On the Cases of Equality in Bobkov's Inequality and Gaussian Rearrangement

E. A. Carlen* and C. Kerce*

School of Mathematics
Georgia Institute of Technology
Atlanta, GA, 30332 U.S.A.

Abstract We determine all of the cases of equality in a recent inequality of Bobkov that implies the isoperimetric inequality on Gauss space. As an application we determine all of the cases of equality in the Gauss space analog of the Faber-Krahn inequality.

Key words: Isoperimetric inequality, rearrangements, Gauss space. Mathematics Classification Numbers: 49N60, 47H19, 26D10.

* carlen@math.gatech.edu, Work partially supported by U.S. N.S.F. grant DMS 92–07703

* jkerce@math.gatech.edu, Work partially supported by U.S. N.S.F. grant DMS 92–07703

Introduction

This paper concerns the cases of equality in several inequalities on Gauss space, which is $\mathbb{R}^n$ equipped with the Gaussian measure γ_n defined by

$$\gamma_n(A) = \int_A \Pi_{j=1}^n \phi(x_j) \mathrm{d}^n x \tag{1.1}$$

where

$$\phi(t) = (2\pi)^{-1/2} e^{-t^2/2} \ . \tag{1.2}$$

We begin by fixing some notation that shall be used throughout the paper. Let $\Phi(t)$ denote the so-called error function:

$$\Phi(t) = \int_{-\infty}^t \phi(x) \mathrm{d}x \ . \tag{1.3}$$

This is a strictly increasing function mapping $\mathbb{R}$ onto $(0,1)$, and hence it has a well defined inverse Φ^{-1} mapping $(0,1)$ onto $\mathbb{R}$. Finally, define the function $I(t)$ by

$$I(t) = \phi\left(\Phi^{-1}(t)\right) \ . \tag{1.4}$$

Notice that

$$\lim_{t\to 0} I(t) = \lim_{t\to 1} I(t) = 0 \ ,$$

and so we can smoothly extend the domain of definition of I to $[0,1]$ by setting $I(0) = I(1) = 0$. Let $\mathcal{B}$ denote the set of all measurable functions on $\mathbb{R}^n$ with values in $[0,1]$.

Bobkov recently proved [2] that for all $f \in \mathcal{B}$,

$$I\left(\int f \mathrm{d}\gamma_n\right) \leq \int \sqrt{I(f)^2 + |\nabla f|^2} \mathrm{d}\gamma_n \ . \tag{1.5}$$

The main result of the present paper is the determination of the cases of equality in Bobkov's inequality. Before stating this result, it is necessary to explain the meaning of the functional on the right hand side of (1.5) for general elements of $\mathcal{B}$.

For this purpose we introduce the Mehler semigroup of contractions P_t on $L^\infty(\mathbb{R}^n)$ which is defined by

$$P_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + (1-e^{-2t})^{1/2} y) \mathrm{d}\gamma_n(y) \ . \tag{1.6}$$

Ledoux [8] gave a proof of (1.5), to which we shall return below, that was based on showing that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \sqrt{(I(P_tf)^2+|\nabla P_tf|^2}\mathrm{d}\gamma_n \leq 0\ . \tag{1.7}$$

Hence we can define a functional $J(f)$ on $\mathcal{B}$ by

$$J(f)=\lim_{t\to 0}\int \sqrt{I(P_tf)^2+|\nabla P_tf|^2}\mathrm{d}\gamma_n\ . \tag{1.8}$$

Throughout this paper, $J(f)$ denotes the functional defined in (1.8). It is easy to see that when f has a distributional gradient ∇f such that $|\nabla f|$ is integrable with respect to γ_n, then

$$J(f)=\int \sqrt{I(f)^2+|\nabla f|^2}\mathrm{d}\gamma_n\ .$$

We shall show this in section 3, where in fact we show that $J(f)$ is finite if and only if f is a bounded variation function of Gauss space. In any case, now that we have defined $J(f)$ for all $f\in\mathcal{B}$, we can state our main result:

Theorem 1 *Let $f\in\mathcal{B}$ be such that*

$$I\left(\int f\mathrm{d}\gamma_n\right)=J(f)\ .$$

then either

$$f(x)=\Phi(\ell\cdot x+b)$$

for some $\ell\in\mathbb{R}^n$ and $b\in\mathbb{R}$; or else, f is the indicator function of a half space H – i.e., a set of the form

$$H=\{\ x|x\cdot\ell<b\ \}$$

for some $\ell\in\mathbb{R}^n$ and $b\in\mathbb{R}$. Here, $\ell\cdot x$ denotes the standard inner product on $\mathbb{R}^n$.

As Bobkov observed, (1.5) implies the isoperimetric inequality on Gauss space. Formally, this is easy to see: If $f=1_A$ is the indicator function of a measuareable set A then $I(f(x))=0$ for all x so that

$$\int \sqrt{I(1_A)^2+|\nabla 1_A|^2}\mathrm{d}\gamma_n=\int|\nabla 1_A|\mathrm{d}\gamma_n\ .$$

When the boundary of A, ∂A, is sufficiently nice, it is clear that

$$\int |\nabla 1_A| \mathrm{d}\gamma_n = \int_{\partial A} (2\pi)^{-d/2} e^{-|x|^2/2} \mathrm{d}\mathcal{H}^{n-1}$$

where $\mathcal{H}^{n-1}$ denotes the $n-1$ dimensional Hausdorff measure. To deal with this issue more precisely, we define the *gaussian perimeter* $\mathrm{per}_{\gamma_n}(A)$ of a set A to be

$$\mathrm{per}_{\gamma_n}(A) = \sup\{ \int_A \mathrm{div}_{\gamma_n}(v(x)) \mathrm{d}\gamma_n \mid v \in \mathcal{S}(I\!R^n, I\!R^n), |v(x)| \leq 1 \text{ for all } x \} \tag{1.9}$$

where the *gaussian divergence* $\mathrm{div}_{\gamma_n}(v(x))$ is given by

$$\mathrm{div}_{\gamma_n}(v(x)) = (\nabla - x) \cdot v(x)$$

so that it is adjoint to the negation of the usual gradient when one integrates by parts with respect to γ_n. $\mathcal{S}(I\!R^n, I\!R^n)$ is the set of all functions which are infinitly differentiable and have the property that their partial derivatives of any order vanish at infinity. This is the direct analog of the Lebesgue measure notion of perimeter for sets in $I\!R^n$ whose indicator functions have bounded variation [11, Ch 5], [5, Ch 5], and in these terms we can formulate a theorem on the cases of equality in the gaussian isoperimetric inequality:

Theorem 2 *Let 1_A be the indicator function of a measureable set A in $I\!R^n$ with $0 < \gamma_n(A) < 1$. Then*

$$J(1_A) = \mathrm{per}_{\gamma_n}(A)$$

and hence,

$$I(\gamma_n(A)) = \mathrm{per}_{\gamma_n}(A)$$

if and only if A is a half-space.

This determination of the cases of equality in the isoperimetric inequality on Gauss space leads to the determination of the cases of equality in a Gauss space analog of the Faber-Krahn inequality. We first recall some definitions. For any measurable function f on $I\!R^n$, define

$$f^*(x) = \sup\{t \mid \gamma_n(\{ y \mid f(y) > t\}) > \Phi(x_1) \} \ . \tag{1.10}$$

Then f^* has the same distribution function as f, but all of its level sets are half-spaces perpendicular to the first coordinate axis in $I\!R^n$, and it is called the *gaussian rearrangement*

or the *gaussian symmetrization* of f. For this rearrangement, it has been shown by Ehrhard [4] that the following inequality holds: For all strictly increasing convex functions Ψ on $\mathbb{R}_+$,

$$\int \Psi(|\nabla f^*|)\mathrm{d}\gamma_n \leq \int \Psi(|\nabla f|)\mathrm{d}\gamma_n \tag{1.11}$$.

This is the Gauss space analog of the Faber-Krahn inequality[6][7], which refers to spherically symmetric decreasing rearrangements and Lebesgue measure. Specifically, if g is a non-negative function on $\mathbb{R}^n$, then its spherically symmetric decreasing rearrangement g° is defined by

$$g^\circ(x) = \sup\{t \mid \mathcal{L}^n(\{\ y \mid f(y) > t\}) > \mathcal{L}^n(B(|x|))\ \} \tag{1.12}$$

where $\mathcal{L}^n$ is n-dimesnional Lebesgue measure, and $B(r)$ is the centered ball of radius r in $\mathbb{R}^n$. The Faber-Krahn inequality itself then states that for all increasing strictly convex functions Ψ on $\mathbb{R}_+$,

$$\int \Psi(|\nabla g^\circ|)\mathrm{d}^n x \leq \int \Psi(|\nabla g|)\mathrm{d}^n x\ . \tag{1.13}$$

The cases of equality in (1.13) have been determined by Brothers and Ziemer [3]. Actually, Faber and Krahn only considered the case $\phi(x) = x^2$, and Brothers and Ziemer imposed stricter conditions on Ψ. We shall return to these matters in the final section, and now we state our final result:

Theorem 3 *Suppose that Ψ is an increasing strictly convex function on $\mathbb{R}_+$, and f is such that*

$$\int \Psi(|\nabla f^*|)\mathrm{d}\gamma_n = \int \Psi(|\nabla f|)\mathrm{d}\gamma_n\ . \tag{1.14}$$

Then $f = f^* \circ U$, where U is a rotaion on $\mathbb{R}^n$

In Gauss space, **Theorem 3** follows more readily from the appropriate isoperimetric inequality than does the analog proved by Brothers and Ziemer. They first show that for equality to obtain in (1.13), the level sets of g must be balls. However, if g has "flat spots", i.e., if the distribution function $m(t) = \mathcal{L}^n(\{g > t\})$ has discontinuities, then these balls need not have the same center. A great deal of effort in [3] goes into the determination of the precise conditions on the distribution function that force all of the level sets to have the same center.

In this case, once we show that (1.14) implies that all of the level sets of f are half spaces we are done: All of these half spaces must have the same same normal direction,

or they would not be ordered by containment, as they must be. On a formal level, the fact that the level sets must all be half spaces follows directly from Theorem 2 and the so-called "co-area formula". We give a simple, rigorous proof of this, relying in part on recent results of Almgren and Lieb [1].

The paper is organized as follows: In section 2, we prove Theorem 1. In section 3, we prove several facts about the definition of J and the space of functions that have bounded variation with respect to Gauss measure, and then, using these, prove Theorem 2. Finally, the fourth section is devoted to the proof of Theorem 3.

Acknowledgments This paper grew out of conversations with S. Bobkov at Georgia Tech and with M. Ledoux at Institute Henri Poincare, and we thank them both. We also thank the organizers of the Spring 1998 special semester "Logarithmic Sobolev Inequalities, Large Deviations and Statistical Mechanics" at I.H.P., namely G. Ben–Arous, F. Comets, L. Gross, L. Saloff–Coste and D. Stroock for inviting us to participate in a very stimulating program. Finally, we thank Michael Loss for a number of valuable conversations.

Section 2

Our starting point is Ledoux's proof of (1.5), which we now sketch. Ledoux's method of proof of (1.5) is to insert $P_t f$ into J and take a derivative with respect to t. The computations are actually a bit complicated, and in fact Ledoux only explicitly considers the one dimensional case. We begin by indicating the principle steps in the computation in the n dimensional case. First, let $K(f)$ denote the function $I(f)^2 + |\nabla f|^2$ and let L the generator of the Mehler semigroup; i.e.,

$$L = \Delta f - x \cdot \nabla f = \mathrm{div}_\gamma(\nabla f) \ . \tag{2.1}$$

Since $P_t f$, denoted f_t here after, is smooth,

$$\frac{\mathrm{d}J}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \int \sqrt{K(f_t)} \mathrm{d}\gamma_n = \int K(f_t)^{-1/2} \left(I(f_t) I'(f_t) L f_t + \nabla f_t \cdot \nabla L f_t \right) \mathrm{d}\gamma_n$$

First, it follows easily from (2.1) that

$$\nabla(Lf) = L(\nabla f) - \nabla f$$

Using this and integrating by parts yields

$$\frac{\mathrm{d}J}{\mathrm{d}t} = - \int \Big\{ \left(\frac{I\, I'}{\sqrt{K}} \right)_i (f_t)_i + \left(\frac{(f_t)_i}{\sqrt{K}} \right)_j (f_t)_{ij} + \left(\frac{|\nabla f_t|^2}{\sqrt{K}} \right) \Big\} \mathrm{d}\gamma_n \tag{2.2}$$

Here, the i and j subscripts denote partial differentiation, and repeated indices are summed. Next, let $H(f)$ denote the Hessian matrix of f; i.e., the i,j entry of $H(f)$ is f_{ij}. Denote the trace of a matrix A by by trA, and the transpose by $A^\dagger$. By direct computation,

$$\left(\frac{(f_t)_i}{\sqrt{K}}\right)_j (f_t)_{ij} = K^{-3/2}\Big\{K(f_t)\ tr(H(f_t)^2) - I\ I' tr(H(f_t)\nabla f_t \otimes \nabla f_t^\dagger) - |H(f_t)\nabla f_t|^2\Big\}\ . \tag{2.3}$$

Next, using the fact that $I''I = -1$, one finds

$$\left(\frac{I\ I'}{\sqrt{K}}\right)_i (f_t)_i = K^{-3/2}\Big\{(I')^2|\nabla f_t|^4 - I^2|\nabla f_t|^2 - |\nabla f_t|^4 - I\ I' tr(H(f_t)\nabla f_t \otimes \nabla f_t^\dagger)\Big\}. \tag{2.4}$$

Now inserting (2.3) and (2.4) into (2.2) and cancelling and gathering terms, the right hand side of (2.2) reduces to

$$-\int K^{-3/2}\left\{tr\left(\Big(I'\nabla f_t \otimes \nabla f_t^\dagger - IH(f_t)\Big)^2\right) + \big(|\nabla f_t|^2 tr(H(f_t)^2) - |H(f_t)\nabla f_t|^2\big)\right\} \mathrm{d}\gamma_n\ .$$

Note that

$$\big(|\nabla f_t|^2 tr(H(f_t)^2) - |H(f_t)\nabla f_t|^2\big) \geq 0$$

by the Schwarz inequality, and hence $J(f_t)$ is a non-increasing function of t. This is Ledoux's argument [8], but carried out directly in n variables at once, instead of one at a time.

Clearly there is equality in Bobkov's inequality if and only if for each $t > 0$, f_t satisfies

$$I'\nabla f \otimes \nabla f^\dagger - IH(f) = 0 \tag{2.5}$$

and

$$|\nabla f|^2 tr(H(f)^2) - |H(f)\nabla f|^2 = 0\ . \tag{2.6}$$

We shall show below that (2.5) imples (2.6), and so we focus on the former. It turns out that (2.5) can be simplified by a change of dependent variable.

Lemma 2.1 *Let $t > 0$ and suppose $f \in \mathcal{B}$ satisfies*

$$I\left(\int f \mathrm{d}\gamma\right) = \int \sqrt{I(f)^2 + |\nabla f|^2}\mathrm{d}\gamma,$$

Then the function h_t defined by $h_t(x) = \Phi^{-1}(P_t f)$ satisfies

$$H(h_t)(x) = 0$$

for every $x \in I\!R^d$, where $H(h_t)$ denotes the Hessian of h_t.

Proof: Note that by the definitions,

$$I(f_t) = \phi(h_t) \tag{2.7}$$

and that

$$f_t = \Phi(h_t) \ . \tag{2.8}$$

From (2.7), it follows that

$$I' \nabla f_t = -h_t \phi(h) \nabla h_t \ . \tag{2.9}$$

From (2.8), it follows that

$$\nabla f_t = \phi(h) \nabla h_t \tag{2.10}$$

and

$$H(f_t) = -h_t \phi(h_t) \nabla h_t \otimes \nabla h_t^t + \phi(h) H(h) \ . \tag{2.11}$$

Using these identities, one readily sees that f_t satisfies (2.5) if and only if $\phi(h_t)^2 H(h_t) = 0$, which is the case if and only if $H(h_t) = 0$ everywhere. ■

We now prove the main result:

Proof of Theorem 1: From Lemma 2.1, we know that for each $t > 0$ that $h_t(x)$ is a linear function, and hence, for each t, there exist an $\ell(t)$ in $I\!R^n$ and a $b(t)$ in $I\!R$ such that

$$h_t(x) = \ell(t) \cdot x + b(t) \ .$$

Now fix any $t_0 > 0$, and let $\ell = \ell(t_0)$ and $b = b(t_0)$. Then by the definition of h_t, $P_{t_0} f(x) = \Phi(\ell \cdot x + b)$. Now it is clear from (1.6) that for any unit vector u and any number c, if one defines $g(x) = 1_{(-\infty, c]}(u \cdot x)$ then one has

$$P_t g(x) = \Phi((1 - e^{-2t})^{-1/2} c - (e^{2t} - 1)^{-1/2} u \cdot x) \ .$$

Now suppose first that $(e^{2t_0}-1)^{-1/2} \geq |\ell|$, and then define $s \geq 0$ by $(e^{2(t_0+s)}-1)^{-1/2} = |\ell|$, and define c by $c = b(1-e^{-2(t_0+s)})^{1/2}$. Then clearly with $u = |\ell|^{-1}\ell$,

$$P_{t_0} f = P_{t_0+s} g$$

for this choice of s and c. This implies that $f = P_s g$, proving the theorem in this case.

Next, suppose that

$$(e^{2t_0}-1)^{-1/2} < |\ell| \ . \tag{2.12}$$

Define s by $(e^{2s}-1)^{-1/2} = |\ell|$, and define c by $c = b(1-e^{-2s})^{1/2}$. Then clearly with $u = |\ell|^{-1}\ell$, $P_{t_0} f = P_s g$ for this choice of s and c. Notice that $s < t_0$, and therefore $P_{t_0-s} f = g$. This, however, is impossible since $P_{t_0-s} f$ is smooth for al f in $\mathcal{B}$ and g is discontinuous. Hence (2.12) is impossible. ■

Section 3

We now turn to the proof of Theorem 2, which follows readily from the facts that the operator P_t is self adjoint in L_2, and that

$$P_t \ \mathrm{div}_{\gamma_n}(v) = e^{-t} \mathrm{div}_{\gamma_n}(P_t \nu) \tag{3.1}$$

for $v \in \mathcal{S}(I\!R^n, I\!R^n)$. This in turn follows easily from (1.6).

We begin by defining a Gauss space analog of the BV (semi)norm: Let $||Df||_{\gamma_n}$ be defined by

$$||Df||_{\gamma_n} = \sup\{\int f \ \mathrm{div}_{\gamma_n}(v(x)) \ \mathrm{d}\gamma_n \mid v \in \mathcal{S}(I\!R^n, I\!R^n), \ |v(x)| \leq 1 \text{ for all } x\}.$$

This is the direct analog of the usual BV norm discussed in [11]. It is shown in Proposition A.1 that if f has a distributional gradient in the usual sense, then

$$||Df||_{\gamma_n} = \int |\nabla f| \mathrm{d}\gamma_n \ .$$

Here what we require for the proof of Theorem 2 is the following:

Lemma 3.1 *For any function $f \in BV(I\!R^n, \gamma_n)$,*

$$\lim_{t \to 0} \int |\nabla f_t| \ \mathrm{d}\gamma_n = ||Df||_{\gamma_n} \ .$$

Proof: Notice that

$$\int |\nabla f_t| \, \mathrm{d}\gamma_n = \sup_{|v|\leq 1} \int (P_t f) \, \mathrm{div}_{\gamma_n}(v) \, \mathrm{d}\gamma_n = \sup_{|v|\leq 1} \int f \mathrm{div}_{\gamma_n}(P_t v) \, \mathrm{d}\gamma_n .$$

Let $L = \sup_{|v|\leq 1} \int f \, \mathrm{div}_{\gamma_n}(v) \, \mathrm{d}\gamma_n$. If $L = \infty$, then $J(f) = \infty$, and we are done, so suppose $L < \infty$. Note that if $|v(x)| \leq 1$ for all x, then also $|P_t v(x)| \leq 1$ for all x. Hence

$$\sup_{|v|\leq 1} \int f \mathrm{div}_{\gamma_n}(P_t v) \mathrm{d}\gamma_n \leq L \; . \tag{3.2}$$

Next, given $\epsilon > 0$, there exists a $v_\epsilon \in \mathcal{S}(I\!R^n, I\!R^n)$ such that

$$\int f \; \mathrm{div}_{\gamma_n}(v_\epsilon) \; \mathrm{d}\gamma_n + \epsilon > L \; .$$

Moreover, since the Mehler semigroup is strongly continuous on $L^1(\gamma_n)$, it follows from (3.1) that there exists a time t_ϵ such that for $0 < t < t_\epsilon$,

$$\Big| \int f \; \mathrm{div}_{\gamma_n}(P_t v_\epsilon) \; \mathrm{d}\gamma_n - \int f \; \mathrm{div}_{\gamma_n}(v_\epsilon) \; \mathrm{d}\gamma_n \Big| < \epsilon .$$

This implies that for all $t \in (0, t_\epsilon)$,

$$\begin{aligned} &\int f \; \mathrm{div}_{\gamma_n}(P_t v_\epsilon) \mathrm{d}\gamma_n + 2\epsilon = \\ &\int f \; \mathrm{div}_{\gamma_n}(P_t v_\epsilon) \mathrm{d}\gamma_n - \int f \; \mathrm{div}_{\gamma_n}(v_\epsilon) \mathrm{d}\gamma_n + \int f \; \mathrm{div}_{\gamma_n}(v_\epsilon) \mathrm{d}\gamma_n + 2\epsilon > \\ &\int f \mathrm{div}_{\gamma_n}(v_\epsilon) \; \mathrm{d}\gamma_n + \epsilon > L \end{aligned}$$

This computation proves that

$$\limsup_{t\to 0} \sup_{|v|\leq 1} \int f_t \; \mathrm{div}_{\gamma_n}(v) \; \mathrm{d}\gamma_n \geq L$$

which combined with (3.2) yields the result. ■

Proof of Theorem 2: Note the obvious bound, true for all $f \in \mathcal{B}$,

$$|\nabla f_t(x)| \leq \sqrt{I(f_t(x))^2 + |\nabla f_t(x)|^2} \leq I(f_t(x)) + |\nabla f_t(x)| \tag{3.3}$$

and recall that if f is the indicator function of some set, then $\lim_{t\to 0} \int I(f_t)\ \mathrm{d}\gamma_n = 0$.

Therefore, Combining Lemma 3.1, the integrated form of (3.3), and the definition of per_{γ_n}, one finds that $J(f) = \mathrm{per}_{\gamma_n}(\{f > 0\})$ whenever f is the indicator function of a set. This fact in conjunction with Theorem 1 proves Theorem 2. ■

Notice also that Lemma 3.1 combined with (3.3) shows that $J(f) < \infty$ if and only if $f \in BV(I\!R^n, \gamma_n)$.

As a final point we observe that

$$J(f) = \int \sqrt{I(f)^2 + |\nabla f|^2}\ \mathrm{d}\gamma_n \tag{3.4}$$

whenever $f \in \mathcal{B} \cap W^{1,1}(I\!R^n, \gamma_n)$; i.e., wheniver $f \in \mathcal{B}$ and f has an integrable distributional gradient. This follows from (3.3) and dominated convergence since both $\lim_{t\to 0} |\nabla f_t(x)| = |\nabla f(x)|$ and $\lim_{t\to 0} f_t(x) = f(x)$for a.e. x, as one easily sees from (1.6).

Section 4

In this section we present the proof of Theorem 3 with the intent of clarifying how the rigorous transition from an isoperimetric inequality to a rearrangement inequality is made. On a formal level, the basic idea is well known and is used in the papers of both Faber and Krahn: To study functionals of f, use coordinates based on f itself. That is, use f as a coordinate function, and then parameterize its level surfaces to obtain the remaining coordinates. Where f is C^1 with a non-vanishing gradient, this coordinate transformation is non-singular, and the Jacobian is easily worked out: If we call dh the distance from a point x on the level surface $\{f = t\}$ to the level surface $\{f = t + dt\}$, then

$$t + dt = f(x + \hat{n}\ dh) \approx f(x) + |\nabla f(x)|\ dh.$$

Recalling that $f(x) = t$, we get $dt = |\nabla f|\ dh$. This leads to the change–of–variables formula commonly called the "co-area formula":

$$\mathrm{d}^n x = |\nabla f|^{-1} \mathrm{d}\mathcal{H}^{n-1} dt\ . \tag{4.1}$$

where $\mathcal{H}^{n-1}$ denotes the $n-1$ dimensional Hausdorff measure on the level surface $\{f = t\}$. Let $\mathcal{H}_\gamma^{n-1}$ denote the $n-1$ dimensional Hausdorff measure computed using the Gaussian measure $\mathrm{d}\gamma_n$ as the reference measure. Then in the same formal way, one has

$$\mathrm{d}\gamma_n = |\nabla f|^{-1} \mathrm{d}\mathcal{H}_\gamma^{n-1} dt \ . \tag{4.2}$$

When we talk of surface measure in this section, we will continue to use the Hausdorff measure in place of per_{γ_n}. It can be shown that $\mathrm{per}_\gamma(A)$ and $\mathcal{H}_\gamma^{n-1}(\partial A)$ agree for certain types of sets $A \subset I\!R^n$, for more detail see [5, Ch 5.8]. Here we are interested in sets which arise from the distribution of functions which satisfy the pointwise convention (4.12) given below. For these functions we show in the appendix that $\mathrm{per}_\gamma(\{f > t\}) = \mathcal{H}_\gamma^{n-1}(\{f = t\})$ for almost every t.

We generalize Krahn's argument [7] to give a formal proof of (1.11), which will be made precise in the second half of this section. For any convex increasing function Ψ on $I\!R_+$, define

$$F_t^{(\Psi)}(f) = \int_{\{f>t\}} \Psi(|\nabla f|) \ \mathrm{d}\gamma_n \ . \tag{4.3}$$

In particular, $F_t^{(1)}(f)$ is the distribution function of f; i.e.,

$$F_t^{(1)}(f) = \int_{\{f>t\}} 1 \ \mathrm{d}\gamma_n = \gamma_n(\{f > t\}) \ .$$

Next, define the quantity

$$\mathcal{N}_t^{(\Psi)}(f) = \int_{\{f=t\}} \Psi(|\nabla f|) \ |\nabla f|^{-1} \mathrm{d}\mathcal{H}_\gamma^{n-1}$$

and notice that by a formal application of (4.2),

$$\mathcal{N}_t^{(\Psi)}(f) = -\frac{\mathrm{d}}{\mathrm{d}t} F_t^{(\Psi)}(f) \ . \tag{4.4}$$

Due to the definition of f^*, we have that $F_t^{(1)}(f) = F_t^{(1)}(f^*)$. Formally differentiating this, we have the identity

$$\mathcal{N}_t^{(1)}(f) = \mathcal{N}_t^{(1)}(f^*) \ . \tag{4.5}$$

Using the fact that $|\nabla f^*|$ is constant on $\{f^* = t\}$, one further deduces

$$\mathcal{N}_t^{(1)}(f^*) = |\nabla f^*|^{-1} \mathcal{H}_{\gamma_n}^{n-1}(\{f^* = t\}). \tag{4.6}$$

Now with the preliminaries done, observe that at least where $|\nabla f|$ does not vanish,

$$\mathcal{H}_\gamma^{n-1}(\{f=t\}) = \int_{\{f=t\}} |\nabla f||\nabla f|^{-1} \mathrm{d}\mathcal{H}_\gamma^{n-1} \tag{4.7}$$

and hence

$$\frac{1}{\mathcal{N}_t^{(1)}(f)} \mathcal{H}_\gamma^{n-1}(\{f=t\}) = \frac{1}{\mathcal{N}_t^{(1)}(f)} \int_{\{f=t\}} |\nabla f| \, |\nabla f|^{-1} \mathrm{d}\mathcal{H}_\gamma^{n-1} \ .$$

By Jensen's inequality

$$\Psi\left(\frac{1}{\mathcal{N}_t^{(1)}(f)} \mathcal{H}^{n-1}(\{f=t\})\right) \leq \frac{1}{\mathcal{N}_t^{(1)}(f)} \int_{\{f=t\}} \Psi(|\nabla f|)|\nabla f|^{-1} \mathrm{d}\mathcal{H}^{n-1} = \frac{\mathcal{N}_t^{(\Psi)}(f)}{\mathcal{N}_t^{(1)}(f)} \tag{4.8}$$

or finally,

$$\mathcal{N}_t^{(\Psi)}(f) \geq \mathcal{N}_t^{(1)}(f) \Psi\left(\frac{\mathcal{H}_\gamma^{n-1}(\{f=t\})}{\mathcal{N}_t^{(1)}(f)}\right). \tag{4.9}$$

Making use of (4.5) one uses the isoperimetric inequality and the fact that $\Psi(x)$ is increasing to deduce that

$$\mathcal{N}_t^{(1)}(f) \Psi\left(\frac{\mathcal{H}_\gamma^{n-1}(\{f=t\})}{\mathcal{N}_t^{(1)}(f)}\right) \geq \mathcal{N}_t^{(1)}(f^*) \Psi\left(\frac{\mathcal{H}_\gamma^{n-1}(\{f^*=t\})}{\mathcal{N}_t^{(1)}(f^*)}\right) \ . \tag{4.10}$$

This final term can be rewritten $\mathcal{N}^{(\Psi)}(f^*)$ since $|\nabla f^*|$ is constant on the set $\{f^*=t\}$.

Now combine (4.4) and (4.10) to get

$$\int \Psi(|\nabla f|)\mathrm{d}\gamma_n = \int_0^\infty \mathcal{N}_t^{(\Psi)}(f)\mathrm{d}t \geq \int_0^\infty \mathcal{N}_t^{(\Psi)}(f^*)\mathrm{d}t = \int \Psi(|\nabla f^*|)\mathrm{d}\gamma_n \ .$$

Moreover, given that there is equality in the Gaussian isoperimetric inequality if and only if $\{f>t\}$ is a half space, the inequality (4.10) is strict for each t such that $\{f>t\}$ is not a half space. Hence (1.11) holds, and the is equality if and only if each $\{f>t\}$ is a half space. However, since we must have $\{f>t\} \subset \{f>s\}$ for all $s<t$, all of these half spaces must be parallel, and hence we must have that $f = f^* \circ U$ for some rotation U.

On a formal level, this is the proof of Theorem 3. However, serious problems are caused by the set of points on which $|\nabla f| = 0$. The formal argument we have given is correct for for what Almgren and Lieb [1] have called "co–area regular" functions, but not in general.

The formal identity (4.5) is not true in general. We show below in Lemma 4.3, that (4.5) is replaced by an inequality, and fortunately the inequality works with the isoperimetric inequality instead of against it. Lemma 4.3 is the main result we need to render our formal argument rigorous. Before coming to this, there are several other technical issues to be dealt with.

We begin with a rigorous integrated version [1] of the co-area formula for Lebesgue measure $\mathcal{L}$ for integrable functions f that have an integrable distributional gradient:

$$\mathcal{L}^n(\{f > t\}) \;=\; \mathcal{L}^n(\{f > t\} \cap \{|\nabla f| = 0\}) \;+\; \int_{s>t} \int_{\{f=s\}} |\nabla f|^{-1} \mathrm{d}\mathcal{H}^{n-1} ds. \qquad (4.11)$$

The Lebesgue measure analog of (4.4) for $\Psi = 1$ would hold if the first term on the right in (4.11) would have zero derivative almost everywhere with respect to t. It is this set of functions which Almgren and Lieb [1] term "co-area regular", and they prove that in one dimension, all functions are co-area regular [1, Theroem 5.4]. However, in higher dimension, this is not the case.

In writing such sets as $\{f = t\}$, we are following the convention of Brothers and Ziemer in the choice of the representative of f: Let $B(x,r)$ be the ball of radius r centered at x, and denote by $\mu(x)$ and $\lambda(x)$ the upper and lower approximate limits of f at x:

$$\mu(x) = \inf\{t \mid \limsup_{r\to 0} \frac{\mathcal{L}^n(\ \{f > t\} \cap B(x,r)\ )}{\mathcal{L}^n(B(x,r))} = 0\}$$

$$\lambda(x) = \sup\{t \mid \limsup_{r\to 0} \frac{\mathcal{L}^n(\ \{f < t\} \cap B(x,r)\ )}{\mathcal{L}^n(B(x,r))} = 0\}.$$

We the define f pointwise by

$$f(x) = \frac{\mu(x) + \lambda(x)}{2}. \qquad (4.12)$$

We need the Gaussian analog of (4.11), which we shall derive from the following form of the co-area formula for Lebesgue measure [3,p. 156]:

Proposition *If $f \in W^{1,p}(\mathbb{R}^n)$, $1 \le p \le \infty$, f is compactly supported, and u is a nonnegative Borel function, then with the above convention (4.12) in force,*

$$\int_{\mathbb{R}^n} |\nabla f| u = \int_{s\ge 0} \int_{\{f=s\}} u \mathrm{d}\mathcal{H}^{n-1} ds,$$

where on the left side the convention $0 \cdot \infty = 0$ is employed.

This is easily adapted to Gauss space. Note that $W^{1,p}(I\!R^n, \gamma_n) \subset W^{1,1}(I\!R^n, \gamma_n)$ for $p > 1$, and hence we only consider $p = 1$ in the following.

Lemma 4.1 *Let $f \in W^{1,1}(I\!R^n, \gamma_n)$, and $u \in L_\infty(I\!R^n, \gamma_n)$. With the convention of Brothers and Ziemer in force,*

$$\int_{I\!R^n} |\nabla f|\ u\ \mathrm{d}\gamma_n = \int_{s\geq 0} \int_{\{f=s\}} u\ d\mathcal{H}^{n-1}_{\gamma_n} ds, \tag{4.13}$$

where on the left side the convention $0 \cdot \infty = 0$ is employed.

Proof: Let $\psi_R(x) = 1_{B_R} + \left(\frac{R+n-|x|}{n}\right)^2 1_{B_{R+n}\setminus B_R}$, where B_R is the ball of radius R in $I\!R^n$. Then $|\nabla \psi_R| \leq 1$ for all $x \in I\!R^n$ and all $R > 0$. Let $N \in I\!N$, where $I\!N$ denotes the natural numbers, then the compactly supported functions defined by $f_N = \psi_N f$ converge to f in $W^{1,p}(I\!R^n, \gamma_n)$ as $N \to \infty$. If we suppose that u in (4.13) is compactly supported, then for N large enough we have $|\nabla f_N|\ u\ = |\nabla f|\ u$, so the theorem is true whenever u has compact support. Lemma 4.1 follows by dominated convergence, since any function $u \in L_\infty$ can be approximated by the function $u_N = u 1_{B_N}$. ■

Next, inserting for the function u the pointwise bounded function

$$u_\delta = (|\nabla f| + \delta)^{-1} 1_{\{f>t\}\cap\{|\nabla f|\neq 0\}}$$

and applying monotone convergence yields:

$$\gamma_n\big(\ \{f > t\} \cap \{|\nabla f| \neq 0\}\ \big) = \int_{\{s>t\}} \int_{\{f=s\}\cap\{|\nabla f|\neq 0\}} |\nabla f|^{-1}\ \mathrm{d}\mathcal{H}^{n-1}_{\gamma_n}\ ds. \tag{4.14}$$

Lemma 4.2 *For f as in Lemma 4.1 and almost every t,*

$$\mathcal{H}^{n-1}_{\gamma_n}(\{f = t\} \cap \{|\nabla f| = 0\}) = 0.$$

Proof: Setting $u = 1_{\{|\nabla f|=0\}}$ in (4.13) gives

$$0 = \int_{\{|\nabla f|=0\}} |\nabla f(x)| \mathrm{d}\gamma_n = \int_{\{s\geq 0\}} \int_{\{f=s\}\cap\{|\nabla f|=0\}} \mathrm{d}\mathcal{H}^{n-1}_{\gamma_n} ds.$$

So $\int_{\{s\geq 0\}} \mathcal{H}^{n-1}_{\gamma_n}(\{f=s\}\cap\{|\nabla f|=0\})ds = 0$, combined with the fact that $\mathcal{H}^{n-1}_{\gamma_n}(\{f=s\}\cap\{|\nabla f|=0\}) \geq 0$ proves the claim. ■

The upshot of Lemma 4.2 is that the set $\{f=s\}\cap\{|\nabla f|\neq 0\}$ can be replaced in (4.14) by the set $\{f=s\}$. This allows us to represent the distribution function of f as

$$\gamma_n(\{f>t\}) = \gamma_n(\{f>t\}\cap\{|\nabla f|=0\}) + \int_{\{s>t\}}\int_{\{f=s\}} |\nabla f|^{-1}\, \mathrm{d}\mathcal{H}^{n-1}_{\gamma_n}\, ds. \qquad (4.15)$$

This formula has important consequences for the formal proof of Theorem 3. The function $F^{(1)}_t(f)$ is the sum of two decreasing functions of t. Surprisingly, as demonstrated in [1], however, there are functions in two or more dimensions for which $\gamma_n(\{f>t\}\cap\{|\nabla f|=0\})$ has a classical non-zero derivative on a set of positive measure. The identity (4.5) fails for this reason. However, we have the follwoing inequality:

Lemma 4.3 *Let f be as in Lemma 4.1 and f^* be the Gaussian rearrangement of f. Then for almost every $t\geq 0$,*

$$\mathcal{N}^{(1)}_t(f) \quad\leq\quad \mathcal{N}^{(1)}_t(f^*) \quad=\quad |\nabla f^*|^{-1}\mathcal{H}^{n-1}_{\gamma_n}(\{f^*=t\}). \qquad (4.16)$$

Proof: We use Lemma 5.4 in [1], which states that all $f\in BV(I\!R)$ are co-area regular. The Gaussian rearrangement of the function f depends only on the first coordinate, so we can define the function $f^*_1(r) = f^*(r,0,\cdots,0)$. For every $t\geq 0$,

$$\gamma_n(\{f^*>t\}\cap\{|\nabla f^*|=0\}) = \gamma_1(\{f^*_1>t\}\cap\{f^{*'}_1=0\}).$$

So by the lemma in [1], the derivative of the right-hand side with respect to t is zero for almost every t. For this reason f^* satisfies

$$-\frac{d}{dt}\gamma_n(\{f^*>t\}) = |\nabla f^*|^{-1}\mathcal{H}^{n-1}_{\gamma_n}(\{f^*=t\})$$

for almost every $t\geq 0$.

On the other hand from the representation (4.15) it is easy to see that

$$-\frac{d}{dt}\gamma_n(\{f>t\}) \geq \int_{\{f=t\}} |\nabla f|^{-1}\mathrm{d}\mathcal{H}^{n-1}_{\gamma_n}$$

by simply neglecting the first term on the right hand side of the equality in (4.15). The combination of these results with the fact that $\gamma_n(\{f > t\}) = \gamma_n(\{f^* > t\})$ for all t proves (4.16). ■

Proof of Theorem 3: First, Lemma 4.2 shows that (4.7) is in fact valid for almost every t and we can arrive at the inequality (4.9) through Jensen's inequality by the argument presented before. Now instead of applying the formal identity (4.5), we apply the inequality of Lemma 4.3 using the fact that $\Psi(x)/x$ is a non-decreasing function of x. Thus one correctly arrives at (4.10).

The proof is now finished as before, and one clearly sees that there is equality only in case for almost every t, $\{f \geq t\}$ is a half space, which, as explained before, suffices to prove Theorem 3. ■

Appendix

In the previous section we used the fact that $\mathrm{per}_\gamma(\{f > t\}) = \mathcal{H}^{n-1}_\gamma(\{f = t\})$ without proof. This result can be extracted, in Lebesgue measure form, from [5] and [11]. The purpose of this appendix is to provide a road map for the extraction from the more complete refrences [5] and [11].

Definition Given a function $f \in BV(I\!R^n, \gamma)$, define the total variation of f to be

$$||Df||_{\gamma_n} = sup\{\int f \ \mathrm{div}_{\gamma_n}(\nu(x)) \ \mathrm{d}\gamma_n \mid \nu \in \mathcal{S}(I\!R^n, I\!R^n), \ |\nu(x)| \leq 1 \text{ for all } x\}.$$

In this language, we can write the gaussian perimeter of a measurable set A as $\mathrm{per}_{\gamma_n}(A) = ||D1_A||_{\gamma_n}$. For a non-negative function $u \in \mathcal{S}(I\!R^n)$ define the functional

$$||Df||_{\gamma_n}(u) = sup\{\int g \ \mathrm{div}_{\gamma_n}(\nu(x)) \ \mathrm{d}\gamma_n \mid \nu \in \mathcal{S}(I\!R^n, I\!R^n), \ |\nu(x)| \leq u(x) \text{ for all } x\}.$$

It can be shown that $||Df||_{\gamma_n}$ is a positive linear functional which is continuous under monotone convergence. Subsequently, by the Riesz Representation Theorem, $||Df||_{\gamma_n}$ defines a Radon measure on $I\!R^n$. This is done, for example, in [5, Ch 5.1] and [11, Ch 5.1]. Moreover, we have the following

Proposition A.1 *If $f \in W^{1,p}(I\!R^n, \gamma)$ for $p \geq 1$, and u is a measurable function such that $||Df||_{\gamma_n}(u) < \infty$, then $||Df||_{\gamma_n}(u)$ can be written*

$$||Df||_{\gamma_n}(u) = \int_{I\!R^n} |\nabla f| \ u \ \mathrm{d}\gamma, \tag{6.1}$$

where the right hand side is interpreted in the distributional sense.

Proof To see this, first suppose $u \in \mathcal{S}(I\!R^n)$ and notice that any function $\nu \in \mathcal{S}(I\!R^n, I\!R^n)$ for which $|\nu| \leq u$ can be rewritten $\nu = \hat{\nu} u$ with $\hat{\nu} \in \mathcal{S}(I\!R^n, I\!R^n)$. Making use of integration by parts, the triangle inequality, and the Cauchy-Shwartz inequality,

$$\Big| \int_{I\!R^n} f \ \mathrm{div}_{\gamma_n}(\nu) \ \mathrm{d}\gamma \Big| = \Big| \int_{I\!R^n} \nabla f \cdot \nu \ \mathrm{d}\gamma \Big| \leq \int_{I\!R^n} |\nabla f| \ u \ \mathrm{d}\gamma.$$

So that $||Df||_{\gamma_n}(u) \leq \int_{I\!R^n} |\nabla f| u \mathrm{d}\gamma$. To see that the equality is attained, consider a smooth sequence $\hat{\nu}_i$, $|\hat{\nu}_i| \leq 1$, which approximates $\frac{\nabla f}{|\nabla f|}$. Then $\nu_i = \hat{\nu}_i u$ is a sequence of functions which is admissible in the definition of $||Df||_{\gamma_n}(u)$ and for which equality is attained in (6.1). Finally, we can extend (6.1) to functions u which are integrable against $|\nabla f|$ by dominated convergence, since $\mathcal{S}(I\!R^n)$ is dense in $L_1(I\!R^n, \gamma)$. ■

To see that $\mathrm{per}_\gamma(\{f > t\}) = \mathcal{H}^{n-1}_\gamma(\{f = t\})$ for almost every t, turn to Lemma 4.1 and compare the equality (4.13) with the equality (6.1). From this we can see that

$$||Df||_{\gamma_n}(1_{\{f>t\}}) = \int_{s \geq t} \int_{\{f=s\}} \mathrm{d}\mathcal{H}^{n-1}_{\gamma_n} ds.$$

The final step is the connection between $||Df||_{\gamma_n}$ and per_{γ_n}, contained in the following Lemma, due to Fleming and Rishel which we modify slightly from [5, Ch 5.5].

Proposition A.2 *Let $f \in BV(I\!R^n, \gamma)$, and u be a measurable function such that $||Df||_{\gamma_n}(u) \leq \infty$, then*

$$||Df||_{\gamma_n}(u) = \int_{s \geq 0} ||D1_{\{f>s\}}||_{\gamma_n}(u) \ ds.$$

If we take $u = 1_{\{f>t\}}$ in the above propositions A.1 and A.2 and apply the co-area formula from Lemma 4.1 we get the result that for every t,

$$\int_{s>t} \mathrm{per}_\gamma(\{f > s\}) \ ds = \int_{s>t} \mathcal{H}^{n-1}_\gamma(\{f = s\}) \ ds.$$

Hence $\mathrm{per}_\gamma(\{f > t\}) = \mathcal{H}^{n-1}_\gamma(\{f = t\})$ for almost every $t \geq 0$.

The reader should note that it can happen that $\mathrm{per}_\gamma(\{f > t\}) \neq \mathcal{H}^{n-1}_\gamma(\{f = t\})$ for some values of t. For instance, consider in dimension $n = 1$ the function $f(x) = x^2$. For $t = 0$ we have $\mathrm{per}_{\gamma_1}(\{f > 0\}) = 0$, while $\mathcal{H}^{n-1}_\gamma(\{f = 0\}) = 1$.

References

[1] Almgren, Frederick J. and Lieb, Elliott H. : *Symmetric Decreasing Rearrangement is Sometimes Continuous*, Journal of the A.M.S., Vol. 2, No. 4, 1989.

[2] Bobkov, S.G. : *An isoperimetric inequality on the discrete cube, and an elementry proof of the isoperimetric inequality in Gauss space*, Annals of Prob., **25**, 206–214 1997.

[3] Brothers, John E. and Ziemer, William P. : *Minimal rearrangements of Sobolev functions*, Journal fur die reine und angewandte Mathematik, **384**, 153-179 1988.

[4] Ehrhard, Antoine : *Inégalités Isopérimetriques et Intégrales de Dirichlet Gaussiennes*, Ann. Scient. Ec. Norm. Sup. , Series 4, 17, 1984

[5] Evans, Lawrence C. and Gariepy, Ronald F. : *Measure theory and fine properties of functions*, Boca Raton: CRC Press, c1992.

[6] Faber, Georg : *Beweis, dass unter allen homogenen Membranen von gleicher Flache und gleicher Spannung die kreisformige den tiefsten Grundton gibt*, Sitzungsberichte der Bayrischen Akademie der Wissenschaften, 1923, 169-172

[7] Krahn, E. : *Uber eine von Rayleigh formulierte Minimaleigenschaft des Kreises*, Mathematische Annalen, Vol. 94, 97-100, 1924.

[8] Ledoux, Michel : *A Short Proof of the Gaussian Isoperimetric Inequality*, Progress in Pobability, Vol. 43, 1998.

[9] Mehler, F.G. : *Uber die Entwicklung einer Funktion von beliebig vielen Variablen nach Laplaceschen Funktionen*, Crelle's Journal, Vol. 66, 1866

[10] Polya and Szego : *Isoperimetric Inequalities in Mathematical Physics*, Princeton University Press, 1951.

[11] Ziemer, William P. : *Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation*, Graduate Texts in Mathathematics , 120, Springer-Verlag, 1989.