Shift inequalities of Gaussian type and norms of barycentres

by

F. Barthe, D. Cordero-Erausquin and
M. Fradelizi (Marne-la-Vallée)

Abstract. We derive the equivalence of different forms of Gaussian type shift inequalities. This completes previous results by Bobkov. Our argument strongly relies on the Gaussian model for which we give a geometric approach in terms of norms of barycentres. Similar inequalities hold in the discrete setting; they improve the known results on the so-called isodiametral problem for the discrete cube. The study of norms of barycentres for subsets of convex bodies completes the exposition.

1. Introduction. Let μ be a measure on $\mathbb{R}$ and consider the product measure μ^∞ on $\mathbb{R}^\infty$. For $h \in \mathbb{R}^\infty$, the translate μ_h^∞ of μ^∞ is defined for Borel subsets $A \subset \mathbb{R}^\infty$ by

$$\mu_h^\infty(A) = \mu^\infty(A + h).$$

Feldman and Shepp proved that μ^∞ and μ_h^∞ are equivalent for every $h \in \ell^2$ if ([S]) and only if ([F]) μ is absolutely continuous with respect to the Lebesgue measure, with a.e. positive density ϱ and finite Fisher information

$$J(\mu) = \int (\varrho'/\varrho)^2 \, d\mu < \infty,$$

where ϱ' is the derivative of ϱ in the sense of distributions. More precisely, if one sets $h = \sqrt{\varrho}$, the condition on the Fisher information is equivalent to $h' \in L^2(dx)$. Under this hypothesis, it is natural to quantify the equivalence of μ^∞ and μ_h^∞. One can look for functions R_h and S_h such that for all A,

$$S_h(\mu^\infty(A)) \le \mu_h^\infty(A) \le R_h(\mu^\infty(A)).$$

This question is extensively studied by Bobkov in [Bo2]. For the canonical Gaussian measure γ, the optimal bounds are given by Kuelbs and Li ([Ku-Li]):

$$\Phi(\Phi^{-1}(\gamma^\infty(A)) - |h|) \le \gamma_h^\infty(A) \le \Phi(\Phi^{-1}(\gamma^\infty(A)) + |h|)$$

where equality occurs when A is a half-space orthogonal to h. In the previous inequality, $\Phi(t) = \int_{-\infty}^{t} \exp(-s^2/2)\,ds/\sqrt{2\pi}$ is the distribution function of γ, Φ^{-1} is its reciprocal and $|\cdot|$ is the Euclidean norm. Bobkov gives a necessary and sufficient condition for a measure μ to satisfy a Gaussian-like shift inequality: with the previous notation, there exists $c > 0$ such that for every Borel set $A \subset \mathbb{R}^\infty$, and for all $h \in \ell^2$,

$$(1) \qquad \Phi(\Phi^{-1}(\mu^\infty(A)) - c|h|) \le \mu_h^\infty(A) \le \Phi(\Phi^{-1}(\mu^\infty(A)) + c|h|)$$

if and only if μ has an a.e. positive density ϱ with

$$\int \exp(\varepsilon \varrho'^2/\varrho^2)\,d\mu < \infty$$

for some $\varepsilon > 0$. The first step of his proof is to show that, if we set $\mu_n = \mu^{\otimes n}$, the shift inequality (1) is equivalent to the following: for every $n \ge 1$ and every smooth, compactly supported function $f : \mathbb{R}^n \to [0,1]$,

$$(2) \qquad \left| \int \nabla f \, d\mu_n \right| \le cI\left(\int f \, d\mu_n \right),$$

where $I = \Phi' \circ \Phi^{-1}$ is the Gaussian isoperimetric function. In fact, this inequality is the infinitesimal form of (1).

The starting point of our work was the following improvement of the differential form (2) of the Gaussian shift inequality: for f as above

$$(3) \qquad \sqrt{\left(\int_{\mathbb{R}^n} I(f)\,d\gamma_n \right)^2 + \left| \int_{\mathbb{R}^n} \nabla f\,d\gamma_n \right|^2} \le I\left(\int_{\mathbb{R}^n} f\,d\gamma_n \right).$$

We learnt this inequality from Beckner during his lectures at the Institut Henri Poincaré in May 1998. His proof consisted in showing that the left hand side term is non-decreasing along the Ornstein–Uhlenbeck semigroup $(P_t)_{t \ge 0}$:

$$P_t(g)(x) = \int_{\mathbb{R}^n} g(e^{-t}x + (1 - e^{-2t})^{1/2}y)\,d\gamma_n(y).$$

Actually, (3) appeared earlier in the article of Bakry and Ledoux [BL] in the following form:

$$I(P_t f)^2 - P_t(I(f))^2 \ge (e^{2t} - 1)|\nabla P_t f|^2.$$

Indeed, for $F = f \circ (1 - e^{-2t})^{1/2}Id$, one has $P_t f(0) = \int F\,d\gamma_n$.

Inequality (3) is a formal inverse of Bobkov's form of Gaussian isoperimetry in [Bo1]:

$$I\left(\int_{\mathbb{R}^n} f\,d\gamma_n \right) \le \int_{\mathbb{R}^n} \sqrt{I^2(f) + |\nabla f|^2}\,d\gamma_n.$$

We will see that (3) has a similar formal behaviour. For this reason, the development will parallel the work about isoperimetry of Bobkov [Bo1], Bakry and Ledoux [BL] and of the first named author and Maurey [Ba-M].

The organization of the paper is as follows: in Section 2, we give a simplified proof of (3). Integration by parts transforms the problem into a very simple question about norms of barycentres. The third section is devoted to the study of probability measures μ on $\mathbb{R}^n$ which satisfy

$$(4) \qquad \sqrt{\left(\int_{\mathbb{R}^n} I(f) \, d\mu \right)^2 + \left| \frac{1}{c} \int_{\mathbb{R}^n} \nabla f \, d\mu \right|^2} \leq I\left(\int_{\mathbb{R}^n} f \, d\mu \right)$$

for some positive c and all functions f as above. It turns out that this is equivalent to the weaker form (2); thus (2) inherits the tensorization property of (4). Moreover, the stronger form (4) is adapted to semigroup methods, which allows us to prove it for a class of Boltzmann measures.

In Section 4, we study the corresponding inequalities in the discrete cube. They provide a simple way to study the average isodiametral problem of Ahlswede and Katona [AK]. We improve the results of Althöfer and Sillke [Al-Si] and recover the fact that among subsets of $\{-1, 1\}^n$ containing half of the points, half-cubes have largest "norm of barycentres".

In Section 5, we show that in the "continuous" cube $[-1, 1]^n$, half-cubes have the same extremal property.

2. The Gaussian model. Let $(\mathbb{R}^n; \langle \cdot, \cdot \rangle, | \cdot |)$ be the Euclidean n-space endowed with the canonical scalar product and norm. Let γ_1 be the measure on $\mathbb{R}$ with density $\varphi(t) = \exp(-t^2/2)/\sqrt{2\pi}$. From now on, $\gamma_n = \gamma_1^{\otimes n}$ will be the standard n-dimensional Gaussian probability. Recall that $\Phi(t) = \int_{-\infty}^t \varphi(u) \, du$ and $I = \varphi \circ \Phi^{-1}$.

Let μ be a measure on $\mathbb{R}^n$. For a measurable $A \subset \mathbb{R}^n$ we denote, whenever it exists, $g_\mu(A) = \int_A x \, d\mu(x)$. Then $g_\mu(A)/\mu(A)$ is the barycentre of A for μ.

The infinitesimal version of the shift inequality for γ_n (see [Bo2]) states that for smooth and compactly supported $f : \mathbb{R}^n \to [0, 1]$, one has

$$(5) \qquad \left| \int_{\mathbb{R}^n} \nabla f \, d\gamma_n \right| \leq I\left(\int_{\mathbb{R}^n} f \, d\gamma_n \right).$$

On integrating by parts, this is equivalent to

$$\left| \int_{\mathbb{R}^n} x f(x) \, d\gamma_n(x) \right| \leq I\left(\int_{\mathbb{R}^n} f \, d\gamma_n \right),$$

which can be extended to any measurable f with values in $[0, 1]$. On characteristic functions of sets, this inequality means that among the sets A of given Gaussian measure, half-spaces are the ones whose barycentres have maximal Euclidean norm. Indeed, if $|e| = 1$ and $A = \{x : \langle x, e \rangle \leq a\}$, then $\gamma_n(A) = \Phi(a)$ and

$$\int\limits_A x\,d\gamma_n(x) = \Big(\int\limits_{-\infty}^{a} t\,d\gamma_1(t)\Big)e = -\varphi(a)\,e = -I(\gamma_n(A))e.$$

But such a property of half-spaces is very general and simple:

LEMMA 1. *Let μ be a probability measure on $\mathbb{R}^n$ such that $\mu(h) = 0$ for every hyperplane h. Let Ψ be a convex function on $\mathbb{R}^n$ and assume that $\int_{\mathbb{R}^n} |x|\,d\mu(x) < \infty$. Let $0 < a < 1$. Then*

(6) $$\sup_{\mu(A)=a} \Psi(g_\mu(A)) = \sup\{\Psi(g_\mu(H)) : H \text{ half-space such that } \mu(H) = a\}.$$

When $\{x : \Psi(x) = \inf \Psi\}$ is empty or reduced to a single point, only the half-spaces can be extremal, up to sets of zero measure in the sense of μ.

Proof. For $y \in \mathbb{R}^n \setminus \{0\}$, the function $t \mapsto \mu(\{\langle x, y \rangle \geq t\})$ is continuous on $\mathbb{R}$. So we may set $H_y = \{x \in \mathbb{R}^n : \langle x, y \rangle \geq m_y\}$, where m_y is such that $\mu(H_y) = a$. If $\mathcal{A}$ is the set of all affine functions dominated by Ψ, then $\Psi(x) = \sup_{l \in \mathcal{A}} l(x)$ for every $x \in \mathbb{R}^n$. For $A \subset \mathbb{R}^n$ with $\mu(A) = a$ and $l \in \mathcal{A}$ we will prove that $l(g_\mu(A)) \leq l(H)$ for some half-space H with $\mu(H) = a$. If l is constant, this is obvious. If l is not constant, there exist $y \in \mathbb{R}^n \setminus \{0\}$ and $c \in \mathbb{R}$ such that $l(x) = \langle x, y \rangle + c$. The main observation is

(7) $$\int\limits_A \langle x, y \rangle d\mu - \int\limits_{H_y} \langle x, y \rangle\, d\mu = \int\limits_{A \setminus H_y} \langle x, y \rangle d\mu - \int\limits_{H_y \setminus A} \langle x, y \rangle\, d\mu$$
$$\leq m_y \mu(A \setminus H_y) - m_y \mu(H_y \setminus A) = 0,$$

which implies $l(g_\mu(A)) \leq l(g_\mu(H_y))$ and proves (6). Let $A \subset \mathbb{R}^n$ be an extremal set in (6). If Ψ does not attain its minimum at $g_\mu(A)$, there is a non-constant $l \in \mathcal{A}$, $l(x) = \langle x, y \rangle + c$, such that $l(g_\mu(A)) = \Psi(g_\mu(A))$. But if $\mu(A \setminus H_y) \neq 0$ then inequality (7) is strict. Now if Ψ attains its minimum at $g_\mu(A)$, then every set of measure a should have the same barycentre; in particular $g_\mu(H_y) = g_\mu(H_{-y})$ for any fixed $y \neq 0$. But as before this would imply $\mu(H_y \setminus H_{-y}) = \mu(H_{-y} \setminus H_y) = 0$, which certainly cannot hold unless $a = \mu(H_y) = 0$ or 1. ∎

REMARK. When Ψ is the Euclidean norm, from the proof above and with the same notation one gets

$$\sup_{\mu(A)=a} |g_\mu(A)| = \sup_{|y|=1} |g_\mu(H_y)| = \sup_{|y|=1} \langle g_\mu(H_y), y \rangle.$$

Now, we give a simple proof of (3): Let $f : \mathbb{R}^n \to [0, 1]$ and consider

$$S = \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : t \leq \Phi^{-1}(f(x))\}.$$

It is clear by Fubini that $\gamma_{n+1}(S) = \int_{\mathbb{R}^n} f(x)\,d\gamma_n(x)$. The components

$(t_g(S), x_g(S))$ of the Gaussian "barycentre" of S, $\int_S (t,x)\, d\gamma_{n+1}(t,x)$, are

$$t_g(S) = \int_{\mathbb{R}^n} \int_{-\infty}^{\Phi^{-1}(f(x))} t\, d\gamma_1(t)\, d\gamma_n(x) = -\int_{\mathbb{R}^n} \varphi(\Phi^{-1}(f(x)))\, d\gamma_n(x),$$

$$x_g(S) = \int_{\mathbb{R}^n} x \int_{-\infty}^{\Phi^{-1}(f(x))} d\gamma_1(t)\, d\gamma_n(x) = \int_{\mathbb{R}^n} x f(x)\, d\gamma_n(x).$$

From the previous lemma and from the computation for half-spaces we get $|(t_g(S), x_g(S))| \le I(\gamma_{n+1}(S))$. Equivalently,

$$\sqrt{\left(\int_{\mathbb{R}^n} I(f)\, d\gamma_n\right)^2 + \left|\int_{\mathbb{R}^n} x f(x)\, d\gamma_n(x)\right|^2} \le I\left(\int_{\mathbb{R}^n} f\, d\gamma_n\right),$$

with equality if and only if the subgraph $\Phi^{-1} \circ f$ is a half-space. When f has suitable properties, integration by parts gives (3) with the same equality cases.

As a conclusion, inequality (3) in $\mathbb{R}^n$ is nothing else than the shift inequality (5) in $\mathbb{R}^{n+1}$. This strengthening of the shift inequality is possible because in the Gaussian case the "shift" function I does not depend on the dimension.

3. Shift inequalities of Gaussian type. We begin by deriving the equivalence of different forms of the shift inequality of Gaussian type. In particular, it is equivalent to the reverse log-Sobolev inequality (see [Be]) involving the entropy

$$\mathrm{Ent}_\mu(f) = \int f \log f\, d\mu - \left(\int f\, d\mu\right) \log\left(\int f\, d\mu\right).$$

Some of the following statements have been shown by Bobkov in the setting of product measures. For this reason, we will skip some details. The interested reader will find them in [Bo2].

In this section we will consider a measure μ on $\mathbb{R}^n$ with a density ϱ a.e. positive which satisfies the following condition: there exists a Borel function $\Lambda : \mathbb{R}^n \to \mathbb{R}^n$ such that we have the following integration by parts:

$$(8) \qquad \int_{\mathbb{R}^n} \nabla f\, d\mu = \int_{\mathbb{R}^n} f\Lambda\, d\mu$$

for every smooth and compactly supported $f : \mathbb{R}^n \to \mathbb{R}$. In other words, the derivative of ϱ in the sense of distributions is a (vector-valued) measure which has a density $-\Lambda$ with respect to μ. If we note by $\nabla\varrho$ the derivative of ϱ in the sense of distributions we have $\varrho\Lambda = -\nabla\varrho$. When the density $\varrho = e^{-V}$ is smooth, the measure μ satisfies the required integration by parts and the above equality holds in $\mathbb{R}^n$, and we have $\Lambda = -(\nabla\varrho)/\varrho = \nabla V$.

For a measure μ of the form $d\mu = \varrho\, dx$, Λ as above and $c > 0$, let us define the following properties:

$P_1(c)$:   For every compactly supported smooth function $f : \mathbb{R}^n \to [0,1]$,

$$\sqrt{\Big(\int_{\mathbb{R}^n} I(f)\, d\mu \Big)^2 + \frac{1}{c^2}\Big| \int_{\mathbb{R}^n} \nabla f\, d\mu \Big|^2} \leq I\Big(\int_{\mathbb{R}^n} f\, d\mu \Big).$$

$P_2(c)$:   For every compactly supported smooth function $f : \mathbb{R}^n \to [0,1]$,

$$\Big| \int_{\mathbb{R}^n} \nabla f\, d\mu \Big| \leq c I\Big(\int_{\mathbb{R}^n} f\, d\mu \Big).$$

$P_3(c)$:   For every Borel set $A \subset \mathbb{R}^n$, $|\int_A \Lambda\, d\mu| \leq c I(\mu(A))$.

$P_4(c)$:   For every Borel set $A \subset \mathbb{R}^n$,

$$\Big| \int_A \Lambda\, d\mu \Big| \leq c \mu(A)\sqrt{\log \frac{1}{\mu(A)}}.$$

$P_5(c)$:   The measure μ is such that $\int_{\mathbb{R}^n} \exp(|\Lambda/c|^2)\, d\mu \leq 2$.

$P_6(c)$:   For every compactly supported smooth function $f : \mathbb{R}^n \to \mathbb{R}^+$,

$$\Big| \int_{\mathbb{R}^n} \nabla f\, d\mu \Big|^2 \leq c^2 \Big(\int_{\mathbb{R}^n} f\, d\mu \Big)\mathrm{Ent}_\mu(f).$$

PROPOSITION 2. *Let μ be a probability measure on $\mathbb{R}^n$ with a.e. positive density ϱ and Λ as above. Then the assertions*

$$P_i : \quad \textit{There exists } c > 0 \textit{ such that } P_i(c)$$

are equivalent for $i = 1, \ldots, 6$. Furthermore, the following control on the constants holds:

$$P_1(c) \; \Leftrightarrow \; P_2(c) \; \Leftrightarrow \; P_3(c),$$
$$P_3(c) \; \Rightarrow \; P_4(c\sqrt{2}) \quad \textit{and} \quad P_4(c) \; \Rightarrow \; P_3(c\sqrt{2}),$$
$$P_1(c) \; \Rightarrow \; P_6(c\sqrt{2}) \quad \textit{and} \quad P_6(c) \; \Rightarrow \; P_4(c),$$
$$P_4(c) \; \Rightarrow \; P_5(c\sqrt{3n}) \quad \textit{and} \quad P_5(c) \; \Rightarrow \; P_6(c\sqrt{8n}).$$

REMARK. As we already mentioned, Bobkov [Bo2] has proved that $P_2(c)$ is equivalent to the global form of the shift inequality: for every Borel set $A \subset \mathbb{R}^n$ and every vector $h \in \mathbb{R}^n$,

$$\Phi(\Phi^{-1}(\mu(A)) - c|h|) \leq \mu(A + h) \leq \Phi(\Phi^{-1}(\mu(A)) + c|h|).$$

Therefore the assertions P_i, $i \leq 6$, are also equivalent to the global form of the shift inequality.

Proof of Proposition 2. Let $c > 0$. We first prove that $P_1(c)$ and $P_2(c)$ are equivalent. Of course $P_1(c) \Rightarrow P_2(c)$. For the converse suppose that for

every smooth function $g : \mathbb{R}^n \to [0,1]$ one has

$$(9) \qquad \left| \int_{\mathbb{R}^n} \nabla g \, d\mu \right| \leq cI \left(\int_{\mathbb{R}^n} g \, d\mu \right).$$

We now follow an argument of Bakry and Ledoux [BL]. Let $f : \mathbb{R}^n \to [0,1]$ be a smooth function, and let ν be its distribution with respect to μ. One may assume that ν is absolutely continuous with respect to Lebesgue's measure on $\mathbb{R}$. Let $r \in \mathbb{R}$ and $\varepsilon > 0$. Define $\psi_\varepsilon(x) = 1$ if $x \leq r$, 0 if $x \geq r + \varepsilon$, and $1 - (x - r)/\varepsilon$ in between. If we apply (9) to $g = \psi_\varepsilon(f)$ and let $\varepsilon \to 0$, we get

$$(10) \qquad \frac{1}{c} N'(r) |\theta(r)| \leq I(N(r))$$

where $N(r) := \nu([0, r]) = \mu(f \leq r)$ and $\theta(r) := \mathbb{E}[\nabla f | f = r]$. Next, we set $k = N^{-1} \circ \Phi$. Differentiation yields $k' \cdot N' \circ k = \varphi$. Thus, inequality (10) applied to $r = k(x)$ becomes

$$(11) \qquad \frac{1}{c} |\theta(k)| \leq k'.$$

Next, we apply (3) to k:

$$\sqrt{\left(\int_{\mathbb{R}} I(k) \, d\gamma_1 \right)^2 + \left(\int_{\mathbb{R}} k' \, d\gamma_1 \right)^2} \leq I \left(\int_{\mathbb{R}} k \, d\gamma_1 \right).$$

Since k is non-decreasing and the law of k under γ_1 is ν, one has, using (11),

$$(12) \qquad \sqrt{\left(\int_0^1 I(r) \, d\nu(r) \right)^2 + \frac{1}{c^2} \left(\int_0^1 |\theta(r)| \, d\nu(r) \right)^2} \leq I \left(\int_0^1 r \, d\nu(r) \right).$$

Recall that ν is the law of f under μ and that

$$\int |\theta(r)| \, d\nu(r) \geq \left| \int \theta(r) \, d\nu(r) \right| = \left| \int \nabla f \, d\mu \right|.$$

So using (12) one gets the result.

To prove that $P_3(c)$ follows from $P_2(c)$, we use $P_2(c)$ and the relation $\int \nabla f \, d\mu = \int f \Lambda \, d\mu$ for approximations of characteristic functions of sets. Conversely, as noticed by Bobkov [Bo2], in the inequality

$$\left| \int_{\mathbb{R}^n} f \Lambda \, d\mu \right| \leq cI \left(\int_{\mathbb{R}^n} f \, d\mu \right),$$

the left hand side is convex in f and the right hand side is concave in f. Therefore it is enough to prove this inequality for characteristic functions of sets. Thus $P_3(c)$ implies $P_2(c)$.

The implications $P_3(c) \Rightarrow P_4(c\sqrt{2}) \Rightarrow P_3(2c)$ follow directly from Lemma 4.2 of [Bo2]:

$$p\sqrt{\frac{1}{2} \log \frac{1}{p}} \leq I(p) \leq p\sqrt{2 \log \frac{1}{p}}$$

where the left hand side inequality holds for $p \in [0, 1/2]$ and the right hand side inequality for $p \in [0, 1]$. Notice that, since $\int_{\mathbb{R}^n} \Lambda \, d\mu = 0$ and I is symmetric, it is enough to prove $P_3(c)$ for sets A with $\mu(A) \leq 1/2$.

We now deal with the inverse log-Sobolev inequality $P_6(c)$. The fact that $P_1(c)$ implies $P_6(c\sqrt{2})$ was explained by Beckner during his lectures at the IHP (see [L] for a written reference). His method is as follows: apply $P_1(c)$ to εf for a bounded and smooth $f : \mathbb{R}^n \to \mathbb{R}^+$ and use $I(\varepsilon) \sim \varepsilon \sqrt{2\log(1/\varepsilon)}$ as $\varepsilon \to 0^+$. Conversely, if $P_6(c)$ holds then $P_4(c)$ also holds: we use again $\int_{\mathbb{R}^n} \nabla f \, d\mu = \int_{\mathbb{R}^n} f \Lambda \, d\mu$ and then apply $P_6(c)$ to characteristic functions.

The equivalence of P_4 and P_5 is shown in [Bo2] for product measures, but the method applies in general. Assume that $P_4(c)$ holds. In particular for any $i \leq n$,

$$\left| \int_A \Lambda_i \, d\mu \right| \leq c\mu(A) \sqrt{\log \frac{1}{\mu(A)}},$$

where $\Lambda_i : \mathbb{R}^n \to \mathbb{R}$ is the ith coordinate of Λ. Applied to level sets of Λ_i, the latter yields $\mu(|\Lambda_i| > t) \leq 2\exp(-t^2/c^2)$. After a standard calculation, one gets

$$\int_{\mathbb{R}^n} \exp\left(\frac{(\Lambda_i)^2}{3c^2} \right) d\mu \leq 2.$$

From Hölder's inequality, we conclude that $\int_{\mathbb{R}^n} \exp(|\Lambda|^2/(3nc^2)) \, d\mu \leq 2$.

Conversely, assume that $P_5(c)$ holds. Then for any $i \leq n$,

$$(13) \qquad \int_{\mathbb{R}^n} e^{(\Lambda_i/c)^2} \, d\mu \leq 2.$$

The argument of [Bo2] proves that for every bounded function $f : \mathbb{R}^n \to \mathbb{R}^+$,

$$\left| \int_{\mathbb{R}^n} f \Lambda_i \, d\mu \right|^2 \leq 8c^2 \left(\int_{\mathbb{R}^n} f \, d\mu \right) \mathrm{Ent}_\mu(f).$$

Summing these inequalities for $i = 1, \ldots, n$, we obtain $P_6(c\sqrt{8n})$. This completes the proof of the proposition. ∎

REMARK. One can easily check that $P_1(c)$ has the tensorization property: if μ on $\mathbb{R}^n$ and ν on $\mathbb{R}^p$ satisfy $P_1(c)$, then so does the product measure $\mu \otimes \nu$. Indeed, let $f : \mathbb{R}^{n+p} \to [0, 1]$ be a smooth compactly supported function. One has

$$\left(\int_{\mathbb{R}^{n+p}} I(f) \, d\mu \otimes \nu \right)^2 + \frac{1}{c^2} \left| \int_{\mathbb{R}^{n+p}} \nabla f \, d\mu \otimes \nu \right|^2$$

$$= \left(\int_{\mathbb{R}^n} d\mu(x) \int_{\mathbb{R}^p} I(f(x,y)) \, d\nu(y) \right)^2 + \frac{1}{c^2} \left| \int_{\mathbb{R}^n} d\mu(x) \int_{\mathbb{R}^p} \nabla_y f(x,y) \, d\nu(y) \right|^2$$

$$+ \frac{1}{c^2} \left| \int_{\mathbb{R}^n} d\mu(x) \nabla_x \int_{\mathbb{R}^p} f(x,y) \, d\nu(y) \right|^2$$

$$\le \left(\int\limits_{\mathbb{R}^n} d\mu(x) \sqrt{\left(\int\limits_{\mathbb{R}^p} I(f(x,y))\, d\nu(y) \right)^2 + \frac{1}{c^2} \left| \int\limits_{\mathbb{R}^p} \nabla_y\, f(x,y)\, d\nu(y) \right|^2} \right)^2$$

$$+ \frac{1}{c^2} \left| \int\limits_{\mathbb{R}^n} d\mu(x)\, \nabla_x \int\limits_{\mathbb{R}^p} f(x,y)\, d\mu(y) \right|^2$$

$$\le \left(\int\limits_{\mathbb{R}^n} d\mu(x) I\left(\int\limits_{\mathbb{R}^p} f(x,y)\, d\nu(y) \right) \right)^2 + \frac{1}{c^2} \left| \int\limits_{\mathbb{R}^n} d\mu(x)\, \nabla_x \int\limits_{\mathbb{R}^p} f(x,y)\, d\nu(y) \right|^2$$

$$\le I^2 \left(\int\limits_{\mathbb{R}^n} d\mu(x) \int\limits_{\mathbb{R}^p} f(x,y)\, d\nu(y) \right) = I^2 \left(\int\limits_{\mathbb{R}^{n+p}} f\, d\mu \otimes \nu \right)$$

where we used successively the Minkowski inequality, $P_1(c)$ for the measure ν and $P_1(c)$ for the measure μ. The above computations can be handled in a more abstract setting, for instance in the case of discrete measures (on the discrete cube or on $\mathbb{Z}^n$). One has only to define a suitable notion of gradient. It follows from Proposition 2 that $P_2(c)$ and $P_3(c)$ also have the tensorization property: if μ satisties $P_2(c)$, then so do the product measures $\mu^{\otimes n}$ for $n \ge 1$. Notice that [Bo2] proved only that the $\mu^{\otimes n}$ satisfy $P_2(4c)$.

The integrability criterion P_5 is an easy way to derive shift inequalities, but with poor constants. In the following, we get sharper constants by extending the semigroup method to Boltzmann measures. One should remark that, as far as we know, among the equivalent forms of the shift inequality, only the reverse log-Sobolev P_6 and the reverse Bobkov P_1 inequalities can be proved with semigroups.

THEOREM 3. *Let μ be a probability on $\mathbb{R}^n$ such that $d\mu = e^{-V} dx$. Assume that V is twice differentiable on $\mathbb{R}^n$ and that its Hessian, considered as an endomorphism of the Euclidean space, satisfies $V'' \ge \varepsilon \mathrm{Id}_{\mathbb{R}^n}$ at infinity for some $\varepsilon > 0$. If $\|V''\|$ is uniformly bounded by $c > 0$, then for every smooth function $f : \mathbb{R}^n \to [0,1]$,*

$$\sqrt{\left(\int\limits_{\mathbb{R}^n} I(f)\, d\mu \right)^2 + \frac{1}{c} \left| \int\limits_{\mathbb{R}^n} \nabla f\, d\mu \right|^2} \le I\left(\int\limits_{\mathbb{R}^n} f\, d\mu \right).$$

Proof. The measure μ is invariant for the semigroup P_t with generator

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

In the case of the Gaussian measure, this is the Ornstein–Uhlenbeck semigroup. Because of the limit assumptions on V, we have $P_t(f) \to \int f\, d\mu$ as $t \to \infty$, for every bounded smooth function f on $\mathbb{R}^n$. We prove by differentiation that the quantity

$$J(t) = \left(\int I(P_t f)\,d\mu\right)^2 + \frac{1}{c}\left|\int \nabla P_t f\,d\mu\right|^2$$

$$= \left(\int I(P_t f)\,d\mu\right)^2 + \frac{1}{c}\left|\int P_t f\nabla V\,d\mu\right|^2$$

is non-decreasing. Its derivative $J'(t)$ is equal to

$$2\int I(P_t f)\,d\mu \int I'(P_t f)LP_t f\,d\mu + \frac{2}{c}\left\langle \int P_t f\nabla V\,d\mu, \int LP_t f\nabla V\,d\mu\right\rangle.$$

We set $F = P_t f$ and use integration by parts for L: for suitable functions u and v, one has

$$2\int uLv\,d\mu = -\int \langle \nabla u, \nabla v\rangle\,d\mu.$$

Since $II'' = -1$, we get

$$J'(t) = -\int I(F)\,d\mu \int I''(F)|\nabla F|^2\,d\mu - \frac{1}{c}\left\langle \int F\nabla V\,d\mu, \int V''\cdot\nabla F\,d\mu\right\rangle$$

$$= \int I(F)\,d\mu \int \frac{|\nabla F|^2}{I(F)}\,d\mu - \frac{1}{c}\left\langle \int F\nabla V\,d\mu, \int V''\cdot\nabla F\,d\mu\right\rangle.$$

But

$$\left|\left\langle \int F\nabla V\,d\mu, \int V''\cdot\nabla F\,d\mu\right\rangle\right| \le \int |\nabla F|\,d\mu \int |V''\cdot\nabla F|\,d\mu \le c\left(\int |\nabla F|\,d\mu\right)^2,$$

so $J'(t) \ge 0$ by Cauchy–Schwarz. ∎

4. Discrete and spherical inequalities. In this section, we first apply the shift inequality for the Gaussian measure to the average isodiametral problem of Ahlswede and Katona [AK]. Let us introduce some notation. We denote by $\varepsilon = (\varepsilon_i)_{i=1}^n$ the elements of $\Omega_n = \{-1,1\}^n$. We consider the uniform probability μ_n on this set, and denote by $\mathbb{E}_n$ the corresponding expectation. If f is a function on Ω_n, we write $\mathbb{E}_n(\varepsilon f)$ for the vector of coordinates

$$\left(\int_{\Omega_n} \varepsilon_i f(\varepsilon)\,d\mu(\varepsilon)\right)_{i=1}^n.$$

The *Hamming distance* on Ω_n is defined by $d(\varepsilon, \eta) = \#\{i : \varepsilon_i \ne \eta_i\}$. Considering Ω_n as a subset of $\mathbb{R}^n$, one also has

$$d(\varepsilon, \eta) = \left|\frac{\varepsilon - \eta}{2}\right|^2 = \frac{n - \langle \varepsilon, \eta\rangle}{2}.$$

The problem is to find subsets $A \subset \Omega_n$ of given cardinality which minimize the average inner distance:

$$\frac{1}{(\#A)^2}\sum_{\varepsilon\in A, \eta\in A} d(\varepsilon, \eta) = \frac{1}{2(\#A)^2}\sum_{\varepsilon,\eta\in A}(n - \langle \varepsilon, \eta\rangle) = \frac{1}{2}\left(n - \left|\frac{\mathbb{E}\varepsilon\mathbf{1}_A}{\mu_n(A)}\right|^2\right).$$

Thus it is equivalent to the maximization of the norm of barycentres. Finding the extremal sets seems difficult. Ahlswede and Althöfer [A-Al] proved that

among the families of sets $A_n \subset \Omega_n$ of cardinality $2^{\lambda n}$, for $\lambda \in (0,1)$, the Hamming spheres are asymptotically optimal. Althöfer and Sillke [Al-Si] gave a lower bound of the inner distance, which holds for all cardinalities. With our notation, their result states that for $A \subset \Omega_n$,

$$|\mathbb{E}_n \varepsilon \mathbf{1}_A| \leq \sqrt{\mu_n(A)(1 - \mu_n(A))}.$$

When $\mu_n(A) = 1/2$, this is an equality for $\{\varepsilon : \varepsilon_1 = 1\}$ and for the other half-cubes. The Gaussian shift inequality has a discrete version which provides an improvement of the latter estimate.

PROPOSITION 4. *Let f be defined on Ω_n with values in $[0,1]$. Then*

$$(14) \qquad (\mathbb{E}_n I(f))^2 + \frac{2}{\pi} |\mathbb{E}_n \varepsilon f|^2 \leq I^2(\mathbb{E}_n f),$$

with equality if and only if f is constant or is the characteristic function of a half-cube.

Proof. For $t \in \mathbb{R}$ let $s(t) = 1$ if $t > 0$ and $s(t) = -1$ else. We define a function F on $\mathbb{R}^n$ by $F(x_1, \ldots, x_n) = f(s(x_1), \ldots, s(x_n))$. The law of F under γ_n is the law of f under μ_n. Furthermore

$$\int_{\mathbb{R}^n} x_k F(x) \, d\gamma_n(x) = \sum_{\varepsilon \in \Omega_n} \int x_k f(\varepsilon) \mathbf{1}_{\{\varepsilon_i x_i > 0, \forall i\}} \, d\gamma_n(x)$$

$$= \sum_{\varepsilon \in \Omega_n} f(\varepsilon) \varepsilon_k \int y_k \mathbf{1}_{\{y_i > 0, \forall i\}} \, d\gamma_n(y) = \sqrt{2/\pi} \, \mathbb{E}_n(\varepsilon_k f(\varepsilon)).$$

The proposition follows from the inequality

$$\sqrt{\left(\int_{\mathbb{R}^n} I(F) \, d\gamma_n \right)^2 + \left| \int_{\mathbb{R}^n} x F(x) \, d\gamma_n(x) \right|^2} \leq I\left(\int_{\mathbb{R}^n} F \, d\gamma_n \right),$$

where equality occurs if and only if the subgraph of $\Phi^{-1} \circ F$ in $\mathbb{R}^{n+1}$ is a half-space. ∎

REMARKS. (i) As in Proposition 2, one can derive from (14) an inverse form of the Gross log-Sobolev inequality in Ω_n.

(ii) The derivation of these discrete inequalities does not loose much information: one can recover a weak form of the inequality on Gauss space by a standard central limit argument. Let f be a continuous compactly supported function on $\mathbb{R}$ with values in $[0,1]$. For $n \in \mathbb{N}$, let

$$f_n(\varepsilon_1, \ldots, \varepsilon_n) = f\left(\frac{\varepsilon_1 + \ldots + \varepsilon_n}{\sqrt{n}} \right).$$

When n tends to infinity, $\mathbb{E}_n f_n$ tends to $\int f \, d\gamma_1$ and $\mathbb{E}_n I(f_n)$ to $\int I(f) \, d\gamma_1$. For the barycentre term, using the permutation invariance of f_n, for all

$i = 1, \ldots, n$ we get $\mathbb{E}_n \varepsilon_i f_n = \mathbb{E}_n(\varepsilon_1 + \ldots + \varepsilon_n) f_n / n$. Thus

$$|\mathbb{E}_n \varepsilon f_n| = \Big(\sum_{i=1}^n (\mathbb{E}_n \varepsilon_i f_n)^2 \Big)^{1/2} = \sqrt{n} \, \Big| \mathbb{E}_n \frac{\varepsilon_1 + \ldots + \varepsilon_n}{n} f_n \Big|$$

$$= \Big| \mathbb{E}_n \frac{\varepsilon_1 + \ldots + \varepsilon_n}{\sqrt{n}} f \Big(\frac{\varepsilon_1 + \ldots + \varepsilon_n}{\sqrt{n}} \Big) \Big|$$

tends to $|\int_{\mathbb{R}} x f(x) \, d\gamma_1(x)|$. Hence Proposition 4 implies that γ_1 satisfies the inequality $P_1(\sqrt{\pi/2})$, which implies the same for γ_n, by the tensorization property noticed in Section 3.

Applying Proposition 4 to characteristic functions of sets, we get

PROPOSITION 5. *Let* $A \subset \{-1, 1\}^n$. *Then*

$$|g_{\mu_n}(A)| = |\mathbb{E}_n \varepsilon \mathbf{1}_A| \leq \sqrt{\pi/2} \, I(\mu_n(A)),$$

with equality if and only if A *is* Ω_n, *a half-cube or the empty set.*

This result improves [Al-Si]. Indeed, for $t \in [0, 1]$, one has

$$\sqrt{\pi} \, I(t) \leq \sqrt{2t(1-t)}.$$

To prove this inequality, one can assume that $1/2 \leq t \leq 1$ since both functions are symmetric with respect to $1/2$. Let $r = \Phi^{-1}(t) > 0$. We have $[-r, r]^2 \subset r\sqrt{2} B_2^2$, where B_2^2 is the Euclidean unit ball of $\mathbb{R}^2$. Thus $\gamma_2([-r, r]^2) = (2\Phi(r) - 1)^2 = 1 - 4\Phi(r) + 4\Phi^2(r)$ is less than $\gamma_2(r\sqrt{2} B_2^2) = \int_0^{r\sqrt{2}} e^{-s^2/2} s \, ds = 1 - e^{-r^2} = 1 - 2\pi \varphi(r)^2$. The result follows.

Up to the multiplicative factor $\sqrt{\pi/2}$, the inequality of Proposition 5 is the best possible dimension free estimate. Indeed, taking $f = \mathbf{1}_{(-\infty, t]}$ in the previous remark, one can see that the Hamming balls

$$A_{n,t} = \Big\{ \varepsilon : \sum_{i=1}^n \varepsilon_i \leq \sqrt{n}\, t \Big\}$$

satisfy $|\mathbb{E}_n \varepsilon \mathbf{1}_{A_{n,t}}| \sim \big| \int_{-\infty}^t x \, d\gamma_1(x) \big| = I(\Phi(t)) \sim I(\mu_n(A_{n,t}))$ when n is large.

We denote by σ_n the uniform probability on S^{n-1}. The Gaussian shift inequality also has a spherical version.

PROPOSITION 6. *Let* f *be defined on* S^{n-1} *with values in* $[0, 1]$. *Then*

$$(15) \qquad \Big(\int_{S^{n-1}} I(f) \Big)^2 + c_n^2 \Big| \int_{S^{n-1}} u f(u) \, d\sigma_n(u) \Big|^2 \leq I^2 \Big(\int_{S^{n-1}} f \Big),$$

with $c_n = \sqrt{2} \, \Gamma((n+1)/2)/\Gamma(n/2) \sim \sqrt{n}$ *when* n *is large and with equality if and only if* f *is constant or is the characteristic function of a half-sphere.*

Proof. We define a function F on $\mathbb{R}^n \setminus \{0\}$ by $F(x) = f(x/|x|)$. The law of F under γ_n is the law of f under σ_n. Furthermore

$$\int_{\mathbb{R}^n} x F(x)\, d\gamma_n(x) = c_n \int_{S^{n-1}} u f(u)\, d\sigma_n(u)$$

where $c_n = |g_{\gamma_n}(H)|/|g_{\sigma_n}(H)|$, H being a half-space of measure $1/2$. A little calculation gives $c_n = \sqrt{2}\, \Gamma((n+1)/2)/\Gamma(n/2) \sim \sqrt{n}$ when n is large. As in the discrete case, the proposition follows. ∎

REMARKS. (i) Inequality (15) is optimal in the sense that, by Poincaré's limit argument, it allows us to recover the initial shift inequality for γ_n.

(ii) Lemma 1 implies that among the subsets of S^{n-1} of given measure the caps have the largest barycentre norm, whereas inequality (15) implies this only for the sets of measure one half.

5. Barycentres in the cube. The aim of this section is to derive an analogue of Proposition 5 for subsets of volume $1/2$ in the "continuous" cube. The proof involves classical methods of convexity theory. If $K \subset \mathbb{R}^n$ is convex with k-dimensional affine hull, $|K|$ will be its k-dimensional Lebesgue measure.

PROPOSITION 7. *Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a non-increasing function. Then the function*

$$F(p) = (p+1) \int_0^{+\infty} t^p f(t)\, dt \quad \text{if } p > -1 \quad \text{and} \quad F(-1) = f(0)$$

is a log-convex function of p on $[-1, +\infty[$. Moreover F is log-affine if and only if f is constant on its support.

Proof. For $s > 0$, we define $h(s) = \sup\{t : f(t) \geq s\}$. Let $p > -1$. Then

$$F(p) = \int_0^{+\infty} (p+1) t^p \int_0^{f(t)} ds\, dt = \int_0^{f(0)} \int_0^{h(s)} (p+1) t^p\, dt\, ds = \int_0^{f(0)} h(s)^{p+1}\, ds.$$

The result follows from Hölder's inequality. ∎

REMARK. As a corollary, we recover a lemma of Milman and Pajor [Mi-P]:

$$\left(\frac{F(p)}{F(-1)} \right)^{1/(p+1)} = \left((p+1) \int_0^{+\infty} t^p \frac{f(t)}{f(0)}\, dt \right)^{1/(p+1)}$$

is an increasing function on $]-1, +\infty[$. Proposition 7 is also related to results of Borell [Bor].

COROLLARY 8. *Let K be a symmetric convex body in $\mathbb{R}^n$ and $u \in S^{n-1}$. Then*

$$\int\limits_K |\langle x, u \rangle|\, dx \le \frac{\sqrt{3}}{2} |K|^{1/2} \Big(\int\limits_K |\langle x, u \rangle|^2\, dx \Big)^{1/2},$$

with equality if and only if K is cylindrical in the direction u, i.e. there exists $x \in \mathbb{R}^n$ such that $K = K \cap u^\perp + [-x, x]$.

Proof. Let $f(t) = |K \cap \{x \in \mathbb{R}^n : \langle x, u \rangle = t\}|$ be the parallel section function. We know by the Brunn–Minkowski theorem that f is non-increasing on $\mathbb{R}^+$ and is constant if and only if K is cylindrical in the direction u. Hence, we can apply Proposition 7 to f. We get $F(1) \le (F(0)F(2))^{1/2}$ (with the same notation). Moreover, Fubini's theorem yields, for $p > -1$,

$$\int\limits_K |\langle x, u \rangle|^p\, dx = 2 \int\limits_0^{+\infty} t^p f(t)\, dt = \frac{2}{p+1} F(p).$$

This proves the corollary. ∎

REMARK. This corollary gives the sharp constant in the comparison between the Legendre ellipsoid of K and its centroid body $Z(K)$ (defined by $\|x\|_{Z(K)^\circ} = (2|K|)^{-1} \int_K |\langle x, u \rangle|\, dx$). The fact that such a constant exists was noticed in Milman–Pajor [Mi-P].

PROPOSITION 9. *Among the Borel subsets A of the unit cube $[-1/2, 1/2]^n$ such that $|A| = 1/2$, the half-cube $\{x_1 \le 0\}$ has the largest norm of barycentre.*

Proof. Let $A \subset [-1/2, 1/2]^n$ be such that $|A| = 1/2$. By the remark after Lemma 1, we have, for $g(A) = \int_A x\, dx$,

$$|g(A)| \le \sup_{u \in S^{n-1}} \langle g(H_u), u \rangle,$$

where $H_u = \{x \in [-1/2, 1/2]^n : \langle x, u \rangle \ge 0\}$. Corollary 8 yields

$$\langle g(H_u), u \rangle = \frac{1}{2} \int\limits_{[-1/2, 1/2]^n} |\langle x, u \rangle|\, dx \le \frac{\sqrt{3}}{4} \Big(\int\limits_{[-1/2, 1/2]^n} |\langle x, u \rangle|^2 dx \Big)^{1/2} = \frac{1}{8}.$$

Indeed, the last integral does not depend on $u \in S^{n-1}$ (the cube is in isotropic position). There is equality in the previous inequality if and only if the cube is cylindrical in the direction u; this happens if and only if u is a basis vector. ∎

References

[A-Al]   R. Ahlswede and I. Althöfer, *The asymptotic behaviour of diameters in the average*, J. Combin. Theory Ser. B 61 (1994), 167–177.

[AK]   R. Ahlswede and G. Katona, *Contributions to the geometry of Hamming spaces*, Discrete Math. 17 (1977), 1–22.

[Al-Si] I. Althöfer and T. Sillke, *An "average distance" inequality for large subsets of the cube*, J. Combin. Theory Ser. B 56 (1992), 296–301.

[BL]   D. Bakry and M. Ledoux, *Lévy–Gromov isoperimetric inequality for an infinite dimensional diffusion generator*, Invent. Math. 123 (1996), 259–281.

[Ba-M]   F. Barthe and B. Maurey, *Some remarks on isoperimetry of Gaussian type*, Ann. Inst. H. Poincaré Probab. Statist. 36 (2000), 419–434.

[Be]   W. Beckner, *Geometric asymptotics and the logarithmic Sobolev inequality*, Forum Math. 11 (1999), 105–137.

[Bo1]   S. G. Bobkov, *An isoperimetric inequality on the discrete cube*, *and an elementary proof of the isoperimetric inequality in Gauss space*, Ann. Probab. 25 (1997), 206–214.

[Bo2]   —, *The size of singular component and shift inequalities*, ibid. 27 (1999), 416–431.

[Bor]   C. Borell, *Complements of Lyapounov's inequality*, Math. Ann. 205 (1973), 323–331.

[F]   J. Feldman, *Examples of non-Gaussian quasi-invariant distributions on Hilbert space*, Trans. Amer. Math. Soc. 99 (1961), 342–349.

[Ku-Li] J. Kuelbs and W. V. Li, *Some shift inequalities for Gaussian measures*, in: Progr. Probab. 43, Birkhäuser, 1998, 233–243.

[L]   M. Ledoux, *The geometry of Markov diffusion generators*, Ann. Fac. Sci. Toulouse 9 (2000), 305–366.

[Mi-P]   V. D. Milman and A. Pajor, *Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space*, in: Geometric Aspects of Functional Analysis, Lecture Notes in Math. 1376, Springer, 1989, 64–104.

[S]   L. A. Shepp, *Distinguishing a sequence of random variables from a translate of itself*, Ann. Math. Statist. 36 (1965), 1107–1112.

Équipe d'Analyse et de Mathématiques Appliquées, ESA 8050
Université de Marne-la-Vallée
Cité Descartes, 5 Boulevard Descartes
Champs-sur-Marne
77454 Marne-la-Vallée Cedex 2, France
E-mail: barthe@math.univ-mlv.fr
      cordero@math.univ-mlv.fr
      fradeliz@math.univ-mlv.fr