AN ISOPERIMETRIC INEQUALITY ON THE DISCRETE CUBE, 
AND AN ELEMENTARY PROOF OF THE ISOPERIMETRIC 
INEQUALITY IN GAUSS SPACE

BY S. G. BOBKOV 
Syktyvkar University

We prove an isoperimetric inequality on the discrete cube which is the 
precise analog of a logarithmic inequality due to Talagrand. As a conse-
quence, the Gaussian isoperimetric inequality is derived.

Let us consider the following inequality: for all \( 0 \leq a, b \leq 1 \),

\[
I \left( \frac{a + b}{2} \right) \leq \frac{1}{2} \sqrt{I(a)^2 + \frac{1}{2} |a - b|^2} + \frac{1}{2} \sqrt{I(b)^2 + \frac{1}{2} |a - b|^2},
\]

where \( I \) is a nonnegative function defined on \([0, 1]\) such that

\[
I(0) = I(1) = 0.
\]

Clearly, if several functions satisfy (1) and (2), then their supremum also 
satisfies (1) and (2). One may wonder therefore if there exists a maximal 
function among those for which (1) and (2) hold, and if so, what the maximal 
function is. This question turns out to be a key to an isoperimetric problem on 
the discrete cube. As we will see, an appropriate functional isoperimetric 
inequality contains in a limit case the well-known isoperimetric inequality in 
Gauss space. For \( x \in [-\infty, +\infty] \), set

\[
\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\{-x^2/2\}, \quad \Phi(x) = \int_{-\infty}^{x} \varphi(t) \, dt.
\]

\( \Phi \) is an increasing bijection from \([-\infty, +\infty]\) to \([0, 1]\). Let \( \Phi^{-1} : [0, 1] \to [-\infty, +\infty] \) be the inverse function.

PROPOSITION 1. The function \( I(p) = \varphi(\Phi^{-1}(p)) \), \( 0 \leq p \leq 1 \), is maximal 
among all nonnegative continuous functions satisfying (1) and (2).

This statement is proved at the end of the present note [in fact, that 
\( I = \varphi(\Phi^{-1}) \) satisfies (1) and (2) implies the maximal property]. Now let us 
rewrite (1) “on functions” as a “two point” analytic inequality. Given an
arbitrary function $f: \{-1, 1\} \rightarrow [0, 1]$, we have, putting in (1) $a = f(-1), b = f(1)$:

$$I(E f) \leq E \sqrt{I(f)^2 + |\nabla f|^2},$$

where mathematical expectations (integrals) are understood with respect to uniform measure $\mu = \frac{1}{2} \delta_{-1} + \frac{1}{2} \delta_{1}$ on $\{-1, 1\}$, and where $\nabla f$ denotes discrete gradient, that is, $|\nabla f| = |(f(1) - f(-1))/2|$. More generally, for functions $f: \{-1, 1\}^n \rightarrow \mathbb{R}$, the modulus of discrete gradient will be defined by

$$|\nabla f(x)|^2 = \frac{1}{3} \sum_{i=1}^{n} |f(x) - f(s_i(x))|^2,$$

where $s_i(x)_j = x_j$, if $j \neq i$, and $s_i(x)_i = -x_i$, if $j = i$ [i.e., $s_i(x)$ is the neighbor of $x$ in the $i$th coordinate]. We first observe a main additivity property of (3).

**Lemma 1.** Given a nonnegative function $I$ on $\{-1, 1\}$, if (3) holds for all $f: \{-1, 1\} \rightarrow [0, 1]$ with respect to a probability measure $\mu$ on $\{-1, 1\}$, then (3) holds for all $f: \{-1, 1\}^n \rightarrow [0, 1]$ with respect to the product measure $\mu_n$, the $n$th power of $\mu$.

The same statement could be made about arbitrary probability measure $\mu$ on $\mathbb{R}$ and its power $\mu^n$ for the usual gradient $\nabla f$ of locally Lipschitz functions $f$ on the Euclidean space (see, for extensions, [1]).

**Proof.** Lemma 1 is easily proved by induction over $n$. Given $f: \{-1, 1\}^{n+1} \rightarrow [0, 1]$, put $f_0(x) = f(x, -1), f_1(x) = f(x, 1)$, where $x \in \{-1, 1\}^n$. We use the notation $E_n \psi = \int \psi \, d\mu_n$. Put

$$p_0 = \mu(\{-1\}), \quad p_1 = \mu(\{1\}), \quad a_0 = E_n f_0, \quad a_1 = E_n f_1.$$

Hence, $E_{n+1} f = p_0 a_0 + p_1 a_1$. Since $|\nabla f(x, -1)|^2 = |\nabla f_0(x)|^2 + \frac{1}{3} |f_0(x) - f_1(x)|^2$ and $|\nabla f(x, 1)|^2 = |\nabla f_1(x)|^2 + \frac{1}{3} |f_0(x) - f_1(x)|^2$, one can write

$$E_{n+1} = E_{n+1} \sqrt{I(f)^2 + |\nabla f|^2}$$

$$= p_0 E_n \sqrt{I(f_0)^2 + |\nabla f_0|^2 + \frac{1}{3} |f_0 - f_1|^2}$$

$$+ p_1 E_n \sqrt{I(f_1)^2 + |\nabla f_1|^2 + \frac{1}{3} |f_0 - f_1|^2}.$$

Next, in order to estimate the right integrals in (4), we apply twice the triangle inequality

$$\int \sqrt{u^2 + v^2} \geq \sqrt{(u)^2 + (v)^2}$$

to $u_0 = \sqrt{I(f_0)^2 + |\nabla f_0|^2}$, $v = (f_0 - f_1)/2$ and to $u_1 = \sqrt{I(f_1)^2 + |\nabla f_1|^2}$, $v = (f_0 - f_1)/2$. Then, we come to

$$E_{n+1} \geq p_0 \sqrt{(E_n u_0)^2 + (E_n v)^2} + p_1 \sqrt{(E_n u_1)^2 + (E_n v)^2}.$$
By the induction assumption, $E_n u_0 \geq I(a_0), E_n u_1 \geq I(a_1)$. In addition, $E_n \mathcal{V} = (a_0 - a_1)/2$. Therefore,

\[ E_{n+1} \geq p_0 \sqrt{I(a_0)^2 + \frac{1}{2} |a_0 - a_1|^2} + p_1 \sqrt{I(a_1)^2 + \frac{1}{2} |a_0 - a_1|^2}. \]

The right-hand side of (5) is estimated, according to (3) in the case $n = 1$, by $I(p_0 a_0 + p_1 a_1) = I(E_{n+1} f)$. Lemma 1 is proved. $\square$

In case $\mu = \frac{1}{2} \delta_{-1} + \frac{1}{2} \delta_1$, $\mu_n$ represents the uniform measure on $\{-1, 1\}^n$. For any set $A \subset \{-1, 1\}^n$, we define its discrete perimeter by

\[ \mu_n^+(A) = \int |\nabla \chi_A| \, d\mu_n, \]

where $\chi_A$ denotes indicator function of the set $A$. Note that, for all $x \in \{-1, 1\}^n$,

\[ |\nabla \chi_A(x)|^2 = \frac{1}{4} \text{card}\{ i \leq n : (x \in A, s_i(x) \notin A) \text{ or } (x \notin A, s_i(x) \in A) \}. \]

Combining Proposition 1 and Lemma 1, we obtain the following statement (note that we do not use the fact that $I$ is maximal in Proposition 1).

**Proposition 2.** Let $I = \varphi(\Phi^{-1})$. Then, for all $f: \{-1, 1\}^n \to [0, 1]$,

\[ I(E f) \leq E \sqrt{I(f)^2 + |\nabla f|^2}, \]

where mathematical expectations are with respect to the uniform measure $\mu_n$.

In particular, for any $A \subset \{-1, 1\}^n$, applying (6) to $f = \chi_A$, we have

\[ \mu_n^+(A) \geq I(\mu_n(A)). \]

The function $I$ in (6) is optimal as a continuous function not depending on the dimension (since it implies the appropriate Gaussian inequality—see Corollary 1) but we do not know how optimal the inequality (7) is. For example, for the sets $A_n(p) = \left\{ x \in \{-1, 1\}^n : \frac{x_1 + \cdots + x_n}{\sqrt{n}} \leq \Phi^{-1}(p) \right\}$, we have $\mu_n(A_n(p)) \to p$ by the central limit theorem, while $\mu_n^+(A_n(p)) \to \sqrt{2} I(p)$, as $n \to \infty$, by de Moivre's local limit theorem. Hence, $A_n(p)$ are not extremal in (7) even in an asymptotic sense.

**Remark 1.** Talagrand studied in [8] the functional

\[ Mf(x) = \left[ \sum_{i=1}^n \left( f(x) - f(s_i(x)) \right)^2 \right]^{1/2} \]

and proved the inequality

\[ E Mf \geq \mathcal{Q} (E f^2 - (E f)^2), \]
where \( f : \{-1, 1\}^n \to [0, 1] \) is arbitrary, \( c \) is a universal constant and \( J(p) = p \sqrt{\log(1/p)} \). The value \( E M_{\mathcal{X}_A} \) can be viewed as "interior" perimeter. Since \( l(p) \sim p \sqrt{2 \log(1/p)} \), as \( p \to 0^+ \), and since \( M f \leq \frac{1}{2} \| \nabla f \| \), (8) implies (7) up to some universal constant in front of \( l(\mu_n(\mathcal{A})) \). In fact, the inequality

\[
(9) \quad E M_{\mathcal{X}_A} \geq c \left( \mu_n(\mathcal{A})(1 - \mu_n(\mathcal{A})) \right),
\]

which is a partial case of (8) for indicator functions \( f = \chi_A \), can be essentially better than (7) for sets \( A \) of small measure. As noted in [8], when \( A \) consists of one point, \( E M_{\mathcal{X}_A} = \sqrt{n} 2^{-n} \) while \( \mu_n(\mathcal{A}) = n 2^{-n} \). For such a set, the right-hand sides of (7) and (9) are of order \( \sqrt{n} 2^{-n} \).

Now, consider a twice differentiable function \( f : \mathbb{R}^n \to [0, 1] \) with bounded first and second partial derivatives and apply (6) to the functions

\[
f_k(x_1, \ldots, x_k) = f\left( \frac{x_1 + \cdots + x_k}{\sqrt{k}} \right), \quad x_1, \ldots, x_k \in \{-1, 1\}^n,
\]

defined on \( \{-1, 1\}^n \). (This argument, when some inequalities for Gaussian measure are derived from appropriate inequalities for Bernoulli independent random variables, is well known; see, e.g., Gross [5]). By the central limit theorem in \( \mathbb{R}^n \),

\[
\int_{\{-1, 1\}^n} f_k \, d\mu_{nk} \to \int_{\mathbb{R}^n} f \, d\gamma_n, \quad k \to \infty,
\]

where \( \gamma_n \) is the canonical Gaussian measure on \( \mathbb{R}^n \), with density \( \varphi_n(y) = \varphi(y_1) \cdots \varphi(y_n), \quad y = (y_1, \ldots, y_n) \in \mathbb{R}^n \). Note also that

\[
|\nabla f_k(x_1, \ldots, x_k)|^2 = \left| \frac{\partial f(x_1, \ldots, x_k)}{\partial x_i} \right|^2 + O\left( \frac{1}{k} \right) \quad \text{as } k \to \infty,
\]

uniformly over all \( x_1, \ldots, x_k \in \{-1, 1\}^n \), where \( \frac{\partial f}{\partial x_i} = \sum_{i=1}^n |\partial f/\partial x_i|^2 \) is continuous and bounded, again by the central limit theorem, we have

\[
\int_{\{-1, 1\}^n} \sqrt{l(f_k)^2 + |\nabla f_k|^2} \, d\mu_{nk} \to \int_{\mathbb{R}^n} \sqrt{l(f)^2 + |Df|^2} \, d\gamma_n \quad \text{as } k \to \infty.
\]

We have thus proved (6) for Gaussian measure under the above assumptions on \( f \). By a simple approximation argument, this inequality extends to all locally Lipschitz functions (which are differentiable almost everywhere by Rademacher’s theorem).

**Corollary 1.** Let \( l = \varphi(*)^{-2} \). Then, for any locally Lipschitz function \( f : \mathbb{R}^n \to [0, 1] \),

\[
(10) \quad l(E f) \leq E \sqrt{l(f)^2 + |Df|^2},
\]

where mathematical expectations are with respect to the Gaussian measure \( \gamma_n \).

In particular, for any Borel measurable set \( A \subset \mathbb{R}^n \),

\[
(11) \quad \gamma_n^+(A) \geq l(\gamma_n(A)).
\]
Here
\[ \gamma_n^+(A) = \liminf_{h \to 0} \frac{\gamma_n(A^h) - \gamma_n(A)}{h} \]
denotes Gaussian perimeter, that is, Minkowski's surface measure with respect to \( \gamma_n \); \( A^h = \{ x \in \mathbb{R}^n : |x - a| < h \text{ for some } a \in A \} \) is an open \( h \)-neighborhood of the set \( A \). Since \( \sqrt{a^2 + b^2} \leq |a| + |b| \), we have, from (10), the following corollary.

**Corollary 2.** For any locally Lipschitz function \( f: \mathbb{R}^n \to [0, 1] \),
\[ l(E f) - E l(f) \leq E |D f|. \]  
(12)

The inequality (11) easily follows from (12), as well as from (10), via approximation of the indicator functions \( f = \chi_A \) by Lipschitz functions \([\text{in the case } \gamma_n(\partial A) = 0, \text{ one may take in (12) } f_h(x) = \max(1 - (1/h)\text{dist}(A^h, x), 0), \ h > 0, \text{ and let } h \to 0\]. The last step shows that the function \( I \) is optimal in (6), or equivalently, in (1), among all continuous nonnegative functions on \([0, 1] \) satisfying (1) and (2). Indeed, if another continuous nonnegative function \( J \) on \([0, 1] \) satisfies (1) and (2), then we obtain as above the inequalities (10) and (11) for \( J \) instead of \( I \). But for half-spaces \( A \) of \( \gamma_n \)-measure \( p \), \( \gamma_n^+(A) = l(p); \text{ hence, } l(p) \geq J(p), \text{ for all } p. \)

**Remark 2.** In the same way, noting that the function \( J \) of (8) is up to a multiplicative constant equivalent to \( I \), one can deduce the inequality
\[ \gamma_n^+(A) \geq d(\gamma_n(A)) \]
with some universal constant \( c \in (0, 1) \) from Talagrand's logarithmic inequality (8). Another approach to the above inequality, based on hypercontractivity, was suggested by Ledoux ([6], Chapter 8). It is of course an interesting question how to prove this inequality with \( c = 1 \) [or its functional forms (10) and (12)] analytically. By semigroup arguments, this was recently performed by Bakry and Ledoux [1].

**Remark 3.** In the original form, the isoperimetric inequality for Gaussian measure stated that, for any Borel measurable set \( A \subset \mathbb{R}^n \), and \( h > 0 \),
\[ \gamma_n(A^h) \geq \Phi(\Phi^{-1}(\gamma_n(A)) + h) \]
(13)
(with equality at any half-space). The first proof, due to Sudakov and Tsirel'son [7] and Borell [3], was based on the isoperimetric property of balls on the sphere (a theorem by Lévy and Schmidt). Ehrhard [4] developed a rearrangement of sets argument in Gauss space \((\mathbb{R}^n, \gamma_n)\) and, as result, obtained (13). Of course, the inequality (11) represents a differential analog of (13). By considering small \( h > 0 \), (11) immediately follows from (13). Converse implication is also simple: the family of functions,
\[ R_n(p) = \Phi(\Phi^{-1}(p) + h), \quad p \in [0, 1], \ h \in \mathbb{R}, \]
possesses the property $R_{h_1 + h_2} = R_{h_1}(R_{h_2})$, and operation $h \to A^n$ possesses the property $A^{h_1 + h_2} = (A^{h_1})^{h_2}$, $h_1, h_2 \geq 0$. Therefore, if $h_1, h_2 \geq 0$ satisfy (13), then $h_1 + h_2$ also satisfies (13). Hence, (13) holds for all $h > 0$, if it holds for $h > 0$ small enough that is true by (11).

REMARK 4. With the help of (13), the inequality (12) was proved in [2]. More generally, for any Borel measurable function $f : \mathbb{R}^n \to [0, 1]$ and $h > 0$, the following holds:

\[
\mathbb{E} M_h f \geq R_h(\mathbb{E} R_{-h} f),
\]

where $M_h f(x) = \sup\{f(y) : |x - y| < h\}$. For smooth $f$, letting $h \to 0$, we have

\[
M_h f(x) = f(x) + |Df(x)|h + O(h^2),
\]

\[
R_h(\mathbb{E} R_{-h} f) = \mathbb{E} f + (I(\mathbb{E} f) - \mathbb{E} I(f))h + O(h^2),
\]

and thus (14) turns into (12). For indicator functions $f = \chi_A$, (14) becomes (13). Consequently, the inequalities (11), (12), (13) and (14) are equivalent to each other. As noted, (11) is a partial case of (10). On the other hand, if for a “good” function $f$ on $\mathbb{R}^n$ with values in $(0, 1)$, one takes $A = \{(x, y) : x \in \mathbb{R}^n, y \in \mathbb{R}, f(y) < f(x)\}$, then (10) also becomes a partial case of (11) but in $(\mathbb{R}^{n+1}, \gamma_{n+1})$. Indeed, in terms of the function $g = \Phi^{-1}(f)$, the inequality (10) reads as

\[
I(\gamma_{n+1}(A)) \leq \int_{\mathbb{R}^n} \varphi_n(x) \varphi(g(x)) \sqrt{1 + |Dg(x)|^2} \, dx.
\]

Also note that

\[
\gamma_{n+1}(A) = \int_{\partial A} \varphi_{n+1}(z) \, dH_n(z) = \int_{\partial A} \varphi_n(x) \varphi(y) \, dH_n(x, y),
\]

where $z = (x, y)$ and where $H_n$ stands for the n-dimensional Hausdorff measure in $\mathbb{R}^{n+1}$ [that is, the right integral in (16) is taken over Lebesgue surface measure on $\partial A$]. This surface is defined by equation $y = g(x)$, and $\sqrt{1 + |Dg(x)|^2} \, dx$ represents the element $dH_n(x, y)$ of measure $H_n$ at the point $(x, y) \in \partial A$; hence, the right-hand sides of (15) and (16) coincide. Therefore, the functional inequality (10) for the measure $\gamma_n$ can be written as the $(n + 1)$-dimensional isoperimetric inequality

\[
I(\gamma_{n+1}(A)) \leq \gamma_{n+1}^+(A).
\]

We are not sure that this argument is quite rigorous enough to derive (10) for dimension $n$ from (11) with $n + 1$ on the class of all smooth functions $f$, but it shows that the Gaussian isoperimetric inequality (11) is, in essence, two-dimensional: if it holds for $n = 2$, then (10) holds as above for $n = 1$; therefore, it holds for all $n$ by additivity property of (10). And, as noted, on indicator functions, (10) gives (11) for all dimensions.
PROOF OF PROPOSITION 1. As already noted, it suffices to show that \( \Phi(P^{-1}) \) satisfies (1). Fix \( c \in (0, 1) \), and introduce the function \( g(x) = l(c + x)^2 + x^2, x \in \Delta(c) = (-\min(c, 1 - c), \min(c, 1 - c)). \) If we put \( c = (a + b)/2, x = (a - b)/2, \) then (1) can be rewritten as

\[
(17) \quad \sqrt{g(0)} \leq \frac{1}{2} \sqrt{g(x)} + \frac{1}{2} \sqrt{g(-x)},
\]

and the condition \( a, b \in (0, 1) \) is equivalent to \( x \in \Delta(c). \) Multiplying by 2 and squaring (17), we get

\[
(18) \quad 4g(0) - (g(x) + g(-x)) \leq 2 \sqrt{g(x)g(-x)}.
\]

Again squaring (18) [there is no need to show that the left-hand side of (18) is nonnegative], we come to

\[
16g(0)^2 - 8g(0)(g(x) + g(-x)) + (g(x)^2 + 2g(x)g(-x) + g(-x)^2)
\]

\[
\leq 4g(x)g(-x),
\]

that is,

\[
(19) \quad 16g(0)^2 + (g(x) - g(-x))^2 \leq 8g(0)(g(x) + g(-x)).
\]

Now rewrite (19) in terms of the function \( h(x) = g(x) - g(0) = l(c + x)^2 + x^2 - l(c)^2: \)

\[
(20) \quad (h(x) - h(-x))^2 \leq 8l(c)^2(h(x) + h(-x)).
\]

**Lemma 2.** (a) \( l \cdot l'' = -1. \) (b) the function \( l'(l')^2 \) is convex on \((0, 1).\)

**Proof.** (a) follows from \( \varphi'(x) = -x\varphi(x), x \in \mathbb{R}. \) (b) \((l')^2' = 2l'l'' = -2(l'/l), \) hence, \((l')^2'' = -2(1''l - l'^2)/l^2 = 2(1 + l'^2)/l^2 \geq 0. \) \( \square \)

**Lemma 3.** The function \( R(x) = h(x) + h(-x) - 2l'(c)^2 x^2 \) is convex on \( \Delta(c). \)

**Proof.** \( R'(x) = 2l(c + x)l'(c + x) - 2l(c - x)l'(c - x) + 4x - 4l'(c)^2 x. \) Hence,

\[
R''(x) = 4 \left[ l'(c + x)^2 + l'(c - x)^2 - l'(c)^2 \right] / 2
\]

is nonnegative since \((l')^2 \) is convex [Lemma 2(b)].

Since \( R \) is even, we have from Lemma 3 that \( R(x) \geq R(0) = 0 \) for all \( x \in \Delta(c), \) therefore,

\[
h(x) + h(-x) \geq 2l(c)^2l'(c)^2 x^2.
\]
Hence, (20) will follow from the stronger inequality \( (h(x) - h(-x))^2 \leq 16I(c)l'(c)^2x^2 \), that is, from
\[
\left| \frac{h(x) - h(-x)}{x} \right| \leq 4I(c)|l'(c)|.
\]
Since \( h(x) - h(-x) = l(c + x)^2 - l(c - x)^2 \), it remains to show that
\[
\left| \frac{l(c + x)^2 - l(c - x)^2}{x} \right| \leq 4I(c)|l'(c)|.
\]
Since \( l \) is symmetric around 1/2, we have \( l(1 - c) = l(c), |l'(1 - c)| = |l'(c)| \), and
\[
\left| l((1 - c) + x)^2 - l((1 - c) - x)^2 \right| = \left| l(c - x)^2 - l(c + x)^2 \right|
\]
Therefore, one may assume \( 0 < c \leq \frac{1}{2} \). Note that \( \Delta(1 - c) = \Delta(c) \). In addition, one may assume \( x > 0 \), since the left-hand side of (21) is an even function of \( x \). Under these assumptions, \( l(c + x) \geq l(c - s) \), because \( l \) increases on \([0, \frac{1}{2}]\), decreases on \([\frac{1}{2}, 1]\) and is symmetric around \( \frac{1}{2} \). Indeed, by these properties, \( l(c + x) \geq l(c - x) \iff 1 - (c + x) \geq c - x \iff 1 \geq 2c \). Consequently, one may rewrite (21) as
\[
\frac{l(c + x)^2 - l(c - x)^2}{x} \leq 4I(c)l'(c),
\]
assuming \( 0 < x < c \leq \frac{1}{2} \). Consider the function \( u(x) = l(c + x)^2 - l(c - x)^2 \). By Lemma 2(a), \( u''(x) = 2(l'(c + x)^2 - l'(c - x)^2) \). As a convex, symmetric around \( 1/2 \) function, \( l'' \) decreases on \((0, \frac{1}{2})\) and increases on \([\frac{1}{2}, 1]\), hence, \( l'(c + x)^2 \leq l'(c - x)^2 \) and thus \( u''(x) \leq 0 \), whenever \( 0 < x < c \leq \frac{1}{2} \). Therefore, \( u \) is a concave nonnegative function on \([0, c]\). But then
\[
\frac{u(x)}{x} = \int_0^1 u'(xt) \, dt
\]
is nonincreasing on \((0, c]\), and it remains to prove (22) at \( x = 0 \). Since
\[
l(c + x)^2 = l(c)^2 + 2l(c)l'(c)x + O(x^2)
\]
as \( x \to 0 \), we have \( u(x)/x \to 4I(c)l'(c) \). Proposition 1 follows. \( \Box \)

Acknowledgments. The author would like to thank Professor M. Ledoux and Professor F. Götze for useful and stimulating discussions.

REFERENCES


DEPARTMENT OF MATHEMATICS
SYKTYVKAR UNIVERSITY
167001 SYKTYVKAR
RUSSIA