The Herbst argument

In the seminal 1975 contribution [5], L. Gross emphasized a logarithmic form of the classical Sobolev inequality for Gaussian measures.

Let $\gamma = \gamma_n$ be the standard Gaussian probability distribution on $\mathbb{R}^n$, with density $\frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}|x|^2}$, $x \in \mathbb{R}^n$, with respect to the Lebesgue measure. The logarithmic Sobolev inequality (for γ) expresses that, for every smooth (locally Lipschitz) function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\int_{\mathbb{R}^n} f^2 d\gamma < \infty$,

$$\int_{\mathbb{R}^n} f^2 \log f^2 \, d\gamma - \int_{\mathbb{R}^n} f^2 \, d\gamma \, \log\left(\int_{\mathbb{R}^n} f^2 \, d\gamma \right) \leq 2 \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma. \tag{1}$$

L. Gross

By elementary convexity, the inequality expresses at a qualitative level that whenever $\int_{\mathbb{R}^n} |\nabla f|^2 d\gamma < \infty$, then f belongs to the Orlicz space $\mathrm{L}^2 \log \mathrm{L}(\gamma)$, thus part of the Sobolev inequalities. After a simple, equivalent change of functions, the Gaussian logarithmic Sobolev inequality is expressed equivalently with respect to the Lebesgue measure λ_n as

$$\begin{aligned}\int_{\mathbb{R}^n} f^2 \log f^2 d\lambda_n - \int_{\mathbb{R}^n} f^2\, d\lambda_n \,\log\bigg(\int_{\mathbb{R}^n} f^2\, d\lambda_n\bigg) \\ \leq \frac{n}{2}\int_{\mathbb{R}^n} f^2\, d\lambda_n \,\log\bigg(\frac{2}{n\pi e}\frac{\int_{\mathbb{R}^n}|\nabla f|^2 d\lambda_n}{\int_{\mathbb{R}^n} f^2\, d\lambda_n}\bigg)\end{aligned} \tag{2}$$

for every smooth function $f : \mathbb{R}^n \to \mathbb{R}$ such that $\int_{\mathbb{R}^n} f^2 d\lambda_n < \infty$.

It is however an important feature that the inequality (1) with respect to $\gamma = \gamma_n$ and the constants therein do not depend on the dimension of the underlying state space. By the specific sub-additivity property of the entropy functional $\int_{\mathbb{R}^n} f^2 \log f^2 d\gamma$, the inequality actually tensorizes and reduces to the one-dimensional case. By affine transformations, the logarithmic Sobolev inequality (1) may be formulated for arbitrary Gaussian measures. Due to its dimension-free character, infinite dimensional Gaussian measures may also be considered. It is a simple matter to check that the inequality is sharp on the exponential functions $f(x) = e^{\langle a,x\rangle - |a|^2}$, $x \in \mathbb{R}^n$, where $a \in \mathbb{R}^n$.

As already investigated by L. Gross, the logarithmic Sobolev inequality may also be considered for arbitrary probability measures μ on the Borel sets of $\mathbb{R}^n$, asking whether there exists a (finite) constant $C > 0$ such that, for every for every smooth function $f : \mathbb{R}^n \to \mathbb{R}$ with $\int_{\mathbb{R}^n} f^2 d\mu < \infty$,

$$\int_{\mathbb{R}^n} f^2 \log f^2\, d\mu - \int_{\mathbb{R}^n} f^2\, d\mu \,\log\bigg(\int_{\mathbb{R}^n} f^2\, d\mu\bigg) \,\leq\, 2C\int_{\mathbb{R}^n}|\nabla f|^2 d\mu. \tag{3}$$

More general contexts may be considered as well, and since their discovery and promotion by L. Gross in the seventies, logarithmic Sobolev inequalities have become a central object of interest and study, in various contexts, information theory, Dirichlet spaces, finite and infinite dimensional diffusion operators, statistical mechanics, optimal transport and partial differential equations, convex geometry, Markov chains etc., where they prove as a most useful tool and property, with numerous applications and illustrations. An account on logarithmic Sobolev inequalities and their applications is the monograph [2].

For a given measure μ, a natural issue is to try to determine conditions under which it satisfies such a logarithmic Sobolev inequality (3), or to express necessary conditions for the inequality to hold true. To the latter question raised by L. Gross, I. Herbst showed, in an unpublished letter to L. Gross attached at the end of this note (courtesy of Professor Gross), that if a probability measure μ on the real line satisfies (3), then it must be strongly integrable as Gaussians, namely there exists $\alpha > 0$ (depending on C) such that

$$\int_{\mathbb{R}} e^{\alpha x^2} d\mu \,<\, \infty. \tag{4}$$

The argument outlined by I. Herbst in his letter amounts to apply the logarithmic Sobolev inequality to a (truncated) exponential $e^{\lambda x^2}$ to deduce a differential inequality on $\int_{\mathbb{R}} e^{\lambda x^2} d\mu$ in $\lambda > 0$ which may be suitably integrated to yield the result.

I. Herbst

1 The Herbst argument

The argument sketched by I. Herbst was taken up again following the same steps by E. Davies and B. Simon [4] in the context of Schrödinger operators and Dirichlet Laplacians. In the landmark contribution [1], S. Aida, T. Masuda, I. Shigekawa broaden the investigation to Lipschitz functions on Dirichlet spaces, and reached in particular the following neat statement (which fully covers (4)).

Assume that μ satisfies the logarithmic Sobolev inequality (3). Whenever $F : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz with Lipschitz coefficient $\|F\|_{\mathrm{Lip}}$, then $\int_{\mathbb{R}^n} F^2 d\mu < \infty$ and for every $\alpha < \frac{1}{2C\|F\|_{\mathrm{Lip}}^2}$,

$$\int_{\mathbb{R}^n} e^{\alpha F^2} d\mu \leq \exp\left(\frac{\alpha}{1 - 2\alpha C\|F\|_{\mathrm{Lip}}^2} \int_{\mathbb{R}^n} F^2 d\mu\right). \tag{5}$$

The condition $\alpha < \frac{1}{2C\|F\|_{\mathrm{Lip}}^2}$ is sharp on the example of $F(x) = x$ on the real line with respect to the standard Gaussian measure γ.

The proof, following the early intuition by I. Herbst, amounts to apply the logarithmic Sobolev inequality to $e^{\lambda F^2}$ to derive a differential inequality on $\int_{\mathbb{R}^n} e^{\lambda F^2} d\mu$ as a function of λ. At a technical level, the proof should take care of the prior existence of the various integrals after a suitable smooth truncation of F, carefully discussed in [1] (basically, work first with a smooth regularization of $\min(\max(F,-k),k)$, $k \geq 1$). Set then $u(\lambda) = \int_{\mathbb{R}^n} e^{\lambda F^2} d\mu$, $\lambda \in \mathbb{R}$, and apply the logarithmic Sobolev inequality (3) to $f^2 = e^{\lambda F^2}$ so to get that

$$\begin{aligned} \lambda \int_{\mathbb{R}^n} F^2 \, e^{\lambda F^2} d\mu - u(\lambda) \log u(\lambda) \; &\leq \; 2C\lambda^2 \int_{\mathbb{R}^n} |\nabla F|^2 F^2 e^{\lambda F} d\mu \\ &\leq \; 2C\lambda^2 \|F\|_{\mathrm{Lip}}^2 \int_{\mathbb{R}^n} F^2 \, e^{\lambda F^2} d\mu \end{aligned}$$

since $|\nabla F| \leq \|F\|_{\mathrm{Lip}}$ (whenever F is smooth). Now $\int_{\mathbb{R}^n} F^2 \, e^{\lambda F} d\mu = u'(\lambda)$, so that the latter yields the first order differential inequality

$$\lambda \big(1 - 2C\lambda \|F\|_{\mathrm{Lip}}^2\big) u'(\lambda) \; \leq \; u(\lambda) \log u(\lambda), \quad \lambda \in \mathbb{R},$$

which, after some care, may be integrated to produce the announced claim (5).

2 The Herbst argument on Laplace transforms

It turns out that applying the preceding strategy to rather $e^{\lambda F}$ yields a more simple calculus on the Laplace transform $\int_{\mathbb{R}^n} e^{\lambda F} d\mu$, $\lambda \in \mathbb{R}$, of F, with derivations of sharp bounds [6]. Let indeed $v(\lambda) = \int_{\mathbb{R}^n} e^{\lambda F} d\mu$, $\lambda \in \mathbb{R}$ (considering first, as before, a smooth truncation of F in order for all the integrals to be well-defined). Applying the logarithmic Sobolev inequality (3) to $f^2 = e^{\lambda F}$, it holds true that

$$\lambda \int_{\mathbb{R}^n} F \, e^{\lambda F} d\mu - v(\lambda) \log v(\lambda) \; \leq \; \frac{C\lambda^2}{2} \int_{\mathbb{R}^n} |\nabla F|^2 e^{\lambda F} d\mu \; \leq \; \frac{C\lambda^2}{2} \, \|F\|_{\mathrm{Lip}}^2 \, v(\lambda).$$

Now $\int_{\mathbb{R}^n} F \, e^{\lambda F} d\mu = v'(\lambda)$ resulting into the simple first order differential inequality

$$\lambda \, v'(\lambda) \; \leq \; v(\lambda) \log v(\lambda) + \frac{C\lambda^2}{2} \, \|F\|_{\mathrm{Lip}}^2 \, v(\lambda)$$

in $\lambda \in \mathbb{R}$. This inequality is easily integrated. Indeed, setting for example $w(\lambda) = \frac{1}{\lambda} \log v(\lambda)$, $\lambda \in \mathbb{R}$, $w(0) = \int_{\mathbb{R}^n} F d\mu$, the preceding simply expresses that

$$w'(\lambda) \; \leq \; \frac{C}{2} \, \|F\|_{\mathrm{Lip}}^2, \quad \lambda \in \mathbb{R}.$$

Hence $\log v(\lambda) \leq \lambda w(0) + \frac{C}{2} \|F\|_{\mathrm{Lip}}^2 \lambda^2$, that is

$$v(\lambda) \; = \; \int_{\mathbb{R}^n} e^{\lambda F} d\mu \; \leq \; e^{\lambda \int_{\mathbb{R}^n} F d\mu + \frac{C}{2} \|F\|_{\mathrm{Lip}}^2 \lambda^2}, \quad \lambda \in \mathbb{R} \tag{6}$$

(with the prior observation that F is integrable). This conclusion is fully optimal as the function $F(x) = x$ on the real line achieves equality under the standard Gaussian measure γ_1.

The Laplace bounds (6) may be compared to the conclusion (5) from [1]. To this task, assume (for simplicity) that $\int_{\mathbb{R}^n} F d\mu = 0$, and write

$$e^{\alpha F^2} = \int_{\mathbb{R}} e^{t\sqrt{2\alpha}F} d\gamma_1(t)$$

so that, by Fubini's theorem and the application of (6),

$$\begin{aligned} \int_{\mathbb{R}^n} e^{\alpha F^2} d\mu &= \int_{\mathbb{R}} \bigg(\int_{\mathbb{R}^n} e^{t\sqrt{2\alpha}F} d\mu \bigg) d\gamma_1(t) \\ &\leq \int_{\mathbb{R}} e^{C\alpha \|F\|_{\mathrm{Lip}}^2 t^2} d\gamma(t) \\ &= \frac{1}{\sqrt{1 - 2C\alpha \|F\|_{\mathrm{Lip}}^2}} \end{aligned} \tag{7}$$

provided that $\alpha < \frac{1}{2C\|F\|_{\mathrm{Lip}}^2}$, a condition which coincides with the one in (5). For a more precise comparison with the latter, by the Poincaré inequality (consequence itself of the logarithmic Sobolev inequality, cf. [2]), if $\int_{\mathbb{R}^n} F d\mu = 0$,

$$\int_{\mathbb{R}^n} F^2 d\mu \leq C \int_{\mathbb{R}^n} |\nabla F|^2 d\mu \leq C\|F\|_{\mathrm{Lip}}^2.$$

Therefore (5) yields

$$\int_{\mathbb{R}^n} e^{\alpha F^2} d\mu \leq \exp\bigg(\frac{\alpha C \|F\|_{\mathrm{Lip}}^2}{1 - 2\alpha C \|F\|_{\mathrm{Lip}}^2} \bigg),$$

a weaker bound than (7).

3 Logarithmic Sobolev inequalities and concentration of measure

The Laplace bounds (6), achieved under the logarithmic Sobolev inequality (3), are a prototypical illustration of (Gaussian) concentration inequalities. Namely, for a given Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}$, under the logarithmic Sobolev inequality (3), F is integrable with respect to μ and (6) reads

$$\int_{\mathbb{R}^n} e^{\lambda(F - \int_{\mathbb{R}^n} F d\mu)} d\mu \leq e^{\frac{C}{2}\|F\|_{\mathrm{Lip}}^2 \lambda^2}, \quad \lambda \in \mathbb{R}.$$

Hence, by Markov's inequality, for every $\lambda, t > 0$,

$$\mu\big(F - \textstyle\int_{\mathbb{R}^n} F d\mu \geq t\big) \,\leq\, e^{-\lambda t + \frac{C}{2}\|F\|_{\mathrm{Lip}}^2 \lambda^2}$$

which, after optimization in λ ($\lambda = \frac{t}{C\|F\|_{\mathrm{Lip}}^2}$) yields

$$\mu\big(F - \textstyle\int_{\mathbb{R}^n} F d\mu \geq t\big) \,\leq\, e^{-\frac{t^2}{2C\|F\|_{\mathrm{Lip}}^2}} \tag{8}$$

for every $t \geq 0$.

Together with the same inequality for $-F$ and the union bound, the deviation inequality (8) yields the concentration inequality[1]

$$\mu\Big(\big|F - \textstyle\int_{\mathbb{R}^n} F d\mu\big| \geq t\Big) \,\leq\, 2\, e^{-\frac{t^2}{2C\|F\|_{\mathrm{Lip}}^2}}, \quad t \geq 0. \tag{9}$$

Again, these exponential bounds are sharp on the Gaussian model.

These deviation and concentration inequalities derived from a logarithmic Sobolev inequality are the illustration of a general principle, labelled as "the Herbst argument" or the "entropy method", from which numerous concentration inequalities may be achieved, discovered, or re-proved with elementary and more direct arguments, from logarithmic Sobolev inequalities or sub-additivity of entropy. An important feature of the approach is the production of dimension free concentration inequalities, a most attractive property in the study of high-dimensional systems and models. A (non-exhaustive) selection of applications and illustrations of these methods is presented in the course and monographs [7, 8, 3] (and the references therein), and more is still going on.

References

[1] S. Aida, T. Masuda, I. Shigekawa. Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.* 126, 83–101 (1994).

[2] D. Bakry, I. Gentil, M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der mathematischen Wissenschaften 348. Springer (2014).

[3] S. Boucheron, G. Lugosi, P. Massart. *Concentration inequalities.* Oxford University Press (2013).

[1] Concentration inequalities may also be deduced from the bound (5) at the expense of extra and non-optimal factors.

[4] E. B. Davies, B. Simon. Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. *J. Funct. Anal.* 59, 335–395 (1984).

[5] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* 97, 1061–1083 (1975)

[6] M. Ledoux. Remarks on logarithmic Sobolev constants, exponential integrability and bounds on the diameter. *J. Math. Kyoto Univ.* 35, 211–220 (1995).

[7] M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. *Séminaire de Probabilités XXXIII, Lecture Notes in Math.* 1709, 120–216. Springer (1999)).

[8] M. Ledoux. *The concentration of measure phenomenon.* Mathematical Surveys and Monographs 89. American Mathematical Society (2001).