a)\, \epsilon,$$

wherefrom finally the claimed inequality with $b(a) = 2b_1(a)$.

Consider now the case when $\mu[f] > 2 + 2a$. It seems more convenient to deal with the map $\mathbb{R}_+ \ni t \mapsto t\ln(t)$. Performing an expansion with first-order remainder, centred at $\mu[f^2]$, one finds a $\theta \in [0,1]$ such that $\mu[g^2]\ln(\mu[g^2])$ equals

$$\mu[f^2]\ln\big(\mu[f^2]\big) + \big(1 + \ln\big[\mu[f^2] + \theta(\mu[g^2] - \mu[f^2])\big]\big)\,\big(\mu[g^2] - \mu[f^2]\big).$$

The same operation performed pointwise yields another measurable function $\widetilde{\theta}$ with values in $[0,1]$ such that one has everywhere

$$g^2\ln(g^2) = f^2\ln(f^2) + \big(1 + \ln\big(f^2 + \widetilde{\theta}(g^2 - f^2)\big)\big)\,(g^2 - f^2).$$

Integrating this against μ and taking into account the preceding equality, it appears that

$$\begin{aligned}&\mathrm{Ent}(g^2,\mu) - \mathrm{Ent}(f^2,\mu) \\ &= \mu\left[\Big(\ln\big(f^2 + \widetilde{\theta}(g^2 - f^2)\big) - \ln\big[\mu[f^2] + \theta\big(\mu[g^2] - \mu[f^2]\big)\big]\Big)(g^2 - f^2)\right]. \quad (6)\end{aligned}$$

However, observe that

$$\begin{aligned}f^2 + \widetilde{\theta}(g^2 - f^2) \geqslant f^2 \wedge g^2 &\geqslant \big(\mu[f] - \mathrm{osc}(f) - 1\big)^2 \\ &\geqslant \big(\mu[f] - a - 1\big)^2 \geqslant \frac{\mu[f]^2}{4}\end{aligned}$$

and similarly

$$\mu[f^2] + \theta\big(\mu[g^2] - \mu[f^2]\big) \geqslant \frac{\mu[f]^2}{4}.$$

So one obtains the pointwise inequality

$$\begin{aligned}&\Big|\ln\big(f^2 + \widetilde{\theta}(g^2 - f^2)\big) - \ln\big(\mu[f^2] + \theta\big(\mu[g^2] - \mu[f^2]\big)\big)\Big| \\ &\qquad\leqslant 4\,\mu[f]^{-2}\,\Big|f^2 + \widetilde{\theta}(g^2 - f^2) - \mu[f^2] - \theta\big(\mu[g^2] - \mu[f^2]\big)\Big|.\end{aligned}$$

Let us look at the last absolute value. It can be majorized by

$$\begin{aligned}\big(f+\sqrt{\mu[f^2]}\big)\big|f-\sqrt{\mu[f^2]}\big| + (f+g)\,|f-g| +& \\ \big(\sqrt{\mu[g^2]}+\sqrt{\mu[f^2]}\big)\ &\Big|\sqrt{\mu[g^2]}-\sqrt{\mu[f^2]}\Big| \\ \leqslant\ 2\big(\mu[f]+a\big)\,a + (2\mu[f]+2a+1)\,\epsilon + (2\mu[f]+2a+1)\,\epsilon& \\ \leqslant\ (2\mu[f]+2a+1)(a+2).&\end{aligned}$$

On the other hand, one has as above

$$\big|g^2-f^2\big| \leqslant \big(2\mu[f]+2a+1\big)\,\epsilon,$$

wherefrom, coming back to (6), it appears that

$$\big|\mathrm{Ent}(g^2,\mu)-\mathrm{Ent}(f^2,\mu)\big| \leqslant 4\,\frac{(a+2)\big(2\mu[f]+2a+1\big)^2}{\mu[f]^2}\,\epsilon$$

and in that case the lemma holds with $b(a)=b_2(a)$, where

$$b_2(a) \coloneqq \sup_{t\geqslant 2+2a} 4\,\frac{(a+2)(2t+2a+1)^2}{t^2} \quad < +\infty. \qquad \square$$

This technical result will be used to measure how certain modifications of μ influence $C(\mu,\nu)$. More precisely, for fixed $n\in\mathbb{N}^*$, for any $0\leqslant i\leqslant n$ put $x_{n,i}\coloneqq -M+i2M/n$ and introduce the probability

$$\mu_n \coloneqq \sum_{0\leqslant i\leqslant n} \mu\big([x_{n,i},x_{n,i+1}[\big)\,\delta_{x_{n,i}}$$

with the convention that $x_{n,n+1}=+\infty$.

Lemma 6. *With the notation of Lemma 5, for all $n\in\mathbb{N}^*$ one has*

$$\big|C(\mu_n,\nu)-C(\mu,\nu)\big| \leqslant b\big(\sqrt{2M}\big)\sqrt{\frac{2M}{n}}.$$

Proof. Calling $\mathcal{C}(\nu)$ the set of absolutely continuous functions f such that $\nu[(f')^2]=1$, one has

$$C(\mu,\nu) = \sup_{f\in\mathcal{C}(\nu)} \mathrm{Ent}(f^2,\mu)$$

and one also has a similar formula for $C(\mu_n,\nu)$. Thus, to obtain the claimed bound, it suffices to see that for all $f\in\mathcal{C}(\nu)$, one has

$$\big|\mathrm{Ent}(f^2,\mu_n)-\mathrm{Ent}(f^2,\mu)\big| \leqslant b\big(\sqrt{2M}\big)\sqrt{\frac{2M}{n}}.$$

To that end, rewrite $\mathrm{Ent}(f^2, \mu_n)$ as $\mathrm{Ent}(f_n^2, \mu)$, where f_n is the function which equals $f(x_{n,i})$ on $[x_{n,i}, x_{n,i+1}[$ for all $0 \leqslant i \leqslant n$. To apply Lemma 5, it remains to evaluate $\mathrm{osc}(f)$ and $\|f_n - f\|_\infty$. These estimates, and consequently also the claimed result, easily follow from the following application of the Cauchy-Schwarz inequality:

$$\forall\, x, y \in [-M, M],$$
$$\begin{aligned}\big|f(y) - f(x)\big| = \left|\int_{[x,y]} f'\, d\lambda\right| &\leqslant \sqrt{\int_{[x,y]} (f')^2\, d\nu}\ \sqrt{\int_{[x,y]} \frac{1}{\nu}\, d\lambda} \\ &\leqslant \eta^{-1/2}\sqrt{|y-x|},\end{aligned}$$

where the last estimate holds for any function belonging to $\mathcal{C}(\nu)$. □

Evidently, the above proof also shows that

$$|D(\mu_n, \nu) - D(\mu, \nu)| \leqslant b\big(\sqrt{2M}\big)\sqrt{\frac{2M}{n}}\ ;$$

so, to get convinced of the equality $C(\mu, \nu) = D(\mu, \nu)$, it suffices to see that $C(\mu_n, \nu) = D(\mu_n, \nu)$ for all $n \in \mathbb{N}^*$. But this problem reduces to the discrete context. Indeed, as before Lemma 1, the values of $f(x_{n,i})$ being fixed, one has to minimise the quantity $\int_{x_{n,i}}^{x_{n,i+1}} (f')^2\, \nu\, d\lambda$ for each given $0 \leqslant i < n$. This optimisation problem is simply solved; the minimal value is

$$\left(\int_{x_{n,i}}^{x_{n,i+1}} \frac{1}{\nu}\, d\lambda\right)^{-1} \big(f(x_{n,i+1}) - f(x_{n,i})\big)^2$$

and is achieved by a function which is monotone on the segment $[x_{n,i}, x_{n,i+1}]$. Hence we are back to the discrete problem on $n+1$ points with the probability $\widetilde{\mu}_n$ and the measure $\widetilde{\nu}_n$ respectively defined by

$$\forall\, 0 \leqslant i \leqslant n, \qquad \widetilde{\mu}_n(i) := \mu_n(x_{n,i})$$
$$\forall\, 0 \leqslant i < n, \qquad \widetilde{\nu}_n\big(\{i, i+1\}\big) := \left(\int_{x_{n,i}}^{x_{n,i+1}} \frac{1}{\nu}\, d\lambda\right)^{-1}.$$

Sections 2 and 3 now allow to conclude.

From a possibly more analytically-minded point of view, remark that Lemmas 5 and 6 could also allow to regularize μ, which could be supposed to admit a $\mathcal{C}^\infty$ density with respect to λ.

Let us now mention another possible approach, directly inspired from the method of sections 2 and 3. A priori two problems arise in this perspective: on the one hand, whether a minimising function exists (even in the case of the Poincaré inequality), and on the other hand, when it exists, whether the set of its global minima and maxima can have infinitely many connected components (this means, the function oscillates infinitely often; this is inconvenient for us, see the considerations before Lemma 2). These problems can be bypassed as follows. We put ourselves back in the framework preceding Lemma 5.

First, the notion of local minimum or maximum introduced in section 2 will be extended to the continuous case, with discrete segments replaced by continuous ones. For $f \in \mathcal{C}$, $\mathcal{M}(f)$ will denote the set of local minima and maxima of f. For $p \in \mathbb{N}^*$, call $\mathcal{C}_p$ the set of functions $f \in \mathcal{C}$ such that $\mathcal{M}(f)$ has at most p connected components. So one verifies that $\mathcal{C}_1$ (respectively $\mathcal{C}_2$) is the set of constant (respectively monotone) functions. Set also $\mathcal{C}_\infty := \cup_{p\in\mathbb{N}^*}\mathcal{C}_p$, for which one has the following preliminary result:

Lemma 7. *One has*

$$C(\mu,\nu) = \sup_{f\in\mathcal{C}_\infty} \frac{\operatorname{Ent}(f^2,\mu)}{\nu[(f')^2]}.$$

Proof. Let $\mathcal{F}$ denote the set of all measurable functions $g : [-M,M] \to \mathbb{R}$ belonging to $\mathbb{L}^1([-M,M],\lambda)$ and for which one can find $n \in \mathbb{N}^*$ and $-M = x_0 < x_1 < \cdots < x_n = M$ such that for all $0 \leqslant i < n$, g has a constant sign on $]x_i, x_{i+1}[$ (0 is considered as having at the same time a positive and negative sign). So $\mathcal{C}_\infty$ is nothing but the set of antiderivatives of elements of $\mathcal{F}$.

It then suffices to verify that $\{g \in \mathcal{F} : \nu[g^2] \leqslant 1\}$ is dense in the $\mathbb{L}^2(\nu)$ sense in the unit ball of this space. Indeed, let $f \in \mathcal{C}$ with $\nu[(f')^2] = 1$. According to the preceding property, there exists a sequence $(g_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{F}$ converging to f'. Put for all $n \in \mathbb{N}$

$$\forall\, x \in [-M,M], \qquad G_n(x) = f(-M) + \int_{-M}^{x} g_n(y)\, dy.$$

Due to the minorization $\nu \geqslant \eta$, it is clear that the G_n converge uniformly to f for large n. And since $\operatorname{osc}(f) < +\infty$, Lemma 5 applies and shows that

$$\lim_{n\to\infty} \operatorname{Ent}(G_n^2,\mu) = \operatorname{Ent}(f^2,\mu),$$

wherefrom follows the equality in the lemma.

To show the claimed density, take $g \in \mathbb{L}^2(\nu)$ with $\nu[g^2] = 1$; for $n \in \mathbb{N}$, put

$$g_n := g\mathbb{1}_{\{\nu\leqslant n, |g|\leqslant n\}}.$$

By dominated convergence, the sequence $(g_n)_{n\in\mathbb{N}}$ converges in $\mathbb{L}^2(\nu)$ to g. Now, for fixed $n \in \mathbb{N}$, the measure $(\nu \wedge n)d\lambda$ is regular (in the sense of inner and outer approximation of Borel sets; see for instance Rudin's book [16]), so one can find a sequence $(\widetilde{g}_{n,m})_{m\in\mathbb{N}}$ in $\mathcal{F}$ such that

$$\lim_{m\to\infty} \int (\widetilde{g}_{n,m} - g_n)^2\, (\nu \wedge n)\, d\lambda = 0.$$

So, setting for all $m \in \mathbb{N}$, $\widehat{g}_{n,m} := \widetilde{g}_{n,m} \mathbb{1}_{\{\nu \leqslant n, |g| \leqslant n\}}$, which still belongs to $\mathcal{F}$, one also has

$$\lim_{m \to \infty} \int (\widehat{g}_{n,m} - g_n)^2 \, d\nu = 0$$

and the claimed density is established. □

The lemma entails that

$$C(\mu, \nu) = \lim_{p \to \infty} \sup_{f \in \mathcal{C}_p} \frac{\operatorname{Ent}(f^2, \mu)}{\nu[(f')^2]}.$$

However, for $p \geqslant 3$ and $f \in \mathcal{C}_p \setminus \mathcal{C}_2$, the considerations from the preceding section applied to f yield $\widetilde{f} \in \mathcal{C}_{p-1}$ and $\widehat{f} \in \mathcal{C}_4$ such that

$$\begin{aligned} \nu[(f')^2] &= \nu[(\widetilde{f}')^2] + \nu[(\widehat{f}')^2] \\ \operatorname{Ent}(f^2, \mu) &= \operatorname{Ent}(\widetilde{f}^2, \mu) + \operatorname{Ent}(\widehat{f}^2, \mu). \end{aligned}$$

Let us make this more precise. For $g \in \mathcal{C}$, a connected component of $\mathcal{M}(g)$ will be called internal if it contains neither $-M$ nor M. The union of the internal connected components of $\mathcal{M}(g)$ will be denoted by $\widetilde{\mathcal{M}}(g)$. One then introduces a set $C_3 \subset \widehat{C}_4 \subset C_4$ by imposing that $\widehat{C}_4 \cap (C_4 \setminus C_3)$ consists of the functions $g \in C_4 \setminus C_3$ such that $\min_{\widetilde{\mathcal{M}}(g)} g \leqslant g(-M), g(M) \leqslant \max_{\widetilde{\mathcal{M}}(g)} g$. The interest of this set $\widehat{C}_4$ will be twofold for us: on the one hand, in the above construction, one has $\widehat{f} \in \widehat{C}_4$, and on the other hand, if $g \in \widehat{C}_4 \setminus C_2$ then $\widetilde{g}$ obtained from the preceding procedure is monotone.

However, the sole fact that $\widehat{f} \in \mathcal{C}_4$ already showed that for $p \geqslant 5$, one has

$$\sup_{f \in \mathcal{C}_p} \frac{\operatorname{Ent}(f^2, \mu)}{\nu[(f')^2]} = \sup_{f \in \mathcal{C}_{p-1}} \frac{\operatorname{Ent}(f^2, \mu)}{\nu[(f')^2]},$$

and by induction, one ends up with the fact that this quantity is nothing but $\sup_{f \in \mathcal{C}_4} \operatorname{Ent}(f^2, \mu)/\nu[(f')^2]$. More precisely, the preceding observations even imply that

$$C(\mu, \nu) = \sup_{f \in \widehat{\mathcal{C}}_4} \frac{\operatorname{Ent}(f^2, \mu)}{\nu[(f')^2]}.$$

So let $(f_n)_{n \in \mathbb{N}}$ be a sequence of elements from $\widehat{\mathcal{C}}_4$ satisfying $\nu[(f_n')^2] = 1$ for all $n \in \mathbb{N}$ and $C(\mu, \nu) = \lim_{n \to \infty} \operatorname{Ent}(f_n^2, \mu)$. Two situations can be distinguished: either one can extract from $(f_n)_{n \in \mathbb{N}}$ a subsequence (still denoted $(f_n)_{n \in \mathbb{N}}$) such that $\big(f_n(0)\big)_{n \in \mathbb{N}}$ converges in $\mathbb{R}$, or one has $\liminf_{n \to \infty} |f_n(0)| = +\infty$. The latter case corresponds to the equality $C(\mu, \nu) = A(\mu, \nu)/2$, whose treatment amounts to that of the Poincaré constant, left to the reader. Thus, from now on, we assume to be in the first situation described above. By weak compactness of the unit ball of $\mathbb{L}^2(\nu)$, one can extract a subsequence of $(f_n)_{n \in \mathbb{N}}$, such that $(f_n')_{n\in\mathbb{N}}$ is weakly convergent in $\mathbb{L}^2(\nu)$. Together with the convergence of $(f_n(0))_{n\in\mathbb{N}}$, this weak convergence implies that the sequence $(f_n)_{n\in\mathbb{N}}$ converges pointwise on $[-M,M]$ to a function f which has a weak derivative f' satisfying $\nu[(f')^2]\leqslant 1$ (because the norm is weakly lower semi-continuous). However, the uniform continuity of the f_n for $n\in\mathbb{N}$ (due to the majorization by $\eta^{-1/2}$ of their Hölder coefficient of order $1/2$) ensures, via Ascoli's theorem, that the convergence of the f_n towards f is in fact uniform on the compact $[-M,M]$. In particular, one obtains

$$\mathrm{Ent}(f^2,\mu)=\lim_{n\to\infty}\mathrm{Ent}(f_n^2,\mu) \;=\; C(\mu,\nu).$$

Discarding the trivial situation that $C(\mu,\nu)=0$ (which corresponds to the cases when μ is a Dirac mass or $\nu=+\infty$ a.s. on the convex hull of the support of μ), one then obtains

$$\frac{\mathrm{Ent}(f^2,\mu)}{\nu[(f')^2]}\geqslant C(\mu,\nu),$$

with strict inequality if $0\leqslant\nu[(f')^2]<1$, wherefrom necessarily $\nu[(f')^2]=1$. So f is a maximising function for (1), which, moreover, belongs to $\widehat{\mathcal{C}}_4$, whereof one is easily convinced: at the cost of extracting a subsequence, one can require that the number (between 0 and 2) of internal connected components is the same for each f_n and that there exists a point in each of these components which converges in $[-M,M]$ for large n, and this allows to see a posteriori that $f\in\widehat{\mathcal{C}}_4$). If f is not already monotone, the procedure of the preceding section can be applied again to construct $\widetilde{f}$ and $\widehat{f}$. As f is maximising, so must be these two functions too; now, owing to f belonging to $\widehat{\mathcal{C}}_4$, $\widetilde{f}$ is necessarily monotone. So these arguments allow to conclude that $C(\mu,\nu)=D(\mu,\nu)$.

Remark 2. The latter proof rests partially on the existence of a maximising function for (1), but, contrary to the approach by Chen and Wang [6, 8] (in the case of the Poincaré constant), we have not tried to exploit the equation it fulfills.

More generally, call $S(\mu)$ the convex hull of the support of μ and $[s_-,s_+]$ its closure in the compactified real line $\mathbb{R}\sqcup\{-\infty,+\infty\}$. Still denoting by ν the density of ν with respect to λ, assume that

$$\int_{S(\mu)}\frac{1}{\nu}\,d\lambda<+\infty.$$

One can then show that if $C(\mu,\nu)>A(\mu,\nu)/2$, a maximising function for (1) exists (but these two conditions are not sufficient as can be seen by taking for μ and ν the standard Gaussian distribution). Indeed, fix $o\in S(\mu)$ and define

$$\forall\; x\in S(\mu),\qquad F(x):=\int_o^x\frac{1}{\nu(y)}\,dy.$$

By the preceding condition, F is continuously extendable to $[s_-, s_+]$. Consider then an absolutely continuous function f whose weak derivative satisfies $\int (f')^2\, d\nu \leqslant 1$. Applying as above a Cauchy-Schwarz, inequality, one gets that

$$\forall\, x, y \in S(\mu), \qquad |f(y) - f(x)| \leqslant \sqrt{|F(y) - F(x)|},$$

and consequently, by Cauchy's criterion, f too is continuously extendable to $[s_-, s_+]$. One can then repeat the preceding arguments on this compact (taking into account that $\nu^{-1}\mathbb{1}_I \in \mathbb{L}^2(S(\mu), \nu)$ for each segment $I \subset [s_-, s_+]$, this alowing to obtain pointwise convergence from the weak compactness of the unit ball of $\mathbb{L}^2(S(\mu), \nu)$), and see that except when $C(\mu, \nu) = A(\mu, \nu)/2$, there exists a maximising function f for (1) (and since it is known that dealing with monotone functions is sufficient, Ascoli's theorem can even be replaced with one of Dini's ones). Performing a variational calculation around this function, one realizes that it satisfies two conditions:

$$\int_{S(\mu)} f\, \ln\Big(\frac{f^2}{\mu[f^2]}\Big)\, d\mu = 0$$

and for a.a. $x \in S(\mu)$,

$$C(\mu, \nu)\nu(x) f'(x) = \int_{[s_-, x]} f\, \ln\Big(\frac{f^2}{\mu[f^2]}\Big)\, d\mu. \tag{7}$$

Obviously, if moreover the function ν is assumed to be absolutely continuous and μ absolutely continuous with respect to λ, a further differentiation yields a second-order equation (non linear in the zeroth order term) satisfied by f.

Last, if in addition $[s_-, s_+] \subset \mathbb{R}$, $\nu(s_-) > 0$ and $\nu(s_+) > 0$, equation (7) allows to recover a Neumann condition for f, namely $f'(s_-) = f'(s_+) = 0$.

5 Extensions

We present here a few generalisations of the preceding results, corresponding to modifications of the quantities featuring in (1).

5.1 Modification of the Energy in the Discrete Case

We shall show here Theorem 3, whose context is now assumed, and we put

$$E(\mu, \nu) \coloneqq \sup_{f \in \mathcal{C}} \frac{\mathrm{Ent}(f^2, \mu)}{\mathcal{E}_\nu\big(f^2, \ln(f^2)\big)}.$$

Considering $\mathbb{Z}$ brings no further difficulty, since, as in section 4, one can without loss consider only the finite situation where $E = \{0, ..., N\}$ with $N \in \mathbb{N}^*$, at the cost of accumulating mass on the endpoints and translating the obtained segment. However, we take this opportunity to point out the most famous infinite example where the preceding constant is finite, namely the Poisson laws on $\mathbb{N}$: fix $\alpha > 0$ and take

$$\forall\, l \in \mathbb{N}, \qquad \mu\big(\{l\}\big) := \frac{\alpha^l}{l!}\exp(-\alpha)$$
$$\nu\big(\{l, l{+}1\}\big) := \mu\big(\{l\}\big).$$

It is then known (see for instance section 1.6 of the book [1] by Ané, Blachère, Chafaï, Fougères, Gentil, Malrieu, Roberto and Scheffer) that $E(\mu,\nu)$ equals α.

To get convinced of Theorem 3, on has to inspect again the three-step proof in sections 2 and 3.

• As in the case of the logarithmic Sobolev inequality, one is brought back, up to a multiplicative constant, to the problem of estimating the Poincaré constant when there exists a minimising sequence $(f_n)_{n\in\mathbb{N}}$ verifying

$$\forall\, n \in \mathbb{N}, \qquad \mathcal{E}_\nu(f_n^2, \ln(f_n^2)) = 1$$
$$\lim_{n\to\infty} |f_n(0)| = +\infty.$$

Indeed, it is well known (see for instance Lemma 2.6.6 in the book by Ané and al. [1]) that

$$\forall\, f \in \mathcal{C}, \qquad \mathcal{E}_\nu(f^2, \ln(f^2)) \geqslant 4\nu[(f')^2];$$

so the first condition above ensures that the oscillations of the f_n are bounded in $n \in \mathbb{N}$ (the situation should have been beforehand reduced to the case when $\nu > 0$). This observation allows to perform finite order expansions showing the following equivalent for large n:

$$\frac{\mathrm{Ent}(f_n^2,\mu)}{\mathcal{E}_\nu\big(f_n^2, \ln(f_n^2)\big)} \sim \frac{\mathrm{Var}(f_n,\mu)}{8\nu[(f_n')^2]},$$

wherefrom one easily deduces

$$\sup_{f\in\mathcal{C}} \frac{\mathrm{Ent}(f^2,\mu)}{\mathcal{E}_\nu\big(f^2, \ln(f^2)\big)} = \frac{A(\mu,\nu)}{8} = \sup_{f\in\mathcal{D}} \frac{\mathrm{Ent}(f^2,\mu)}{\mathcal{E}_\nu\big(f^2, \ln(f^2)\big)}.$$

Thus it suffices to consider the situations where there exists a minimising sequence $(f_n)_{n\in\mathbb{N}}$ such that

$$\forall\, n \in \mathbb{N}, \qquad \mathcal{E}_\nu\big(f_n^2, \ln(f_n^2)\big) = 1$$
$$\limsup_{n\to\infty} |f_n(0)| < \infty,$$

in which cases one can extract a subsequence that converges toward a maximiser for the supremum we are interested in.

• Calling f this maximiser, one is easily convinced that it cannot vanish, at least in the relevant situations where $E(\mu,\nu) > 0$. Performing then a variational computation around f shows it to verify for each $i \in E$ the following equation:

$$\begin{aligned}
&\mu(i)\, f(i)\, \ln\Big(\frac{f^2(i)}{\mu[f^2]}\Big) \\
&\quad = E(\mu,\nu)\Big[f(i)\big[\nu\big(\{i,i{+}1\}\big)\big(\ln(f^2(i)) - \ln(f^2(i{+}1))\big) \\
&\qquad + \nu\big(\{i{-}1,i\}\big)\big(\ln(f^2(i)) - \ln(f^2(i{-}1))\big)\big] \\
&\qquad + \frac{\nu\big(\{i,i{+}1\}\big)\big(f^2(i) - f^2(i{+}1)\big) + \nu\big(\{i{-}1,i\}\big)\big(f^2(i) - f^2(i{-}1)\big)}{f(i)}\Big]
\end{aligned}$$

(as usual, $\nu(\{-1,0\}) = 0 = \nu(\{N,N{+}1\})$, hence the terms $f(-1)$ and $f(N{+}1)$ never show up). If μ does not vanish, the form of this equation enables to apply the arguments of the end of section 3, taking advantage of the fact that a maximising function for $E(\mu,\nu)$ cannot take the same value at three consecutive points, unless it is constant (which won't do either). Remark also that contrary to sections 2 and 3, this equation does not allow to recursively compute f from the values of $f(0)$ and $\mu[f^2]$, for the right-hand side is not injective as a function of $f(i{+}1)$ (for $0 \leqslant i < N$), but only as a function of $f^2(i{+}1)$. But this could be forseeen, since the signs of the functions really play no role in the quantities considered here. There remain the cases when μ vanishes at some (interior) points; they cannot be discarded as before Lemma 1. The simplest is to bypass the argument of the consecutive three points with same value, by adapting the second proof of the preceding section (by classifying the functions according to the maximal number of segments included in their set of local extrema); this is immediate enough.

• The last point to be verified, which is also the most important, is the possibility of modifying Lemma 2; namely, with the notations therein, is it true that for all $s \in \,]f(i_1), f(i_0)[$,

$$\mathcal{E}_\nu\big(f^2, \ln(f^2)\big) \geqslant \mathcal{E}_\nu\big((\widetilde{f}'_s)^2, \ln((\widetilde{f}'_s)^2)\big) + \mathcal{E}_\nu\big((\widehat{f}'_s)^2, \ln((\widehat{f}'_s)^2)\big) \tag{8}$$

for any function f with a constant sign (the situation should have been reduced to that case). This question amounts to asking if for all $0 \leqslant x \leqslant y \leqslant z$, one has

$$\varphi_{x,z}(y) \leqslant (z-x)\big(\ln(z) - \ln(x)\big), \tag{9}$$

where $\varphi_{x,z}$ is the function defined by

$$\forall\, y \in [x,z], \qquad \varphi_{x,z}(y) \coloneqq (y-x)\big(\ln(y) - \ln(x)\big) + (z-y)\big(\ln(z) - \ln(y)\big).$$

Now, differentiating this function twice shows it to be strictly convex, and (9) then derives from the fact that $\varphi_{x,z}(x) = \varphi_{x,z}(z) = (z-x)\big(\ln(z) - \ln(x)\big)$. One also derives therefrom that equality in (8) can hold only if $\widetilde{f}'_s(a)\widehat{f}'_s(a) = 0$ for every edge $a \in A$.

The other arguments of section 3 are valid without modification, since they only involve entropy. Theorem 3 follows.

5.2 Modification of the Energy in the Continuous Case

Our aim here is to prove Theorem 4. Recall that $H : \mathbb{R}_+ \to \mathbb{R}_+$ is a convex function such that $H(0) = 0$ and $H'(0) = 1$ (besides these two equalities, we shall only use the bound $x \leqslant H(x)$, valid for all $x \geqslant 0$). In particular, it appears that

$$\forall\, f \in \mathcal{C}, \qquad \mathcal{E}_{H,\nu}(f) \geqslant \nu\big[(f')^2\big]. \tag{10}$$

For μ a probability and ν a measure on $\mathbb{R}$, put

$$F(\mu,\nu) := \sup_{f\in\mathcal{C}} \frac{\mathrm{Ent}(f^2,\mu)}{\mathcal{E}_{H,\nu}(f)} \ \in\ \bar{\mathbb{R}}_+.$$

In view of the second proof in the preceding section, the only non immediate point in the proof of Theorem 4 concerns the cases that can be reduced to that of the Poincaré constant. Indeed, after having supposed without loss that μ is supported in $[-M,M]$ and that $\nu \geqslant \eta$, with $M,\eta > 0$, we have to see that if $(f_n)_{n\in\mathbb{N}}$ is a maximising sequence for $F(\mu,\nu)$ such that

$$\begin{aligned} \forall\, n\in\mathbb{N}, \qquad \mathcal{E}_{H,\nu}(f) &= 1 \\ \lim_{n\to\infty} |f_n(0)| &= +\infty, \end{aligned}$$

then $F(\mu,\nu) = A(\mu,\nu)/2$. But, again, such a sequence will satisfy $\nu[(f')^2] \leqslant 1$ for all $n \in \mathbb{N}$, and the oscillations of the f_n will be bounded, allowing to obtain for large n the equivalent

$$\mathrm{Ent}(f_n^2,\mu) \sim \frac{\mathrm{Var}(f_n,\mu)}{2}.$$

By extracting a subsequence (first, by relative compactness of the f_n, then, by Ascoli's theorem), one may suppose that the f_n converge uniformly to $f \in \mathcal{C}$, with $\nu[(f')^2] \leqslant 1$, wherefrom

$$\begin{aligned} F(\mu,\nu) = \lim_{n\to\infty} \mathrm{Ent}(f_n^2,\mu) &= \lim_{n\to\infty} \frac{\mathrm{Var}(f_n,\mu)}{2} \\ &= \frac{\mathrm{Var}(f,\mu)}{2} \leqslant \frac{\mathrm{Var}(f,\mu)}{2\nu[(f')^2]} \leqslant \frac{A(\mu,\nu)}{2}. \end{aligned}$$

However, the reverse inequality always holds. Indeed, note first that one may content oneself in only dealing, for the supremum defining $A(\mu,\nu)$, with functions having a weak derivative essentially bounded in the sense of the Lebesgue measure on $[-M, M]$. This is because only functions such that $\nu[(f')^2] < +\infty$ need to be considered, and such functions can be approximated in the traditional way. Let $f \in \mathcal{C}$ with $f \geqslant 0$ and f' bounded. For $n \in \mathbb{N}$, consider $f_n := n + f$. The oscillation of f being finite, for large n one has $\mathrm{Ent}(f_n^2, \mu) \sim \mathrm{Var}(f_n, \mu)/2 = \mathrm{Var}(f, \mu)/2$. On the other hand, since $H'(0) = 1$, one has by dominated convergence

$$\lim_{n\to\infty} \mathcal{E}_{H,\nu}(f_n) = \lim_{n\to\infty} \int H\left(\frac{(f')^2}{(n+f)^2}\right)(n+f)^2\, d\nu = \int (f')^2\, d\nu.$$

It ensues therefrom that

$$\frac{\mathrm{Var}(f, \mu)}{2\nu[(f')^2]} \leqslant F(\mu, \nu),$$

then the claimed inequality, by taking the supremum over such functions f.

Similar results hold when $\mathcal{C}$ is replaced with $\mathcal{D}$. It therefore suffices to deal with sequences $(f_n)_{n\in\mathbb{N}}$ maximising for $F(\mu, \nu)$, satisfying $\mathcal{E}_{H,\nu}(f_n) = 1$ for all $n \in \mathbb{N}$, and such that $\lim_{n\to\infty} f_n(0)$ exists in $\mathbb{R}$. But in this situation, the arguments in the second proof in section 4 easily adapt (after one has noted that for each function $f \in \mathcal{C}$ which splits as $\widetilde{f} + \widehat{f}$, with $\widetilde{f}, \widehat{f} \in \mathcal{C}$ and $\widetilde{f}'\widehat{f}' = 0$ a.s., one trivially has $\mathcal{E}_{H,\nu}(f) = \mathcal{E}_{H,\nu}(\widetilde{f}) + \mathcal{E}_{H,\nu}(\widehat{f})$).

Remark 3. One may wonder if there is a link between the discrete modified logarithmic Sobolev inequalities, and the continuous ones as above. As an attempt to shed light on such a link, consider again the approximation procedure used in the first proof of section 4. Thus we work with a probability μ of the form $\sum_{0\leqslant n\leqslant N} \mu(n)\delta_n$. The constant $F(\mu, \nu)$ can then be rewritten

$$\sup_{f\in\mathcal{C}} \frac{\mathrm{Ent}(f^2, \mu)}{\mathcal{E}_J(f)} \tag{11}$$

with for each $f \in \mathcal{C}$ in the discrete context

$$\mathcal{E}_J(f) := \sum_{0\leqslant n<N} J_{n,n+1}\big(f(n), f(n+1)\big)$$

et where the maps $(J_{n,n+1})_{0\leqslant n<N}$ are defined on $\mathbb{R}^2$ by

$$\forall\ x, y \in \mathbb{R}, \qquad J_{n,n+1}(x, y) := \inf_{\substack{g\in\mathcal{C}([n,n+1]):\\ g(n)=x,\, g(n+1)=y}} \int_n^{n+1} H\left(\left(\frac{g'}{g}\right)^2\right) g^2\, \nu\, d\lambda.$$

Obviously, the supremum (11) is not changed by restricting it to monotone functions, since this “discrete” problem can be interpreted in the continuous context where this property has just been verified. But one could certainly also show it directly; note in particular that for any $0 \leqslant n < N$ and all real numbers réels $x \leqslant y \leqslant z$, one has indeed$J_{n,n+1}(x,z) \geqslant J_{n,n+1}(x,y) + J_{n,n+1}(y,z)$ (it suffices to split any function going from x to z as the sum of two functions, the first one being its restriction going from x to y and remaining there).

This leads to ponder on the possibility of rewriting $\mathcal{E}_\nu$ as an $\mathcal{E}_J$, for a suitable choice of the continuous measure μ (the discrete one being given), and of the function H.

5.3 Modification of the Entropy

We now aim to change the entropy term in (1); this leads to logarithmic Sobolev inequalities modified in another sense (see for instance Chafaï [5]). This will give the opportunity to test the limits of the arguments in section 3. We shall content ourselves by treating the discrete case with the usual energy given by the quadratic form $\mathcal{C} \ni f \mapsto \nu[(f')^2]$, although one may think that similar considerations should allow to extend the following to the continuous situation or to energies modified as above. Let $\varphi : \mathbb{R}_+ \to \mathbb{R}$ be a convex function, of class $\mathcal{C}^3$ on $]0,+\infty[$. The corresponding modified entropy is the functional which to any map $f \in \mathcal{C}$, $f \geqslant 0$ associates the quantity (positive by Jensen's inequality)

$$\mathrm{E}_\varphi[f] = \mu\big[\varphi(f)\big] - \varphi\big(\mu[f]\big).$$

Unfortunately the expression $\mathrm{E}_\varphi(f^2)$ is no longer quadratically homogeneous in f (unless it is proportional to the usual entropy in f^2). To remedy this flaw, we shall need two additional hypotheses. Call ψ the map defined by

$$\forall\, x > 0, \qquad \psi(x) \coloneqq x\varphi'(x) - \varphi(x).$$

One says that ψ is asymptotically concave if for some $R > 0$ the function ψ remains below its tangents at points larger than R:

$$\forall\, y \geqslant R, \forall\, x > 0, \qquad \psi(x) \leqslant \psi(y) + \psi'(y)(y-x).$$

This notably implies that ψ is concave on $[R,+\infty[$ (which is not sufficient, but becomes sufficient if moreover $\lim_{x\to+\infty} \psi(x) - x\psi'(x) = +\infty$). We shall first suppose ψ to be asymptotically concave. The second additional hypothesis states the existence of a constant $\eta > 0$ such that for any $0 < x < \eta$, one has $\varphi''(x) + x\varphi'''(x) \geqslant 0$ (if φ is $\mathcal{C}^3$ on $\mathbb{R}_+$, this is ensured by $\varphi''(0) > 0$; more generally, if one does not even want to suppose φ to be of class $\mathcal{C}^3$ on $\mathbb{R}_+^*$, it can be seen that it suffices to suppose that the map $x \mapsto x\varphi''(x)$ is increasing on some interval $]0,\eta[$). An example of a function φ satisfying these conditions is $\mathbb{R}_+ \ni x \mapsto x\ln\big(\ln(e+x)\big)$.

Remark that

$$\forall\, x > 0, \qquad \psi'(x) = x\varphi''(x) \;\geqslant\; 0$$

and that this quantity decreases for $x \geqslant R$; hence it admits a limit $L \geqslant 0$ at $+\infty$. So $\varphi''(x) \leqslant (1+L)/x$ for x large, which shows that up to a constant factor, $\varphi(x)$ is dominated by $x\ln(x)$. Somehow, the usual entropy is an upper bound for the modified entropies to be considered here.

For μ a probability on $E = \{0, ..., N\}$ and ν a measure on the corresponding set A of edges, we are interested in the quantity

$$G(\mu, \nu) \coloneqq \sup_{f \in \mathcal{C}} \frac{\mathrm{E}_\varphi(f^2)}{\nu[(f')^2]}$$

and our aim here is to prove

Proposition 2. *One has as usual*

$$G(\mu, \nu) = \sup_{f \in \mathcal{D}} \frac{\mathrm{E}_\varphi(f^2)}{\nu[(f')^2]}.$$

The main annoyance comes from the inhomogeneity of E_φ, which a priori forbids to only consider maximising sequences for $G(\mu, \nu)$ with energy bounded above and below by a strictly positive constant. To remedy to that, observe that nothing here hinders us from supposing μ and ν to be strictly positive on E. This property ensures the existence of a constant $b_1 > 0$ such that

$$\forall\, g \in \mathcal{C}, \qquad \nu[(g')^2] = 1 \quad \Rightarrow \quad \mu[g^2] \geqslant b_1.$$

Fix a function g satisfying $\nu[(g')^2] = 1$ and consider the function

$$F\,:\, \mathbb{R}_+^* \ni t \mapsto \mathrm{E}_\varphi[tg^2]/t. \tag{12}$$

A computation gives its derivative as

$$\forall\, t > 0, \qquad F'(t) = t^{-2}\left(\mu[\psi(tg^2)] - \psi(t\mu[g^2])\right).$$

So by our hypothesis that ψ is asymptotically concave, F is decreasing on $[R/b_1, +\infty[$. This shows that

$$G(\mu, \nu) = \sup_{f \in \mathcal{C}\,:\,\nu[(f')^2] \leqslant R/b_1} \frac{\mathrm{E}_\varphi(f^2)}{\nu[(f')^2]},$$

which enables us to only consider maximising sequences $(f_n)_{n\in\mathbb{N}}$ satisfying $\nu[(f_n')^2] \leqslant R/b_1$ for all $n \in \mathbb{N}$. One can also suppose that these functions f_n are positive. Write $f_n = \sqrt{t_n}g_n$, with $t_n > 0$ (discarding the trivial cases that $t_n = 0$) and $g_n \in \mathcal{C}$ satisfying $\nu[(g_n')^2] = 1$. Extracting a sub-sequence reduces to the situation when the sequences $(t_n)_{n\in\mathbb{N}}$ and $\left(f_n(0)\right)_{n\in\mathbb{N}}$ are respectively convergent in $[0, R/b_1]$ and $\bar{\mathbb{R}}_+$. Several cases will be distinguished:

• If $\lim_{n\to\infty} t_n = 0$, we shall verify that we may without loss suppose that $\lim_{n\to\infty} f_n(0) > 0$. Indeed, our second hypothesis on ψ ensures that for $g \in \mathcal{C}$, $g \geqslant 0$, the function F defined in (12) is increasing on $]0, \eta/\max g^2]$. This is obtained via a second-order expansion with remainder: for fixed $t > 0$, there exists a function $\theta_t : E \to]0, t \max g^2[$ such that

$$\psi(tg^2) = \psi(t\mu[g^2]) + \psi'(t\mu[g^2])t(g^2 - \mu[g^2]) + \frac{\psi''(\theta_t)}{2}t^2(g^2 - \mu[g^2])^2.$$

When this inequality is integrated with respect to μ, it appears that $F'(t)$ is positive as soon as $t \max g^2 \leqslant \eta$. On the other hand, there exists a constant $b_2 > 0$ such that if g satisfies $\nu[(g')^2] = 1$, then $\text{osc}(g) \leqslant b_2$ and hence, if moreover g is positive, $\max g^2 \leqslant (g(0) + b_2)^2$. Consequently, if one constructs a new sequence $(\widetilde{t}_n)_{n\in\mathbb{N}}$ by setting

$$\forall\, n \in \mathbb{N}, \qquad \widetilde{t}_n := \begin{cases} t_n & \text{si } t_n(g_n(0) + b_2)^2 > \eta \\ \eta/(g_n(0) + b_2)^2 & \text{else,} \end{cases}$$

the sequence $(\widetilde{f}_n)_{n\in\mathbb{N}}$ defined by $\widetilde{f}_n := \widetilde{t}_n g_n$ for $n \in \mathbb{N}$ remains maximising for $G(\mu, \nu)$. We consider from now on this sequence, still called $(f_n)_{n\in\mathbb{N}}$. Then one has

$$\forall\, n \in \mathbb{N}, \qquad t_n(g_n(0) + b_2)^2 \geqslant \eta,$$

that is to say $f_n^2(0) + 2b_2\sqrt{t_n}f_n(0) + b_2^2 t_n \geqslant \eta$, which prevents the convergence $\lim_{n\to\infty} f_n(0) = 0$.

One can now perform a second-order expansion with remainder for $\mathrm{E}_\varphi(f_n^2)$; there exists a new function θ_n valued in $[f_n(0) - \sqrt{t_n}b_2, f_n(0) + \sqrt{t_n}b_2]$ and such that

$$\mathrm{E}_\varphi(f_n^2) = \mu\left[\varphi''(\theta_n)(f_n^2 - \mu[f_n^2])^2\right]/2.$$

First consider the case that $l := \lim_{n\to\infty} f_n(0)$ is finite. Since $l > 0$, one has uniformly on E

$$\lim_{n\to\infty} \varphi''(\theta_n)(f_n + \sqrt{\mu[f_n^2]})^2/2 = 2l^2\varphi''(l^2).$$

If $l^2\varphi''(l^2) > 0$, one draws therefrom the equivalent for large n

$$\mathrm{E}_\varphi(f_n^2) \sim 2l^2\varphi''(l^2)\mu[(f_n - \sqrt{\mu[f_n^2]})^2] \leqslant 2l^2\varphi''(l^2)\,\mathrm{Var}(f_n, \mu),$$

wherefrom

$$\lim_{n\to\infty} \frac{\mathrm{E}_\varphi(f_n^2)}{\nu[(f_n')^2]} \leqslant 2l^2\varphi''(l^2) \limsup_{n\to\infty} \frac{\mathrm{Var}(f_n, \mu)}{\nu[(f_n')^2]} \leqslant 2l^2\varphi''(l^2)A(\mu, \nu).$$

Similarly, one gets

$$\lim_{n\to\infty} \frac{\mathrm{E}_\varphi(f_n^2)}{\nu[(f_n')^2]} = 0$$

when $l^2\varphi''(l^2) = 0$. So it appears that one always has

$$G(\mu,\nu) \leqslant \sup_{l>0} 2l^2\varphi''(l^2)A(\mu,\nu) = \sup_{l>\eta} 2l^2\varphi''(l^2)A(\mu,\nu),$$

where the latter equality comes from the map $x \mapsto x\varphi''(x)$ being increasing on $]0,\eta]$. Conversely the inequality $G(\mu,\nu) \geqslant \sup_{l>\eta} 2l^2\varphi''(l^2)A(\mu,\nu)$ is satisfied under all circumstances: for all l larger than some given η, in the supremum defining $G(\mu,\nu)$, it suffices to consider functions of the form $l+\epsilon f$, with $f \in \mathcal{C}$ and $\epsilon > 0$ which is made to tend to 0. The above argument also holds if $\lim_{n\to\infty} f_n(0) = +\infty$, by existence and finiteness of $L = \lim_{x\to+\infty} x\varphi''(x)$. Thus, in all cases, the convergence entails the equality $G(\mu,\nu) = \sup_{l>\eta^2} 2l\varphi''(l)A(\mu,\nu)$. Then, one also has

$$\sup_{f\in\mathcal{D}} \frac{\mathrm{E}_\varphi(f^2)}{\nu[(f')^2]} = \Big(\sup_{l>\eta^2} 2l\varphi''(l)\Big) \sup_{f\in\mathcal{D}} \frac{\mathrm{Var}(f,\mu)}{\nu[(f')^2]} = \Big(\sup_{l>\eta^2} 2l\varphi''(l)\Big) A(\mu,\nu),$$

the claimed identity (2) follows.

• If $\lim_{n\to\infty} t_n \in\ ]0, R/b_1]$, one is back in a more classical framework, and, as already happened several times, two sub-cases will be considered.

- If $\lim_{n\to\infty} f_n(0) = +\infty$, the boundedness in $n \in \mathbb{N}$ of the oscillations of the f_n and the convergence $\lim_{t\to+\infty} x\varphi''(x) = L$ allow again to perform a second-order expansion with remainder, yielding for large n the equivalent

$$\mathrm{E}_\varphi(f_n^2) \sim \frac{L}{2}\,\mathrm{Var}(f_n,\mu)$$

if $L > 0$. On the other hand, if $L = 0$, it appears that

$$\mathrm{E}_\varphi(f_n^2) \ll \mathrm{Var}(f_n,\mu).$$

Since $A(\mu,\nu) < +\infty$, the latter possibility implies that one is in the trivial situation that $G(\mu,\nu) = 0$. If $L > 0$, one also obtains $G(\mu,\nu) = LA(\mu,\nu)/2$. So one is reduced to the case of the Poincaré inequality.

- If $\lim_{n\to\infty} f_n(0)$ exists in $\mathbb{R}$, one easily shows existence of some minimising function. But the proof of Lemma 4 immediately adapts to this situation, in view of the form of the modified entropy E_φ. Then the quickest way to conclude that (2) holds is to adapt the second proof of section 4.

References

1. Cécile Ané, Sébastien Blachère, Djalil Chafaï, Pierre Fougères, Ivan Gentil, Florent Malrieu, Cyril Roberto and Grégory Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

2. F. Barthe and C. Roberto. Sobolev inequalities for probability measures on the real line. *Studia Math.*, 159(3):481–497, 2003. Dedicated to Professor Aleksander Pełczyński on occasion of his 70th birthday (Polish).
3. S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.*, 163(1):1–28, 1999.
4. Eric A. Carlen. Superadditivity of Fisher's information and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 101(1):194–211, 1991.
5. Djalil Chafaï. Entropies, convexity, and functional inequalities. *J. Math. Kyoto Univ.*, 44(2):325–363, 2004.
6. Mu-Fa Chen and Feng-Yu Wang. Estimation of spectral gap for elliptic operators. *Trans. Amer. Math. Soc.*, 349(3):1239–1267, 1997.
7. Mufa Chen. Analytic proof of dual variational formula for the first eigenvalue in dimension one. *Sci. China Ser. A*, 42(8):805–815, 1999.
8. Mufa Chen. Variational formulas and approximation theorems for the first eigenvalue in dimension one. *Sci. China Ser. A*, 44(4):409–418, 2001.
9. Ivan Gentil, Arnaud Guillin and Laurent Miclo. Modified logarithmic Sobolev inequalities and transportation inequalities. Probab. Theory Related Fields 133 (2005), no. 3, 409–436.
10. Leonard Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.
11. Laurent Miclo. Quand est-ce que des bornes de Hardy permettent de calculer une constante de Poincaré exacte sur la droite? *Annales de la Faculté des Sciences de Toulouse*, Sér. 6, no. 17(1):121–192, 2008.
12. Laurent Miclo. On eigenfunctions of Markov processes on trees. *Probab. Theory Related Fields*, 142(3-4):561–594, 2008.
13. O. S. Rothaus. Logarithmic Sobolev inequalities and the spectrum of Sturm-Liouville operators. *J. Funct. Anal.*, 39(1):42–56, 1980.
14. O. S. Rothaus. Diffusion on compact Riemannian manifolds and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 42(1):102–109, 1981.
15. O. S. Rothaus. Logarithmic Sobolev inequalities and the spectrum of Schrödinger operators. *J. Funct. Anal.*, 42(1):110–120, 1981.
16. Walter Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.
17. Laurent Saloff-Coste. Lectures on finite Markov chains. In *Lectures on probability theory and statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.
18. Liming Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3):427–438, 2000.