ELSEVIER

Statistics & Probability Letters 44 (1999) 281–290

STATISTICS & PROBABILITY LETTERS

www.elsevier.nl/locate/stapro

A converse Gaussian Poincaré-type inequality for convex functions

S.G. Bobkov[a,1], C. Houdré[b,*,2]

[a] Department of Mathematics, Syktyvkar University, 167001 Syktyvkar, Russia
[b] Southeast Applied Analysis Center, School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30322, USA

Received August 1998; received in revised form November 1998

Abstract

A converse Poincaré-type inequality is obtained within the class of smooth convex functions for the Gaussian distribution.

MSC: primary 60E15; 60E99; secondary 26D10

Keywords: Poincaré-type inequalities; log-Sobolev inequalities; Gaussian distribution

These notes form a companion to a previous paper of the authors where the question of obtaining converse Poincaré-type inequalities was investigated. By a Poincaré-type inequality for a given probability measure μ on $\mathbb{R}$, one usually means an analytic inequality of the form

$$\mathrm{Var}_\mu(f) \leqslant K \|f'\|_2^2, \tag{1}$$

which relates the μ-variance of an arbitrary smooth function f on $\mathbb{R}$ to the $L^2(\mu)$-norm of its derivative f', and where the constant K does not depend on f. For example, as first noted in Klaassen (1985), the above inequality is satisfied by the double exponential distribution μ with the optimal constant $K = 4$. In fact, as already mentioned in Borovkov and Utev (1983), if ξ is a random variable with d.f. F and density p such that, for some x_0, $1 - F(x) \leqslant c\,p(x)$, if $x \geqslant x_0$, and $F(x) \leqslant c\,p(x)$, if $x \leqslant x_0$, then Eq. (1) holds with $K = 4c^2$.

For the double exponential measure, Eq. (1) can be inverted in the class of all convex f, in the sense that for some $K > 0$,

$$\mathrm{Var}_\mu(f) \geqslant K \|f'\|_2^2,$$

* Corresponding author. Tel.: +1-404-894-4398; fax: +1-404-894-4409.
E-mail address: houdre@math.gatech.edu (C. Houdré)

[1] Research supported in part by the Russian Foundation for Fundamental Research. This author greatly enjoyed the hospitality of the School of Mathematics, Georgia Institute of Technology, Atlanta, where part of this work was carried out.

[2] Research supported in part by the NSF Grant No. DMS-98032039. This author greatly enjoyed the hospitality of CIMAT, Gto, Mexico where part of this work was carried out.

0167-7152/99/$ - see front matter
PII: S0167-7152(99)00019-X

(see Bobkov and Houdré, 1997a). Actually, such a converse "convex" Poincaré-type inequality holds for many other symmetric probability measures μ with sufficiently heavy tails. It is however not satisfied by the canonical Gaussian measure $\mu=\gamma$, and at the same time, the inequality (1) for $\mu=\gamma$ holds true with $K=1$. Thus, the classical Gaussian Poincaré-type inequality is relatively rough, and for a suitable "invertible" estimate of $\mathrm{Var}_\gamma(f)$, we have to look for a quantity somewhat different from $\|f'\|_2^2$. As it turns out, the L^2-norm of the derivative can be replaced by a suitable Orlicz norm:

Theorem 1. *There exist positive numerical constants K_0 and K_1 such that, for every convex function $f\colon \mathbb{R}\to\mathbb{R}$,*

$$K_0\|f'\|_\psi^2 \leqslant \mathrm{Var}_\gamma(f) \leqslant K_1\|f'\|_\psi^2, \tag{2}$$

where $\|\cdot\|_\psi$ is the Orlicz norm in $L^\psi(\gamma)$ with $\psi(x)=x^2/[\log(1+x)]$, $x\geqslant 0$.

Recall that the Orlicz norm of a measurable function u in $L^\psi(\gamma)$ is defined by

$$\|u\|_\psi = \inf\left\{\lambda>0\colon E\psi\left(\frac{|u|}{\lambda}\right)\leqslant 1\right\},$$

where the above expectation is with respect to γ. One easily verifies that ψ above is a Young function, i.e., it is positive, convex, increasing in $x\geqslant 0$, with $\psi(0)=0$, so that $\|\cdot\|_\psi$ is indeed a norm.

The inequality on the right in (2) is not really new. It holds true for all absolutely continuous f, without the convexity assumption, and is a consequence of an inequality of Talagrand (1994), Theorem 1.6 on the discrete cube: for every function $f:\{0,1\}^n\to\mathbb{R}$,

$$\mathrm{Var}_{\mu_p^n}(f)\leqslant K(p)\sum_{i=1}^n \|\Delta_i f\|_\psi^2. \tag{3}$$

Here, the variance and the Orlicz norm are understood with respect to the n-fold tensor product μ_p^n of the Bernoulli measure on $\{0,1\}$ with parameter $p\in(0,1)$, and $\Delta_i f$ denotes the increment of f along the ith coordinate. Applying the above discrete inequality (with $p=1/2$, for definiteness) to the functions $f_n(x)=f((2(x_1+\cdots+x_n)-n)/\sqrt{n})$, $x\in\{0,1\}^n$, with f smooth enough, and letting $n\to\infty$, we obtain by the central limit theorem the right inequality in (2) with $K_1=4K(1/2)$.

In addition to the proof of Eq. (3) given in Talagrand (1994), it is also noted there that this inequality could also be obtained, by duality, from Gross (1976) logarithmic Sobolev inequality on the discrete cube (in the case $p=1/2$). For an arbitrary probability metric space it is shown, in the appendix at the end of this paper, that the right inequality in (2) implies a log-Sobolev inequality. However, we do not know if the converse statement is true in general, i.e., we do not know if a log-Sobolev inequality implies the right inequality in (2). In this connection, it might be worthwhile to note that, in contrast to (2), the Gaussian log-Sobolev inequality cannot be inverted in the class of all convex functions. Indeed, if we apply

$$E_\gamma f^2\log f^2 - E_\gamma f^2\log E_\gamma f^2 \geqslant K E_\gamma f'^2,$$

to the functions $e^{\varepsilon f}$, $\varepsilon>0$, with f convex and then let $\varepsilon\to 0$, we arrive at $2\mathrm{Var}_\gamma(f)\geqslant KE_\gamma f'^2$, which is clearly false.

We can now pass to the proof of the left inequality in (2). To denote the variance, we also write $\mathrm{Var}(f,\mu)$ instead of $\mathrm{Var}_\mu(f)$, and $\mathrm{Var}(\mu)$ instead of $\mathrm{Var}(f,\mu)$ when the function is identity: $f(x)=x$. We also write $E_\mu f$ to denote the expectation of f with respect to μ. As usual,

$$\Phi(x)=\gamma((-\infty,x])=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^x \mathrm{e}^{-t^2/2}\,\mathrm{d}t,\quad x\in\mathbb{R},$$

denotes the distribution function of γ, and $\varphi(x)=(1/\sqrt{2\pi})\mathrm{e}^{-x^2/2}$ denotes its density.

It is much easier to prove the left inequality in (2) assuming additionally that f is monotone. To treat the general case, we will have to consider an inequality such as

$$\mathrm{Var}(f,\gamma) \geqslant \int_{\mathbb{R}} f'^2 \,\mathrm{d}Q \tag{4}$$

in the class of all convex functions f on $\mathbb{R}$ for a suitable finite positive symmetric measure Q on $\mathbb{R}$, absolutely continuous with respect to γ. Before starting, let us note one sufficient condition for Eq. (4). First, we can assume that for some $a \in \mathbb{R}$, f is non-increasing on $(-\infty, a]$ and is non-decreasing on $[a, +\infty)$. Then, writing

$$\gamma = \Phi(a)\gamma_a^- + (1-\Phi(a))\gamma_a^+,$$

where γ_a^- and γ_a^+ are respectively the left and the right conditional restriction of γ to the half-lines $(-\infty, a]$ and $[a, +\infty)$, i.e.,

$$\gamma_a^-(A) = \frac{\gamma(A \cap (-\infty, a])}{\Phi(a)}, \qquad \gamma_a^+(A) = \frac{\gamma(A \cap [a, +\infty))}{1-\Phi(a)}, \quad A \subset \mathbb{R}, \text{ Borel},$$

we have

$$\begin{aligned}\mathrm{Var}(f,\gamma) &= \Phi(a)\mathrm{Var}(f,\gamma_a^-) + (1-\Phi(a))\mathrm{Var}(f,\gamma_a^+) + \Phi(a)(1-\Phi(a))(E_{\gamma_a^+} f - E_{\gamma_a^-} f)^2 \\ &\geqslant \Phi(a)\mathrm{Var}(f,\gamma_a^-) + (1-\Phi(a))\mathrm{Var}(f,\gamma_a^+).\end{aligned}$$

Therefore, Eq. (4) will follow from

$$\mathrm{Var}(f,\gamma_a^-) \geqslant \frac{1}{\Phi(a)} \int_{-\infty}^{a} f'^2 \,\mathrm{d}Q, \tag{5}$$

$$\mathrm{Var}(f,\gamma_a^+) \geqslant \frac{1}{1-\Phi(a)} \int_{a}^{+\infty} f'^2 \,\mathrm{d}Q. \tag{6}$$

Since Eqs. (5) and (6) are equivalent by the symmetry of γ (and of Q), and since f is non-decreasing on $[a, +\infty)$, in order to get Eq. (4), it will be sufficient to establish Eq. (6) in the class $\mathscr{F}_+$ of non-decreasing convex functions on $\mathbb{R}$. To do so, we use:

Lemma 1. *Let ξ be a random variable with finite second moment, and continuous distribution function, and let Q be a finite positive measure on $\mathbb{R}$. The following are equivalent*:
(a) $\mathrm{Cov}(f(\xi), g(\xi)) \geqslant \int_{\mathbb{R}} f'g' \,\mathrm{d}Q$, *for all* $f, g \in \mathscr{F}_+$, *such that* $Ef(\xi)^2$ *and* $Eg(\xi)^2$ *are finite.*
(b) $\mathrm{Var}(f(\xi)) \geqslant \int_{\mathbb{R}} f'^2 \,\mathrm{d}Q$, *for all* $f \in \mathscr{F}_+$.
(c) $\mathrm{Var}((\xi - a)^+) \geqslant Q([a, +\infty))$, *for all* $a \in \mathbb{R}$.

When Q is the law of ξ, the above statement is Lemma 1 in Bobkov and Houdré (1997a), but the proof there extends to arbitrary Q. Just note that the functionals

$$(f,g) \to \mathrm{Cov}(f(\xi), g(\xi)), \qquad (f,g) \to \int f'g',$$

are bilinear, and that the function $a \to \mathrm{Cov}((\xi - a)^+, (\xi - b)^+)$ is non-decreasing in $a \in (\infty, b]$ (note also the fact that any $f \in \mathscr{F}_+$ is a mixture of functions $f_a(x) = (x-a)^+$).

Now, by Lemma 1, the inequality (6) for the class $\mathscr{F}_+$ is equivalent to the same inequality with $f(x) = (x-b)^+, b \geqslant a$, that is, to

$$\mathrm{Var}((x-b)^+, \gamma_a^+) \geqslant \frac{1}{1-\Phi(a)} \int_b^\infty \mathrm{d}Q = \frac{Q([b, +\infty))}{1-\Phi(a)},$$

that is, to

$$\frac{\mathrm{Var}((x-b)^+,\gamma_a^+)}{\gamma_a^+([b,+\infty))} \geqslant \frac{Q([b,+\infty))}{1-\Phi(b)}. \tag{7}$$

Let us recall another elementary statement (which was observed in the proof of Theorem 1 in Bobkov and Houdré (1997a) and which is needed here only for $\mu=\gamma$).

Lemma 2. *Let μ be a probability measure on $\mathbb{R}$ such that $\mu([b,+\infty))>0$, for all $b\in\mathbb{R}$. Then, for all $a\leqslant b$,*

$$\frac{\mathrm{Var}((x-b)^+,\mu_a^+)}{\mu_a^+([b,+\infty))} \geqslant \mathrm{Var}((x-b)^+,\mu_b^+) = \mathrm{Var}(\mu_b^+).$$

In order words, the left-hand side of the above inequality is minimized for $a=b$. Using Lemma 2, the left-hand side of (7) can be estimated from below by $\mathrm{Var}(\gamma_b^+)$, so that Eq. (7) and therefore Eq. (4) will follow from

$$\mathrm{Var}(\gamma_b^+) \geqslant \frac{Q([b,+\infty))}{1-\Phi(b)}.$$

Thus, the measure Q in Eq. (4) can be chosen via the identity

$$Q([b,+\infty)) = (1-\Phi(b))\mathrm{Var}(\gamma_b^+). \tag{8}$$

Since

$$\begin{aligned}\mathrm{Var}(\gamma_b^+) &= \mathrm{Var}((x-b)^+,\gamma_b^+)\\ &= \frac{1}{1-\Phi(b)} E_\gamma (x-b)^{+2} - \left(\frac{1}{1-\Phi(b)} E_\gamma (x-b)^+\right)^2,\end{aligned}$$

the equality (8) is just

$$Q([b,+\infty)) = E_\gamma (x-b)^{+2} - \frac{1}{1-\Phi(b)} (E_\gamma (x-b)^+)^2. \tag{9}$$

This leads us to:

Lemma 3. *For all convex functions f on $\mathbb{R}$,*

$$\mathrm{Var}_\gamma(f) \geqslant K \int_{-\infty}^{+\infty} f'(x)^2 \frac{1}{1+x^2}\,\mathrm{d}\gamma(x).$$

Proof. As it follows from the above discussion, it only remains to show that, for all $b\in\mathbb{R}$,

$$Q([b,+\infty)) \geqslant K \int_b^{+\infty} \frac{1}{1+x^2}\,\mathrm{d}\gamma(x),$$

where K is a universal constant and where the function Q is defined via Eq. (8). For $b=-\infty$, the above inequality is trivial since, by Eq. (8), $Q((-\infty,+\infty))=\mathrm{Var}(\gamma)=1$; so only the behavior, as $b\to+\infty$, needs to be considered on both sides of the above inequality. That is we just need to compute the correct asymptotics the right-hand side of Eq. (9), as $b\to+\infty$. We start with the second term there. Integrating by parts, we have

$$E_\gamma (x-b)^+ = \int_b^\infty (x-b)\varphi(x)\,\mathrm{d}x = \varphi(b) - b(1-\Phi(b)).$$

Now,

$$1-\Phi(b)=\frac{\varphi(b)}{b}-\frac{\varphi(b)}{b^3}+3\int_b^\infty\frac{\varphi(x)}{x^4}\,\mathrm{d}x$$

$$\geqslant\frac{\varphi(b)}{b}-\frac{\varphi(b)}{b^3}+\mathrm{O}\left(\frac{\varphi(b)}{b^5}\right).$$

Hence,

$$E_\gamma(x-b)^+\leqslant\frac{\varphi(b)}{b^2}\left(1+\mathrm{O}\left(\frac{1}{b}\right)\right),$$

and thus

$$(E_\gamma(x-b)^+)^2\leqslant\frac{\varphi(b)^2}{b^4}\left(1+\mathrm{O}\left(\frac{1}{b}\right)\right).$$

But, since

$$1-\Phi(b)\geqslant\frac{\varphi(b)}{b}\left(1-\frac{1}{b^2}\right),$$

it follows that

$$\frac{1}{1-\Phi(b)}(E_\gamma(x-b)^+)^2\leqslant\frac{\varphi(b)}{b^3}\left(1+\mathrm{O}\left(\frac{1}{b}\right)\right).$$

Let us estimate the first term on the right-hand side of Eq. (9):

$$E_\gamma(x-b)^{+2}=\int_b^\infty(x-b)^2\,\mathrm{d}\Phi(x)$$

$$=2\int_b^\infty(x-b)(1-\Phi(x))\,\mathrm{d}x$$

$$\geqslant2\int_b^\infty(x-b)\varphi(x)\left(\frac{1}{x}-\frac{1}{x^3}+\mathrm{O}\left(\frac{1}{x^5}\right)\right)\,\mathrm{d}x$$

$$=2\int_b^{+\infty}\varphi(x)\left[1-\frac{1}{x^2}-\frac{b}{x}+\frac{b}{x^3}+\mathrm{O}\left(\frac{1}{x^4}\right)\right]\,\mathrm{d}x.$$

Denote by A this last integral and proceed to estimate every single one of its components.

$$\mathrm{I}:\quad\int_b^\infty\varphi(x)\,\mathrm{d}x=1-\Phi(b)=\varphi(b)\left[\frac{1}{b}-\frac{1}{b^3}+\mathrm{O}\left(\frac{1}{b^5}\right)\right].$$

$$\mathrm{II}:\quad\int_b^\infty\frac{\varphi(x)}{x^2}\,\mathrm{d}x=\int_b^\infty\frac{1}{x^3}\,\mathrm{d}(-\varphi(x))=\frac{\varphi(b)}{b^3}+\mathrm{O}\left(\frac{\varphi(b)}{b^5}\right).$$

$$\mathrm{III}:\quad\int_b^\infty\frac{\varphi(x)}{x}\,\mathrm{d}x=\int_b^\infty\frac{1}{x^2}\,\mathrm{d}(-\varphi(x))$$

$$=\frac{\varphi(b)}{b^2}-2\int_b^\infty\frac{\varphi(x)}{x^3}\,\mathrm{d}x$$

$$= \frac{\varphi(b)}{b^2} + 2\int \frac{1}{x^4}\,\mathrm{d}(-\varphi(x))$$

$$= \frac{\varphi(b)}{b^2} - 2\frac{\varphi(b)}{b^4} + \mathrm{O}\left(\frac{\varphi(b)}{b^6}\right).$$

Hence,

$$b\int_b^{\infty} \frac{\varphi(x)}{x}\,\mathrm{d}x = \frac{\varphi(b)}{b} - 2\frac{\varphi(b)}{b^3} + \mathrm{O}\left(\frac{1}{b^5}\right).$$

$$\mathrm{IV}: \quad \int_b^{\infty} \frac{\varphi(x)}{x^3}\,\mathrm{d}x = \int_b^{+\infty} \frac{1}{x^4}\,\mathrm{d}(-\varphi(x)) = \frac{\varphi(b)}{b^4} + \mathrm{O}\left(\frac{\varphi(b)}{b^6}\right),$$

thus,

$$b\int_b^{\infty} \frac{\varphi(x)}{x^3}\,\mathrm{d}x = \frac{\varphi(b)}{b^3} + \mathrm{O}\left(\frac{\varphi(b)}{b^5}\right).$$

Finally,

$$\mathrm{V}: \quad \int_b^{\infty} \frac{\varphi(x)}{x^4}\,\mathrm{d}x = \mathrm{O}\left(\frac{\varphi(b)}{b^5}\right).$$

Combining these five estimates, we have

$$A = \varphi(b)\left[\left(\frac{1}{b} - \frac{1}{b^3}\right) - \left(\frac{1}{b} - \frac{2}{b^3}\right) - \frac{1}{b^3} + \frac{1}{b^3} + \mathrm{O}\left(\frac{1}{b^5}\right)\right]$$

$$= \frac{\varphi(b)}{b^3}\left(1 + \mathrm{O}\left(\frac{1}{b^2}\right)\right).$$

Hence,

$$E_\gamma(x-b)^{+2} \geqslant 2\frac{\varphi(b)}{b^3}\left(1 + \mathrm{O}\left(\frac{1}{b^2}\right)\right),$$

$$Q([b,+\infty)) \geqslant 2\frac{\varphi(b)}{b^3} - \frac{\varphi(b)}{b^3} + \mathrm{O}\left(\frac{\varphi(b)}{b^4}\right)$$

$$\geqslant \frac{\varphi(b)}{b^3}\left(1 + \mathrm{O}\left(\frac{1}{b}\right)\right)$$

$$\geqslant K\frac{1-\Phi(b)}{b^2}\left(\text{or } \geqslant K\int_b^{+\infty} \frac{\varphi(x)}{x^2}\,\mathrm{d}x\right),$$

as $b \to +\infty$. This finishes the proof of Lemma 3. □

To continue, recall that given a Young function N, the corresponding Orlicz norm is defined by

$$\|u\|_N = \inf\left\{\lambda > 0\colon EN\left(\frac{|u|}{\lambda}\right) \leqslant 1\right\}. \tag{10}$$

Lemma 4. *For all measurable functions u, v on some probability space,*

$$|Euv| \leqslant 3\|u\|_{x\log(1+x)}\|v\|_{\mathrm{e}^x-1}.$$

Proof. One may assume $u, v \geqslant 0$. By homogeneity, let $\|v\|_{e^x-1} = 1$ so that by Eq. (10) $Ee^v = 2$. But

$$\sup_{Ee^v=2} Euv = Eu \log\left(\frac{2u}{Eu}\right),$$

which is just a well-known functional representation of the entropy. Thus to prove the result, it is enough to show that for all $u \geqslant 0$ such that $Eu > 0$,

$$Eu \log\left(\frac{2u}{Eu}\right) \leqslant 3\|u\|_{x\log(1+x)}.$$

Again, by homogeneity, let $\|u\|_{x\log(1+x)} = 1$, that is let

$$Eu \log(1+u) = 1.$$

By Jensen's inequality, $Eu \log(1+u) \geqslant Eu \log(1+Eu)$, so

$$Eu \log(1+Eu) \leqslant 1.$$

Hence $Eu \leqslant 2$, since $2\log 2 > 1$ and since the function $x\log(1+x)$ is increasing. Therefore, using also that fact that $x \log x \geqslant -1/e$ for all $x > 0$, we have

$$\begin{aligned} Eu \log\left(\frac{2u}{Eu}\right) &= Eu \log 2 + Eu \log u - Eu \log Eu \\ &\leqslant 2\log 2 + Eu \log(1+u) + e^{-1} \\ &= 2\log 2 + 1 + e^{-1} \leqslant 3, \end{aligned}$$

and Lemma 4 is proved. □

Lemma 5. *For every measurable function* $g : \mathbb{R} \to \mathbb{R}$,

$$\|g\|_\psi^2 \leqslant K \int_{-\infty}^{+\infty} g(x)^2 \frac{1}{1+x^2}\, d\gamma(x),$$

where K is universal constant, and $\|\cdot\|_\psi$ is the Orlicz norm in the space $L^\psi(\gamma)$ with respect to the Young function $\psi(x) = x^2/[\log(1+x)]$, $x \geqslant 0$.

Proof. By homogeneity, let $\|g\|_\psi = 1$; also we can assume that $g \geqslant 0$. Applying Lemma 4 to $u = \psi(g(x))/[1+x^2]$ and $v(x) = 1+x^2$, we have

$$\begin{aligned} 1 = \int_{-\infty}^{+\infty} \psi(g)\, d\gamma &= \int_{-\infty}^{+\infty} \frac{\psi(g)}{v} v\, d\gamma \\ &\leqslant 3 \left\| \frac{\psi(g)}{v} \right\|_{x\log(1+x)} \|v\|_{e^x-1} \\ &= C \left\| \frac{\psi(g)}{v} \right\|_{x\log(1+x)}, \end{aligned}$$

where $C = 3\|v\|_{e^x-1} < +\infty$ is universal. Thus,

$$\left\| \frac{\psi(g)}{v} \right\|_{x\log(1-x)} \geqslant \frac{1}{C} = c$$

which is equivalent to saying that

$$E_\gamma \frac{\psi(g)}{cv}\log\left(1+\frac{\psi(g)}{cv}\right) \geqslant 1.$$

Since $v \geqslant 1$, it follows that

$$E_\gamma \frac{\psi(g)}{v}\log\left(1+\frac{\psi(g)}{c}\right) \geqslant c.$$

From the very definition of ψ, we thus have

$$E_\gamma \frac{g^2}{v\log(1+g)}\log\left(1+\frac{g^2}{c(\log(1+g))}\right) \geqslant c. \tag{11}$$

Now, note for some $\alpha = \alpha(c) > 0$,

$$\log\left(1+\frac{x^2}{c\log(1+x)}\right) \leqslant \alpha\log(1+x), \quad \text{for all } x \geqslant 0. \tag{12}$$

Indeed, this inequality holds true near the points $x = 0$ and $x = +\infty$. Finally, using Eq. (12) in Eq. (11), we get $E_\gamma(g^2/v) \geqslant c/\alpha$, that is,

$$\int_{-\infty}^{+\infty} \frac{g(x)^2}{1+x^2}\,\mathrm{d}\gamma(x) \geqslant \frac{c}{\alpha}. \qquad \square$$

Proof of Theorem 1 (the left inequality (2))**.** Combine Lemmas 3 and 5 (with $g = f'$). $\square$

Appendix

Let (M, ρ) be a metric space equipped with a Borel probability measure μ. For every function f on M, one defines its "modulus of gradient"

$$|\nabla f(x)| = \limsup_{\rho(x,y)\to 0^+} \frac{|f(x)-f(y)|}{\rho(x,y)}, \quad x \in M,$$

with the convention that $|\nabla f(x)| = 0$, when x is an isolated point in M. We apply this definition to the class $\mathscr{A}$ of all functions f on M which have a finite Lipschitz constant on every ball in M: and in this case the function $|\nabla f|$ is finite and Borel measurable (cf. Bobkov and Houdré (1997b) for details). Now for ψ as in Theorem 1, consider inequalities of the following two types:

$$\mathrm{Var}(f) \leqslant K \|\, |\nabla f|\, \|_\psi^2, \tag{13}$$

$$Ef^2 \log f^2 - Ef^2 \log Ef^2 \leqslant K' E|\nabla f|^2, \tag{14}$$

where the expectations, the variance and the Orlicz norm are with respect to μ, where f is an arbitrary function in $\mathscr{A}$, and where the constants K and K' do not depend on f. Here we prove that:

Proposition 1. *The inequality* (13) *implies the inequality* (14) *with $K' \leqslant cK$, where c is a universal constant.*

It is probable that the converse statement in the above proposition is not true. The L^ψ-norm is weaker than the L^2-norm. So, comparing (13) with (14) is very similar to the analogous problem of comparing the inequality (a) $E|f - Ef| \leqslant KE|\nabla f|$, involving the L^1-norm of the modulus of the gradient, with a Poincaré-type inequality (b) $E|f - Ef|^2 \leqslant K'E|\nabla f|^2$, involving the L^2-norm of the modulus of gradient. It is well known that (a) implies (b) which is a version of a Cheeger's inequality. However, the converse is not true. Already on the real line $M = \mathbb{R}$ there exist probability measures which satisfy (b) but do not satisfy (a) with a finite K.

Lemma 6. $\|u\|_\psi \leqslant 2\|u\|_2$, *for all* $u \in L^\psi(\mu)$.

Proof. It is easy to verify that $\psi(y) \leqslant |y| + y^2$, for all $y \in \mathbb{R}$. Thus, if $\|u\|_2^2 \equiv Eu^2 = 1$, then $E\psi(u) \leqslant 2$. Hence, $E\psi(\frac{1}{2}u) \leqslant \frac{1}{2}E\psi(u) \leqslant 1$ (since ψ is convex and $\psi(0) = 0$). □

Lemma 7. *For all measurable functions u and v,* $\|uv\|_\psi \leqslant 24\|u\|_{e^{x^2}-1}\|v\|_2$.

Proof. We use Young's inequality:

$$xy \leqslant \int_0^x \alpha(t)\,dt + \int_0^y \beta(s)\,ds \equiv A(x) + B(y), \quad x, y \geqslant 0,$$

where β is an arbitrary continuous increasing function on $[0, +\infty)$ with $\beta(0) = 0$, $\beta(+\infty) = +\infty$, and where α is the inverse of β. Set $\beta(s) = \sqrt{2\log(1+s)}$ so that $B(y) \leqslant y\sqrt{2\log(1+y)}$. Since $\alpha(t) = e^{t^2/2} - 1$, we also have $A(x) \leqslant xe^{x^2/2}$. Thus,

$$xy \leqslant xe^{x^2/2} + y\sqrt{2\log(1+y)}, \quad x, y \geqslant 0. \tag{15}$$

Now, by the very definition of ψ, $\psi(\lambda a) \leqslant \lambda^2\psi(a)$, whenever $\lambda \geqslant 1$. Hence, using also the convexity of ψ, we have $\psi(a+b) \leqslant 2\psi(a) + 2\psi(b)$, for all $a, b \geqslant 0$. Applying this inequality to $a = xe^{x^2/2}$, $b = y\sqrt{2\log(1+y)}$, we get from Eq. (15):

$$\psi(xy) \leqslant 2\psi(xe^{x^2/2}) + 2\psi(y\sqrt{2\log(1+y)}), \quad x, y \geqslant 0. \tag{16}$$

Since $1 + xe^{x^2/2} \geqslant e^{x^2/2}$, we have $\psi(xe^{x^2/2}) = x^2e^{x^2}/[\log(1+xe^{x^2/2})] \leqslant 2e^{x^2}$. Setting $y = e^{x_1^2/2} - 1$, we also have

$$\psi(y\sqrt{2\log(1+y)}) = \psi(x_1(e^{x_1^2/2} - 1)) \leqslant \psi(x_1e^{x_1^2/2}) \leqslant 2e^{x_1^2} = 2(1+y)^2.$$

These two estimates together with (16) give $\psi(xy) \leqslant 4e^{x^2/2} + 4(1+y)^2$. Therefore, if $u, v \geqslant 0$ (this can be assumed) and $\|u\|_{e^{x^2}-1} = \|v\|_2 = 1$, that is, if $Ee^{u^2} = 2$ and $Ev^2 = 1$, then $E\psi(uv) \leqslant 4Ee^{u^2} + 4E(1+v)^2 \leqslant 24$. This implies the inequality of the lemma. □

Proof of Proposition 1. Introduce the functional

$$\mathscr{L}(f) = \sup_{c\in\mathbb{R}} [E(f+c)^2 \log(f+c)^2 - E(f+c)^2 \log E(f+c)^2],$$

which is defined for all measurable f, is non-negative, and is finite if and only if $Ef^2\log f^2 < +\infty$. Since the modulus of the gradient $|\nabla f|$ is invariant under translations $f \to f + c$, the inequality (14) can formally be strengthened and written as $\mathscr{L}(f) \leqslant K'E|\nabla f|^2$. On the other hand, as shown in Bobkov and Götze (1999), Proposition 4.1,

$$\tfrac{2}{3}\|f - Ef\|_N^2 \leqslant \mathscr{L}(f) \leqslant \tfrac{5}{2}\|f - Ef\|_N^2, \tag{17}$$

where the Orlicz norm is generated by the Young function $N(x) = x^2\log(1+x^2)$. Thus, up to a constant, Eq. (14) is equivalent to the inequality

$$\|f - Ef\|_N^2 \leqslant K''E|\nabla f|^2, \tag{18}$$

where N is as above. To deduce Eq. (18) from Eq. (13), we may assume that f is bounded, $Ef = 0$ and $E|\nabla f|^2 = 1/K$ (and that $K > 0$). Apply Eq. (13) to the function $g = f\sqrt{\log(1+f^2)}$. Since $|(x\sqrt{\log(1+x^2)})'| \leqslant 2\sqrt{\log(1+x^2)}$ and thus $|\nabla g| \leqslant 2\sqrt{\log(1+f^2)}\,|\nabla f|$, we get:

$$\begin{aligned} EN(f) &= E(f\sqrt{\log(1+f^2)})^2 \\ &\leqslant (Ef\sqrt{\log(1+f^2)})^2 + K\|2\sqrt{\log(1+f^2)}\,|\nabla f|\,\|_\psi^2. \end{aligned} \tag{19}$$

Once more by Eq. (13), applied to f, and by Lemma 6,

$$Ef^2 \leqslant K\| \, |\nabla f| \, \|_\psi^2 \leqslant 4K\| \, |\nabla f| \, \|_2^2 = 4. \tag{20}$$

Hence, by Cauchy–Schwarz and Jensen's inequalities,

$$(Ef\sqrt{\log(1+f^2)})^2 \leqslant Ef^2 E\log(1+f^2) \leqslant Ef^2 \log(1+Ef^2) \leqslant 4\log 5.$$

This gives a bound for the first term in Eq. (19). In order to estimate the second term, note that, for the Young function $\theta(x) = \mathrm{e}^{x^2} - 1$, we have according to Eq. (20) that $E\theta(\sqrt{\log(1+f^2)}) = Ef^2 \leqslant 4$, so that $\|\sqrt{\log(1+f^2)}\|_\theta \leqslant 4$. Hence, by Lemma 7,

$$\|\sqrt{\log(1+f^2)}\,|\nabla f|\,\|_\psi \leqslant 24\|\sqrt{\log(1+f^2)}\|_\theta \,\| \, |\nabla f| \, \|_2 \leqslant 96/\sqrt{K}.$$

Combining these bounds, we obtain from Eq. (19) that $EN(f) \leqslant 4\log 5 + (2.96)^2 < 200^2$. Hence, $\|f\|_N \leqslant 200$, since $N(tx) \leqslant t^2 N(x)$ for $|t| \leqslant 1$. This proves Eq. (18) with $K'' = 200^2 K$. By the second inequality in (17), we finally get $\mathscr{L}(f) \leqslant \frac{5}{2} 200^2 K$. Thus, Proposition 1 is proved with $c = \frac{5}{2} 200^2$. $\square$

References

Bobkov, S.G., Götze, F., 1999. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. Bielefeld University, Preprint SFB 97-076; J. Func. Anal., in press.

Bobkov, S.G., Houdré, C., 1997a. Converse Poincaré-type inequalities for convex functions. Statist. Probab. Lett. 34, 37–42.

Bobkov, S.G., Houdré, C., 1997b. Some connections between isoperimetric and Sobolev-type inequalities. Mem. Amer. Math. Soc. 129 (616).

Borovkov, A.A., Utev, S.A., 1983. On an inequality and a related characterization of the normal distribution. Theory Probab. Appl. 28, 209–218.

Gross, L., 1976. Logarithmic Sobolev inequalities. Amer. J. Math. 97, 1061–1083.

Klaassen, C.A.J., 1985. On an inequality of Chernoff. Ann. Probab. 13, 966–974.

Talagrand, M., 1994. On Russo's approximate zero-one law. Ann. Probab. 22 (3), 1576–1587.