A CHARACTERIZATION OF GAUSSIAN MEASURES VIA THE ISOPERIMETRIC PROPERTY OF HALF–SPACES *

S. G. Bobkov † and C. Houdré ‡

July 25, 1995

Abstract

If the half–spaces of the form $\{x \in \mathbf{R}^n : x_1 \leq c\}$ are extremal in the isoperimetric problem for the product measure μ^n, $n \geq 2$, then μ is Gaussian.

1 Introduction

Let μ be a probability measure on the real line $\mathbf{R}$, and let μ^n be the n–fold tensor product of μ with itself. Let A be a Borel set in $\mathbf{R}^n$, and for any $h > 0$, let A^h be its open h–neighbourhood (with respect to the Euclidean metric), i.e., $A^h = \{x \in \mathbf{R}^n : \|x - a\|_2 < h$ for some $a \in A\}$. The isoperimetric problem for $(\mathbf{R}^n, \mu^n)$ consists in minimizing

$$\mu^n(A^h), \tag{1.1}$$

*Key words: Isoperimetry, Gaussian measure

†Research supported in part by ISF NXZ000, NXZ300 and Grant No. 94–1.4–57 from the Grant Center of Fundamental Sciences in Sankt–Petersburg University

‡Research supported in part by an NSF Postdoctoral Fellowship. This author enjoyed the hospitality of the Steklov Mathematical Institute (Sankt–Petersburg branch) and of the Department of Mathematics, University of Syktyvkar, Russia, while part of this research was carried out.

among all the Borel sets A of μ^n–measure greater or equal to p, where $p \in (0,1)$ and $h > 0$ are fixed numbers.

When μ is a Gaussian measure, (1.1) attains its minimum at any half-space of measure p, and therefore, this can be expressed as the isoperimetric inequality

$$\mu^n(A^h) \geq \mu^n(B^h), \tag{1.2}$$

where B is a standard half–space $\{x \in \mathbf{R}^n : x_1 \leq c\}$ of μ^n–measure p, where $h > 0$ is arbitrary, and where clearly c depends only on p. This deep property of Gaussian measures was discovered by V.N.Sudakov and B.S.Tsirel'son [7] and independently by C.Borell [3]. Their proofs are similar and rely on the isoperimetric property of the balls on the sphere (the Lévy–Schmidt theorem). It should however be noted here, that the first instance in which isoperimetric methods are applied to study Gaussian processes appears in Landau and Shepp [5]; and that they already relied on the Lévy–Schmidt theorem. They showed, there, the extremal property of the half–spaces in another closely related problem of isoperimetric nature. Later a different proof of the extremal property of the half–spaces in the isoperimetric problem was given, using rearrangement techniques in Gauss space, by A.Ehrhard [4].

Theorem 1.1a *Let $n \geq 2$. Let μ be a probability measure on $\mathbf{R}$ such that* (1.1) *attains its minimum at the standard half–spaces, for all $p \in (0,1)$ and $h > 0$. Then, if it is not a unit mass at a point, μ is Gaussian.*

The case $n = 1$ essentially differs from the case $n \geq 2$, since on the real line, many interesting measures satisfy (1.2). For example, when μ has a continuous positive density, necessary and sufficient conditions for (1.2) to hold are known (see [2], Sec.13). In particular μ has to be symmetric about its median and has to have finite exponential moment. In fact, μ possesses these two properties without any further assumption (see Proposition 2.6 below). Moreover, should μ be symmetric about zero and with finite variance, the hypotheses of Theorem 1.1a can be weakened.

Theorem 1.1b *Let $n \geq 2$ and let $p = 1/2$. Let μ be a symmetric about zero probability measure on $\mathbf{R}$, with finite variance, and such that* (1.1) *attains its minimum at the standard half–spaces, for all $h > 0$. Then, if it is not a unit mass at zero, μ is Gaussian.*

It is worthwhile here to note the crucial rôle of the Euclidean distance in this characterization. For example, if $\|x-a\|_2$ is replaced in the definition of the enlargement A^h by the supremum distance $\|x-a\|_\infty$, then (1.2) holds for a wide family of log-concave distributions [1] (see also [2], Sec.15). In connection with the concentration of measure phenomenon, inequalities similar to (1.2) and for various types of enlargements have been studied by many authors (see, e.g., M.Talagrand [8], M.Ledoux [6] and the references therein).

Clearly, the inequality (1.2) becomes stronger when the dimension n grows, so in essence, Theorem 1.1a,b concern the case of the plane ($n=2$). Moreover, under the assumptions of Theorem 1.1b, one can derive from (1.2) that μ is Gaussian by applying (1.2) to the half–plane

$$A(t)=\left\{(x_1,x_2):\frac{x_1+x_2}{\sqrt{2}}\leq t\right\},\quad t=0.$$

Proof of Theorem 1.1b. Indeed, let ξ and η be independent random variables defined on some probability space $(\Omega,\mathcal{F},\mathbf{P})$, with common distribution μ. Then, the minimal value of the right–hand side of (1.2) provided $\mu(B)\geq 1/2$ is attained at $B=\{x\in\mathbf{R}^2: x_1\leq 0\}$ and is equal to $\mathbf{P}\{\xi<h\}$. Analogously, the minimal value of the left–hand side of (1.2) provided $\mu(A(t))\geq 1/2$ is attained for $t=0$. Since $(A(t))^h=A(t+h)$, we have $\mathbf{P}\{(\xi+\eta)/\sqrt{2}<h\}\geq \mathbf{P}\{\xi<h\}$, for all $h>0$. Therefore,

$$\mathbf{Var}\left(\frac{\xi+\eta}{\sqrt{2}}\right)=4\int_0^{+\infty}\mathbf{P}\left\{\frac{\xi+\eta}{\sqrt{2}}>h\right\}hdh\leq 4\int_0^{+\infty}\mathbf{P}\{\xi>h\}\,hdh=\mathbf{Var}\left(\xi\right).$$

But $\mathbf{Var}\left((\xi+\eta)/\sqrt{2}\right)=\mathbf{Var}\left(\xi\right)$, hence for almost all (with respect to Lebesgue measure) $h>0$, $\mathbf{P}\left\{(\xi+\eta)/\sqrt{2}>h\right\}=\mathbf{P}\left\{\xi>h\right\}$. Therefore, this equality extends to all $h>0$, and thus the random variables $(\xi+\eta)/\sqrt{2}$ and ξ are identically distributed. That is, the characteristic function f of ξ satisfies equation $f^2(t/\sqrt{2})=f(t)$, for all t real. It is then easy to see that this equation characterizes the Gaussian distributions.

Theorem 1.1b is thus proved. Now, in order to establish Theorem 1.1a, we study the one-dimensional case and prove that (1.2) implies the assumptions of Theorem 1.1b. We do not know if the assumption on the finiteness of the variance can be removed in this result.

2 Necessary conditions when n=1

Given a probability measure μ on the real line $\mathbf{R}$, we set

$F(x) = \mu((-\infty, x]), \quad x \in (-\infty, +\infty]$,

$Im(F) = \{F(x) > 0 : x \in (-\infty, +\infty]\}$,

$S(F) = \{x \in (-\infty, +\infty] : F(y) < F(x), \text{ for all } y < x\}$,

$F^{-1}(p) = \inf\{x \in (-\infty, +\infty] : F(x) \geq p\}, \quad p \in (0, 1]$.

$F^{-1}(p)$ is the minimal quantile of order p of F, indeed, since the function F is right–continuous, the above infimum can be replaced by minimum. In particular, for $p \in Im(F)$, $F^{-1}(p)$ is the least solution to $F(x) = p$. Thus, $F(F^{-1}(p)) \geq p$, whenever $p \in (0, 1]$, and $F(F^{-1}(p)) = p$, for all $p \in Im(F)$. $S(F)$ (without the point $x = +\infty$) is a subset of the (closed) support of μ. It is also easy to see that $\mu(S(F)) = 1$.

Lemma 2.1 *F is an increasing bijection from $S(F)$ to $Im(F)$, and F^{-1} restricted to $Im(F)$ is its inverse. Moreover, F^{-1} is left–continuous on* $(0, 1)$.

Lemma 2.2 *For all $p \in (0, 1]$,*

a) $F(F^{-1}(p)) = p \Longleftrightarrow p \in Im(F)$.

b) $x \geq F^{-1}(p) \Longleftrightarrow F(x) \geq p$, whenever $x \in (-\infty, +\infty]$;

c) $x \leq F^{-1}(p) \Longleftrightarrow F(x) \leq p$, whenever $x \in S(F)$,.

Both these lemmas are elementary and so their proofs are omitted. Now, let F and G be the respective distribution function of the probability measures μ and ν.

Lemma 2.3 *The map $U = G^{-1}(F)$ transforms μ into ν if and only if $Im(G) \subset Im(F)$.*

Proof. For the "if part", one can restrict U to $S(F)$. Let $p = F(x)$, $x \in S(F)$, $q = G(t)$, $t \in (-\infty, +\infty]$, so that $p \in Im(F)$, $q \in Im(G)$, hence $q \in Im(F)$. Hence, by lemma 2.2 b) and c):

$$U(x) = G^{-1}(F(x)) = G^{-1}(p) \leq t \Leftrightarrow G(t) \geq p \Leftrightarrow F(x) \leq q \Leftrightarrow x \leq F^{-1}(q).$$

Therefore, $\mu\{U \leq t\} = F(F^{-1}(q)) = q$, since $q \in Im(F)$. The "only if" statement is trivial.

Lemma 2.4 *Assume that, for all $p \in (0, 1)$ and $h > 0$,*

$$F(F^{-1}(p) + h) \leq G(G^{-1}(p) + h). \tag{2.1}$$

Then, the map $U = G^{-1}(F)$ transforms μ into ν, and for all $x \in S(F)$, $h > 0$,

$$U(x+h) \leq U(x) + h. \tag{2.2}$$

Proof. Letting in (2.1) $h \to 0$, gives

$$F(F^{-1}(p)) \leq G(G^{-1}(p)), \tag{2.3}$$

whenever $p \in (0,1)$. Since $F(F^{-1}(1)) = G(G^{-1}(1)) = 1$, (2.3) also holds for $p = 1$. Let $p \in Im(G)$, then by lemma 2.2a), $G(G^{-1}(p)) = p$, hence by (2.3) $F(F^{-1}(p)) \leq p$. But, as noted before, $F(F^{-1}(p)) \geq p$, and so $F(F^{-1}(p)) = p$. Again by lemma 2.2a), we obtain $p \in Im(F)$. Hence, $Im(G) \subset Im(F)$, and by Lemma 2.3, the map U transforms μ into ν. Now take $x \in S(F)$. By Lemma 2.1, $F^{-1}(F(x)) = x$. Applying (2.1) to $p = F(x)$, it thus follows that

$$F(x+h) \leq G(U(x) + h).$$

Since $F(x+h) \geq F(x) > 0$, and since G^{-1} is non–decreasing, we therefore get

$$U(x+h) \leq G^{-1}(G(U(x)+h)).$$

It now remains to show that $G^{-1}(G(U(x)+h)) \leq U(x) + h$. First note that $G^{-1}(G(y)) \leq y$, for all y such that $G(y) > 0$. Also, in case $y = U(x) + h$, we have $G(y) \geq G(U(x)) = G(G^{-1}(F(x))) \geq F(x) > 0$, since $G(G^{-1}(p)) \geq p$, for all $p \in (0,1]$ and since F is positive on $S(F)$. Lemma 2.4 is proved.

Denote by $m_p(\cdot)$ the minimal quantile of order p of a random variable. We are now ready to establish:

Proposition 2.5 *Given two random variables ξ and λ, the inequality*

$$\mathbf{P}\{\lambda \leq m_p(\lambda) + h\} \geq \mathbf{P}\{\xi \leq m_p(\xi) + h\} \tag{2.4}$$

holds for all $p \in (0,1)$ and $h > 0$, if and only if there exists a Lipschitz, non–decreasing map (a contraction) U from $\mathbf{R}$ to $\mathbf{R}$ such that λ and $U(\xi)$ are identically distributed.

Proof. Assume that (2.4) is fulfilled, that is, assume that (2.1) is fulfilled for F and G the respective distribution function of ξ and λ. By Lemma 2.4, the map $U = G^{-1}(F)$ restricted to $S = S(F) \backslash \{+\infty\}$ transforms the distribution of ξ into the distribution of λ (recall that $\mathbf{P}\{\xi \in S\} = 1$). U is non–decreasing and, according to (2.3), U is finite and Lipschitz on S, with Lipschitz constant $K \leq 1$. By the Kirszbraun–McShane (Hahn–Banach) theorem, U can be extended to the whole real line, without changing the Lipschitz constant K. But it is clear that, on the real line, such an extension can be also chosen non–decreasing. Indeed, in a unique way, and by continuity, U extends to $clos(S)$, so U is Lipschitz and non–decreasing on $clos(S)$. The complement $T = \mathbf{R} \setminus clos(S)$ is open and therefore is the union of at most countably many disjoint open intervals. If (a, b) is a finite interval from this decomposition of T, define U on (a, b) linearly so that $U(a^+) = U(a)$, $U(b^-) = U(b)$. If (a, b) is infinite, say with, $b = +\infty$, put $U(x) = U(a) + (x - a)$, for all $x > a$ (U is defined in a similar way when $a = -\infty$). Clearly, this extension of U is a Lipschitz, non–decreasing function on the whole real line. The proof of the converse statement is elementary.

Finally, we prove:

Proposition 2.6 *Let μ be a probability measure on the real line $\mathbf{R}$, such that* (1.1) *attains its minimum at the intervals $A = (-\infty, x]$, for all $p \in (0, 1)$ and $h > 0$. Then, μ is symmetric around its median and has finite exponential moment.*

Proof. First, write (1.2) for the minimal intervals $A = [a, +\infty)$, $B = (-\infty, x]$ of measure p, and get (2.4) for the random variable ξ with law μ and for $\lambda = -\xi$. Therefore, by Proposition 2.5, for some non–decreasing Lipschitz function U, λ and $U(\xi)$ are identically distributed. Hence, λ and $V(\lambda)$ are identically distributed, where $V(x) = U(-x)$ is also Lipschitz function. Let λ' be an independent copy of λ. Since $|V(\lambda') - V(\lambda)| \leq |\lambda' - \lambda|$, and since sides of this inequality are random variables with the same distribution, we get $|V(x) - V(y)| = |x - y|$, for almost all (x, y) with respect to $\nu \otimes \nu$, where ν is law of λ. Now, by Fubini's theorem, for some point y_0, $|V(x) - V(y_0)| = |x - y_0|$, for ν–almost all x. Therefore, since V is non–increasing, for some real a, $V(x) = -x - 2a$, for $\nu-$almost all x. That is, the distributions of $\lambda + a$ and so $\xi - a$ are symmetric around 0.

To prove the exponential integrability, assume that μ is non–degenerate, symmetric around 0, and assume that the value

$$R_h(p) = \inf_{\mu(A) \geq p} \mu(A^h), \quad 0 < p < 1, \ h > 0, \tag{2.5}$$

is attained at the interval $A = (-\infty, x]$, where $x = F^{-1}(p)$, and where F is the distribution function of μ. Since $A^h = (-\infty, x + h)$, we have $R_h(p) = F(F^{-1}(p) + h - 0)$. Now, we show that for all $h > 0$, $0 < p, q < 1$, such that $p + q < 1$,

$$R_h(p+q) \leq R_h(p) + R_h(q). \tag{2.6}$$

Indeed, let $A = (-\infty, x]$ be the extremal set in (2.5) for p, and since μ is symmetric, one can take a set $B = [y, \infty)$ as extremal for q. The assumption $p + q < 1$ implies $x \leq y$. The case $x = y$ is possible, but then $A \cup B = \mathbf{R}$, so

$$R_h(p) + R_h(q) = \mu(A^h) + \mu(B^h) \geq \mu(A) + \mu(B) \geq 1,$$

and so (2.6) is fulfilled. In case $x < y$, $A \cup B$ has measure $p + q$, and via the identity $(A \cup B)^h = A^h \cup B^h$, we obtain

$$R_h(p+q) \leq \mu((A \cup B)^h) \leq \mu(A^h) + \mu(B^h) = R_h(p) + R_h(q).$$

Next, we show that $\liminf_{p \to 0^+} R_h(p)/p > 1$, for all $h > 0$ large enough. Indeed, assume that, given $h > 0$, this $\liminf = 1$. Then, for any $\varepsilon > 0$, the set E_ε of all the points $p \in (0,1)$ satisfying $R_h(p) \leq (1+\varepsilon)p$ is infinite, and $0 \in clos(E_\varepsilon)$. Therefore, for any $p \in (0,1)$, one can choose a sequence $p_n \in E_\varepsilon$ (it is possible for some elements of this sequence to coincide) such that $r_n = p_1 + \cdots + p_n \to p$, as $n \to \infty$. Applying (2.6) to r_n, we get

$$R_h(r_n) \leq R_h(p_1) + \cdots + R_h(p_n) \leq (1+\varepsilon)(p_1 + \cdots + p_n) \leq (1+\varepsilon)p.$$

Letting $n \to \infty$, and using the left–continuity of the function R_h (via the last statement of Lemma 2.1), we obtain $R_h(p) \leq (1+\varepsilon)p$, for all p. Since $\varepsilon > 0$ is arbitrary, $R_h(p) \leq p$, hence $R_h(p) = p$, for all $p \in (0,1)$. But for h large enough, this is clearly impossible. Indeed, since μ is non–degenerate, taking $x, y \in \mathbf{R}$ such that $0 < F(x) < F(y)$, and then $R_h(p) > p$, for $p = F(x)$, $h > y - F^{-1}(p)$ leads to a contradiction.

Therefore, one can find $h > 0$, $p_0 \in (0,1)$, $c > 1$, such that $R_h(p) = F(F^{-1}(p) + h - 0) \geq cp$, for all $p \in (0, p_0]$ (of course, necessarily $cp_0 < 1$). Thus, $F(F^{-1}(p) + 2h) \geq cp$, that is, $F^{-1}(cp) - F^{-1}(p) \leq 2h$. In particular, $F^{-1}(c^k p) - F^{-1}(c^{k-1}p) \leq 2h$, for all $k = 1, \cdots, n$, if $c^{n-1}p \leq p_0$. Summing over k gives $F^{-1}(c^n p) - F^{-1}(p) \leq 2nh$. Applying this inequality to $p = p_0 c^{-n}$ gives

$$F^{-1}(p_0 c^{-n}) \geq -2nh + F^{-1}(p_0).$$

This easily implies that $F(x) \geq \exp(ax)$, for some $a > 0$, as $x \to -\infty$.

References

[1] Bobkov, S.G. Extremal properties of half-spaces for log–concave distributions. To appear in: *Ann. Probab.*

[2] Bobkov, S.G., Houdré, C. Some connections between Sobolev-type inequalities and isoperimetry. Preprint (1995).

[3] Borell, C. The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* 30 (1975), 207–211.

[4] Ehrhard, A. Symétrisation dans l'espace de Gauss. *Math. Scand.* 53 (1983), 281–301.

[5] Landau, H.J., Shepp, L.A. On the supremum of a Gaussian process. *Sankhyá Ser.A* 32 (1970), 369–378.

[6] Ledoux, M. Isoperimetry and Gaussian Analysis. *École d'été de Probabilités de Saint-Flour.* Preprint (1994).

[7] Sudakov, V.N., Tsirel'son, B.S. Extremal properties of half–spaces for spherically invariant measures. *J. Soviet Math.* 9 (1978), 9–18. Translated from: *Zap. Nauch. Sem. L.O.M.I.* 41 (1974), 14–24 (In Russian).

[8] Talagrand, M. Concentration of measure and isoperimetric inequalities in product spaces. Preprint (1994).

Sergei. G. Bobkov
Department of Mathematics
Syktyvkar University
167001 Syktyvkar
Russia

Christian Houdré
Center for Applied Probability
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30322, USA