Gaussian Isoperimetric Inequality

Ruilin Li and Luning Li

1 Introduction

You have probably heard of the classical isoperimetric problem: among all closed curves in the plane with equal perimeter, which curve encloses the largest area? Equivalently, the question can be formulated as the following: among all curves in the plane enclosing the same area, which one, if any, has the smallest perimeter. The answer to this question is the circle, and this is true even in a much more general setting. Among all Borel sets in $\mathbb{R}^n$ with equal Lebesgue measure, balls have the smallest "boundary measure".

One can ask similar question about the Gaussian measure in $\mathbb{R}^n$, which leads to the Gaussian isoperimetric inequality. This time, the solution is not balls but half-spaces.

Before formally stating and proving the Gaussian isoperimetric inequality, let me first define a few terms that will be useful in the rest of this write-up. First, for $x \in \mathbb{R}^n$, the probability density function of the standard Gaussian distribution $\varphi_n(x)$ is:

$$\varphi_n(x) = \frac{1}{\sqrt{2\pi}} \exp(-\|x\|^2/2)$$

where $\|\cdot\|$ is the Euclidean norm.

The cumulative distribution function of one dimensional Gaussian distribution is (for convenience, I will omit the subscript when referring to one-dimensional Gaussian density):

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)dt.$$

Let $A \subset \mathbb{R}^n$ be a Borel set, then its n-dimensional Gaussian measure is:

$$\gamma_n(A) = \int_A \varphi_n(x)dx.$$

To state the Gaussian isoperimetric inequality formally, we still need to define what is the perimeter, or the "boundary measure", of a set. Here we will use the lower Minkovski content. Let $A \subset \mathbb{R}^n$ be a Borel set, then its n-dimensional lower Minkovski content with respect to the n-dimensional Gaussian measure is defined as:

$$\gamma_n^+(A) = \liminf_{h \to 0^+} \frac{\gamma_n(A^h) - \gamma_n(A)}{h}$$

where $A^h = \{x \in \mathbb{R}^n : \|x - a\| < h \text{ for some } a \in A\}$ is called the h-extension of A.

Intuitively, the lower Minkovski content tells you how fast the volume of a set grows as the "radius" of it increases, thus giving a measure of surface area.

Finally we can state our main result.

Theorem 1 (Gaussian Isoperimetric Inequality). *Let $A \subset \mathbb{R}^n$ be a Borel set, then for any $h > 0$*

$$\Phi^{-1}(\gamma_n(A^h)) \geq \Phi^{-1}(\gamma_n(A)) + h. \tag{1}$$

The connection between the above theorem and the isoperimetric problem in Gaussian space is not immediate. The proposition below makes the connection more explicit:

Proposition 1. *The Gaussian isoperimetric inequality is equivalent to the following statement: let $A \subset \mathbb{R}^n$ be a Borel set, and $H \subset \mathbb{R}^n$ be a half-space[1], such that $\gamma_n(A) = \gamma_n(H)$, then for any $h > 0$,*

$$\gamma_n(A^h) \geq \gamma_n(H^h). \tag{2}$$

Proof. For any half-space H, Let $R \in SO(n)$ be a rotation matrix such that $R(H)$ is in the form $R(H) = \{x \in \mathbb{R}^n : x^1 < t\}$ for some number t, where x^1 is the first coordinate of x. Since Gaussian probability density is invariant under rotation,

$$\gamma_n(H) = \gamma_n(R(H)) = \gamma_n\big((-\infty, t) \times \mathbb{R}^{n-1}\big) = \Phi(t).$$

It is clear that H^h is still a half-space (so is $R(H^h)$), and that $R(H^h) = \{x \in \mathbb{R}^n : x_1 < t + h\}$. Similar to above, $\gamma_n(H^h) = \Phi(t+h)$. The equivalence is immediate once we notice that:

$$\Phi^{-1}(\gamma_n(A)) + h = \Phi^{-1}(\gamma_n(H)) + h = t + h = \Phi^{-1}(\gamma_n(H^h)).$$

As an example of how one can apply the Gaussian isoperimetric inequality, we use it to heuristically estimate the order of Gaussian concentration inequality for Lipschitz function. Let $F : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function with Lipschitz seminorm $\|F\|_{\mathrm{Lip}} = L$. We endow $\mathbb{R}^n$ with the standard n-dimensional Gaussian measure, and let M be the median of F with respect to this measure. That is

$$\mathbb{P}(F \leq M) \geq \frac{1}{2} \qquad \text{and} \qquad \mathbb{P}(F \geq M) \geq \frac{1}{2}.$$

[1] A half-space is a set that can be written as $H = \{x \in \mathbb{R}^n : x \cdot u > r\}$ for some fixed $u \in \mathbb{R}^n$, and $r \in \mathbb{R}$, where $x \cdot u$ denotes the dot product between x and u

Let $A = \{x \in \mathbb{R}^n : F(x) \leq M\}$. Since F is continuous, A is closed. Therefore, for any $y \in A^h$, there exists an $x \in A$, such that $|y - x| < h$. By Lipschitz continuity, $|F(y) - F(x)| < Lh$. In addition, $x \in A$, so $F(x) - M \leq 0$. Hence $F(y) - M < Lh$. Now we apply the isoperimetric inequality,

$$\begin{aligned}\mathbb{P}(F - M < Lh) &\geq \gamma_n(A^h) \\ &\geq \Phi(\Phi^{-1}(\gamma_n(A)) + h) \\ &= \Phi(\Phi^{-1}(1/2) + h) \\ &= \Phi(h).\end{aligned}$$

Similarly,

$$\mathbb{P}(F - M > -Lh) \geq \Phi(h).$$

Notice that the intersection of the above two events is $\{y : |F(y) - M| < Lh\}$, and the union the entire $\mathbb{R}^n$. Therefore

$$\mathbb{P}(|F - M| < Lh) = \mathbb{P}(F - M < Lh) + \mathbb{P}(F - M > -Lh) - 1 \geq 2\Phi(h) - 1$$

which implies

$$\mathbb{P}(|F - M| \geq Lh) \leq 2(1 - \Phi(h)) \leq 2\exp\Big(-\frac{h^2}{2}\Big)$$

Therefore, the tail behavior of F around its median is of order $\exp(-h^2/2)$, so we also expect the tail behavior of F around its expectation to be of order $\exp(-h^2/2)$. That is, there exists a constant K, such that for large h

$$\mathbb{P}(|F - \mathbb{E}F| \geq Lh) \leq K\exp\Big(-\frac{h^2}{2}\Big)$$

which agrees to the concentration inequality.

Remark 1. Subtracting both sides of equation (2) with $\gamma_n(A) = \gamma_n(H)$, dividing them with h, and taking the limit infimums tells us the boundary measure of H is less than or equal to that of A. Moreover, we have:

$$\lim_{h\to\infty} \frac{\gamma_n(H+h) - \gamma_n(H)}{h} = \lim_{h\to\infty} \frac{\Phi(t+h) - \Phi(t)}{h} = \varphi(t) = \varphi(\Phi^{-1}(\gamma_n(A))$$

so the Gaussian isoperimetric inequality implies the following differential form:

$$\gamma_n^+(A) \geq \varphi(\Phi^{-1}(\gamma_n(A))). \tag{3}$$

As we will see later, this inequality also implies the original isoperimetric inequality, but the proof is more difficult.

2 Two Proofs of the Theorem

2.1 Generalize Isoperimetric Inequality From Discrete Cubes to Gauss Space

The following proof was given by S.G. Bobkov. The sketch of the proof is the following: we will first prove some calculus inequality, extend it by induction to multivariate case, use CLT to get functional inequality for Gaussian measures, and show that its equivalent to the standard formulation. Although this proof does contain some technical details, it is quite concise compared to the other proof we are giving in the next section.

We will start by considering the following function, for $p \in [0,1]$

$$I(p) = \varphi(\Phi^{-1}(p)).$$

Notice that $I(p)$ is not defined when $p = 0$ or $p = 1$, set $I(0) = I(1) = 0$, so that this function is continuous. The following proposition gives us the crucial property of this function:

Proposition 2. *For any* $0 \leq a, b \leq 1$,

$$I\left(\frac{a+b}{2}\right) \leq \frac{1}{2}\sqrt{I(a)^2 + \left(\frac{a-b}{2}\right)^2} + \frac{1}{2}\sqrt{I(b)^2 + \left(\frac{a-b}{2}\right)^2}. \tag{4}$$

The proof of this proposition is included in the appendix. In fact, proposition 1 can be viewed as an isoperimetric inequality on the one-dimensional discrete cube. To see this, consider the probability space $(\{-1,1\}, \mathscr{P}(\{-1,1\}), \mu)$, where $\mu(-1) = \mu(1) = \frac{1}{2}$, and a function $f : \{-1,1\} \to [0,1]$. Then by setting $a = f(-1)$, $b = f(1)$, (4) can be written as:

$$I(\mathbb{E}f) \leq \mathbb{E}\sqrt{I(f)^2 + |\nabla f|^2} \tag{5}$$

where the expectation is taken with respect to μ, and

$$|\nabla f| \equiv \left|\frac{f(1) - f(-1)}{2}\right| = \left|\frac{a-b}{2}\right|$$

is the norm of the discrete gradient of f. In general, we define the square of the norm of the discrete gradient of a function $f : \{-1,1\}^n \to \mathbb{R}$ as:

$$|\nabla f|^2 = \frac{1}{4}\sum_{i=1}^{n} |f(x) - f(s_i(x))|^2$$

where $s_i((x^1, \ldots, x^i, \ldots, x^n)) = (x^1, \ldots, -x^i, \ldots, x^n)$ is obtained by negating the ith coordinate of x while keeping the others unchanged. Eventually we want to apply this inequality to Gaussian measure and indicator functions. The following lemma is an n-dimensional version of proposition 1.

Lemma 1. *For a non-negative function F defined on $[0,1]$, if for all $g : \{-1,1\} \to [0,1]$*

$$F(\mathbb{E}g) \le \mathbb{E}\sqrt{F(g)^2 + |\nabla g|^2} \tag{6}$$

where the expectation is taken with respect to a probability measure μ defined on $\{-1,1\}$, then (6) also holds for all $f : \{-1,1\}^n \to [0,1]$, and the expectation is taken with respect to the product measure $\mu_n \equiv \mu^n$

Proof. By assumption, (6) holds when $n = 1$. We to show that (6) holds for any $n \in \mathbb{N}$ implies that it also holds for $n+1$.

Given μ and f, let $p_0 = \mu(-1)$, $p_1 = \mu(1)$, $f_0, f_1 : \{-1,1\}^n \to [0,1]$ defined as $f_0(x) = f(x,-1)$, $f_1(x) = f(x,1)$. Then by definition, we have:

$$\begin{aligned}
|\nabla f(x,0)|^2 &= \frac{1}{4}\sum_{i=1}^{n+1} |f(x,0) - f(s_i(x,0))|^2 \\
&= |\nabla f_0|^2 + \frac{1}{4}|f(x,0) - f(x,1)|^2 \\
&= |\nabla f_0(x)|^2 + \frac{1}{4}|f_0(x) - f_1(x)|^2.
\end{aligned}$$

Similarly, $|\nabla f(x,1)|^2 = |\nabla f_1(x)|^2 + \frac{1}{4}|f_0(x) - f_1(x)|^2$. Next, for $k \in \mathbb{N}$, let $\mathbb{E}_k$, be the expectation taken with respect to $\mu_k \equiv \mu^k$, then by Fubini's theorem, we can first integrate the first n variables and then integrate the last one:

$$\begin{aligned}
\mathbb{E}_{n+1}\sqrt{F^2(f) + |\nabla f|^2} &= \mathbb{E}\left(\mathbb{E}_n\sqrt{F^2(f) + |\nabla f|^2}\right) \\
&= p_0\mathbb{E}_n\sqrt{F^2(f_0) + |\nabla f_0|^2 + \frac{1}{4}|f_0 - f_1|^2} \\
&\quad + p_1\mathbb{E}_n\sqrt{F^2(f_1) + |\nabla f_1|^2 + \frac{1}{4}|f_0 - f_1|^2}
\end{aligned} \tag{7}$$

Now we use Minkovski's inequality to give an lower bound to the right-hand side in the above equation. Let u, v, be two non-negative functions, then the Minkovski's inequality for $p = \frac{1}{2}$, applied to the functions u^2 and v^2, implies that

$$\left(\int \sqrt{u^2 + v^2}\right)^2 \ge \left(\int \sqrt{u^2}\right)^2 + \left(\int \sqrt{v^2}\right)^2.$$

If we take the square root on both sides, and set the integral to be with respect to μ_n, then

$$\mathbb{E}_n\sqrt{u^2 + v^2} \ge \sqrt{(\mathbb{E}_n u)^2 + (\mathbb{E}_n v)^2}.$$

Now let $a_0 = \mathbb{E}_n f_0$, $a_1 = \mathbb{E}_n f_1$, $u_0 = \sqrt{F^2(f_0) + |\nabla f_0|^2}$, $u_1 = \sqrt{F^2(f_1) + |\nabla f_1|^2}$ and $v = \frac{1}{2}(f_0 - f_1)$. Then the above inequality implies:

$$\begin{aligned}\mathbb{E}_n \sqrt{F^2(f_0) + |\nabla f_0|^2 + \frac{1}{4}|f_0 - f_1|^2} &= \mathbb{E}_n \sqrt{u_0^2 + v^2} \\ &\geq \sqrt{(\mathbb{E}_n u_0)^2 + (\mathbb{E}_n v)^2}.\end{aligned}$$

Clearly $\mathbb{E}_n v = \frac{a_0 - a_1}{2}$, and by the induction hypothesis $\mathbb{E}_n u_0 \geq F(\mathbb{E}_n f_0) = F(a_0)$, so

$$\mathbb{E}_n \sqrt{F^2(f_0) + |\nabla f_0|^2 + \frac{1}{4}|f_0 - f_1|^2} \geq \sqrt{F(a_0)^2 + \left(\frac{a_0 - a_1}{2}\right)^2}.$$

Similarly,

$$\mathbb{E}_n \sqrt{F^2(f_1) + |\nabla f_1|^2 + \frac{1}{4}|f_0 - f_1|^2} \geq \sqrt{F(a_1)^2 + \left(\frac{a_0 - a_1}{2}\right)^2}.$$

By the above inequalities and (7),

$$\mathbb{E}_{n+1} \sqrt{F^2(f) + |\nabla f|^2} \geq p_0 \sqrt{F(a_0)^2 + \left(\frac{a_0 - a_1}{2}\right)^2} + p_1 \sqrt{F(a_1)^2 + \left(\frac{a_0 - a_1}{2}\right)^2}.$$

By our assumption, the right-hand side of the above expression is greater than or equal to $F(p_0 a_0 + p_1 a_1)$, but recall that

$$p_0 a_0 + p_1 a_1 = p_0(\mathbb{E}_n f_0) + p_1(\mathbb{E}_n f_1) = \mathbb{E}_{n+1} f.$$

Therefore

$$\mathbb{E}_{n+1} \sqrt{F^2(f) + |\nabla f|^2} \geq F(\mathbb{E}_{n+1} f).$$

□

Before we further generalize (6), let me first remind you what the multivariate central limit theorem says.

Theorem 2 (Multivariate Central Limit Theorem). *Let $X_1, \ldots, X_k : \Omega \to \mathbb{R}^n$ be independently and identically distributed random variables, with mean μ and covariance matrix Σ. Then as $k \to \infty$*

$$\frac{(X_1 - \mu) + \cdots + (X_k - \mu)}{\sqrt{k}} \xrightarrow{d} \mathcal{N}_n(0, \Sigma)$$

where $\mathcal{N}_n(0, \Sigma)$ stands for the n-dimensional Gaussian distribution with mean 0 *and covariance matrix Σ.*

Remark 2. We will use the following fact for the next lemma. If f is a bounded continuous function, and $\{S_k\}_{k\geq 1}$ are random variables such that $S_k \xrightarrow{d} Z$ as $k \to \infty$, for some random variable Z, then $\mathbb{E}_{S_k} f \to \mathbb{E}_Z f$.

Lemma 2. *Let $f : \mathbb{R}^n \to [0,1]$ be a continuously differentiable function with bounded partial derivatives. If F is a continuous function that satisfies the condition in (6), then*

$$F(\mathbb{E}_G f) \leq \mathbb{E}_G \sqrt{F(f)^2 + |\nabla f|^2}. \tag{8}$$

where E_G is the expectation taken with respect to the standard Gaussian measure in $\mathbb{R}^n$, and ∇f is the usual gradient of a differentiable function.

Proof. Let $k \in \mathbb{N}$. We look at the probability space $(\{-1,1\}^{nk}, \mathscr{P}(\{-1,1\}^{nk}), \mu^{nk})$. Let $X_1, \ldots, X_k : \{-1,1\}^{nk} \to \mathbb{R}^n$ be random variables on this probability space such that, for any $\omega = (\omega^1, \ldots, \omega^{nk}) \in \{-1,1\}^{nk}$, $X_i(\omega) = (\omega^{(i-1)k+1}, \ldots, \omega^{ik})$. In other words $X_1, \ldots, X_k$ divide ω into k blocks of vectors with length n, and each of the random variables project ω to one of these blocks. It is clear that $X_1, \ldots, X_k$ are independently and identically distributed, and each has covariance matrix the identity matrix. Now we define $f_k : \{-1,1\}^{nk} \to [0,1]$ with

$$f_k(x_1, \ldots, x_k) = f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big).$$

By the central limit theorem in $\mathbb{R}^n$, $(X_1 + \cdots + X_k)/\sqrt{k}$ converge in distribution to a standard Gaussian distribution in $\mathbb{R}^n$, since f is bounded and continuous

$$\mathbb{E}_{nk} f_k = \mathbb{E}_{nk} f_k(X_1, \ldots, X_k) = \mathbb{E}_{nk} f\Big(\frac{X_1 + \cdots + X_k}{\sqrt{k}}\Big) \to \mathbb{E}_G f.$$

Since F is continuous, we also have $F(\mathbb{E}_{nk} f_k) \to F(\mathbb{E}_G f)$. By our assumption,

$$F(\mathbb{E}_{nk} f_k) \leq \mathbb{E}_{nk} \sqrt{F(f_k)^2 + |\nabla f_k|^2}.$$

Hence

$$F(\mathbb{E}_G f) \leq \mathbb{E}_{nk} \sqrt{F(f_k)^2 + |\nabla f_k|^2}. \tag{9}$$

In addition, write $x = (x_1, \ldots, x_k) \in \mathbb{R}^{nk}$. For each $i \in \{1, \ldots, k\}$, let $s_i(x) = (x_{1,i}, \ldots, x_{k,i})$. It is clear that $\Big|[(x_1 + \cdots + x_k) - (x_{1,i} + \cdots + x_{k,i})]/\sqrt{k}\Big| = 2/\sqrt{k}$. By Taylor's theorem, we have

$$
\begin{aligned}
|\nabla f_k|^2 &= \frac{1}{4}\sum_{i=1}^{nk} |f_k(x) - f_k(s_i(x))|^2 \\
&= \frac{1}{4}\sum_{i=1}^{nk} \left| f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big) - f\Big(\frac{x_{1,i} + \cdots + x_{k,i}}{\sqrt{k}}\Big)\right|^2 \\
&= \frac{1}{4}\sum_{i=1}^{k}\sum_{j=1}^{n} \left| \frac{2}{\sqrt{k}}\partial_j f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big) + R_{ij}(k)\right|^2.
\end{aligned}
$$

Taylor's theorem tells us that $\sqrt{k}R_{ij}(k) \to 0$ as $k \to \infty$. After taking the square, $R_{ij}(k)^2$ has order $1/k$, and $(2R_{ij}(k)\partial_j f)/\sqrt{k}$ also has order $1/k$. Therefore,

$$
\begin{aligned}
|\nabla f_k|^2 &= \sum_{i=1}^{k}\sum_{j=1}^{n} \frac{1}{k}\left| \partial_j f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big)\right|^2 + \mathscr{O}(1/k) \\
&= \sum_{i=1}^{k} \frac{1}{k}\left| \nabla f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big)\right|^2 + \mathscr{O}(1/k) \\
&= \left| \nabla f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big)\right|^2 + k\mathscr{O}(1/k).
\end{aligned}
$$

Therefore, we can write $|\nabla f_k|^2 = \left|\nabla f(\frac{x_1+\cdots+x_k}{\sqrt{k}})\right|^2 + R(k)$, with $R(k) \to 0$ as $k \to \infty$. Without changing the notation of $R(k)$, we also have:

$$
\sqrt{F(f_k)^2 + |\nabla f_k|^2} = \sqrt{F\Big(f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big)\Big)^2 + \left|\nabla f\Big(\frac{x_1 + \cdots + x_k}{\sqrt{k}}\Big)\right|^2 + R(k)}.
$$

The $R(k)$ also goes to zero as $k \to \infty$. When we apply central limit theorem and compute the expectation of the above expression, the integral of $R(k)$ will vanish when $k \to \infty$. Moreover, the gradient of f is bounded and continuous, so the above function of $(x_1 + \cdots + x_k)/\sqrt{k}$ is also bounded and continuous. As a result, we have the following convergence: as $k \to \infty$

$$
\mathbb{E}_{nk}\sqrt{F(f_k)^2 + |\nabla f_k|^2} \to \mathbb{E}_G\sqrt{F(f)^2 + |\nabla f|^2}.
$$

Above and (9) implies:

$$
F(\mathbb{E}_G f) \le \mathbb{E}_G\sqrt{F(f)^2 + |\nabla f|^2}.
$$

□

It's tempting to set $F = I$ and f to be the indicator function of a set A, but there are still some obstacles. First, lemma 2 applies when f is a C^1 function, but most of the time indicator functions are not. Second, the relationship between the norm of the gradient and the boundary measure is still unclear. To solve the first problem, we start from approximating Lipschitz functions with smooth functions.

From now on all the expectations are taken with respect to the standard Gaussian distribution in $\mathbb{R}^n$

Lemma 3. *The conclusion in lemma 2 still holds when* $f : \mathbb{R}^n \to [0,1]$ *is a Lipschitz function.*

Proof. Given a Liptschitz function described in the lemma, we smooth it with a mollifier function $\psi : \mathbb{R}^n \to \mathbb{R}^+$. The mollifier satisfies the following condition:

- $\psi \in C^\infty$.
- ψ is compactly supported.
- $\psi_\varepsilon(x) \equiv \varepsilon^{-n}\psi(x/\varepsilon) \to \delta(x)$ as $\varepsilon \to 0$, where δ is the Dirac delta function centered at 0.
- $\int_{\mathbb{R}^n} \psi_\varepsilon(x)dx = 1$ for any $\varepsilon > 0$.

For ant $\varepsilon > 0$, we can define a smoothed version of f:

$$f_\varepsilon(y) \equiv \int_{\mathbb{R}^n} \psi_\varepsilon(y-x)f(x)dx.$$

When we take partial derivative of f_ε, the partial differential operator is passed to ψ_ε instead of f, so f_ε is C^∞. As $\varepsilon \to 0$

$$f_\varepsilon(y) \to \int_{\mathbb{R}^n} \delta(y-x)f(x)dx = f(y).$$

Moreover, by Hölder's inequality with $p = 1$, $q = \infty$, f_ε is uniformly bounded by $\sup_{x\in\mathbb{R}^n} f(x)$. By the bounded convergence theorem and the continuity of F

$$F(\mathbb{E}f_\varepsilon) \to F(\mathbb{E}f).$$

Next we look at the gradient. Take the partial derivative with respect to y^j and apply Fubini's theorem, and let dx^{-j} denote integration with respect to all but the jth coordinate of x.

$$\begin{aligned}
\frac{\partial}{\partial y^j} f_\varepsilon(y) &= \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \frac{\partial}{\partial y^j} \psi_\varepsilon(y-x)f(x)dx^j dx^{-j} \\
&= \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} -\frac{\partial}{\partial x^j} \psi_\varepsilon(y-x)f(x)dx^j dx^{-j} \\
&= \int_{\mathbb{R}^{n-1}} \left\{ \left[-\psi_\varepsilon(y-x)f(x) \right]_{x^j=-\infty}^{\infty} + \int_{\mathbb{R}} \psi_\varepsilon(y-x)\frac{\partial}{\partial x^j} f(x)dx^j \right\} dx^{-j} \\
&= \int_{\mathbb{R}^n} \psi_\varepsilon(y-x)\frac{\partial}{\partial x^j} f(x)dx \qquad \to \qquad \frac{\partial}{\partial x^j} f(y)
\end{aligned}$$

Above shows that the partial derivatives of f_ε converge to the partial derivatives of f at points where f is differentiable. Since f is Lipschitz, by Rademacher's theorem, it it differentiable almost everywhere, and its gradient is bounded by the corresponding Lipschitz constant. Therefore, ∇f_ε is bounded, and thus $\sqrt{F(f_\varepsilon)^2+|\nabla f_\varepsilon|^2}$ is bounded. As $\varepsilon \to 0$:

$$\mathbb{E}\sqrt{F(f_\varepsilon)^2+|\nabla f_\varepsilon|^2} \to \mathbb{E}\sqrt{F(f)^2+|\nabla f^2|}.$$

Since lemma 2 holds for all f_ε, we conclude that

$$F(\mathbb{E}f) \leq \mathbb{E}\sqrt{F(f)^2+|\nabla f|^2}.$$

for all Liptchitz function $f : \mathbb{R}^n \to [0,1]$. $\square$

Finally, we are ready to prove the differential version of the Gaussian isoperimetric inequality. The three lemmas above imply that, for any Lipschitz function $f : \mathbb{R}^n \to [0,1]$

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

which in turn implies that

$$I(\mathbb{E}f) \leq \mathbb{E}I(f)+\mathbb{E}|\nabla f|. \tag{10}$$

Let's $A \subset \mathbb{R}^n$ be a Borel set. For $h > 0$, define the function:

$$f_h(x) = \max\left\{1-\frac{d(x,A)}{h}, 0\right\}$$

where $d(x,A) \equiv \inf_{a\in A}|x-a|$. We claim that f_h is Lipchitz with Lipchitz constant $1/h$. Given any $x, y \in A$, $\varepsilon > 0$, we assume without loss of generality that $f_h(x) \geq f_h(y)$. Let $a_x^\varepsilon \in A$, such that $|x-a_x^\varepsilon| < d(x,A)+\varepsilon$. Then

$$\begin{aligned} f_h(x)-f_h(y) &\leq \left(1-\frac{d(x,A)}{h}\right)-\left(1-\frac{d(y,A)}{h}\right) \\ &= \frac{d(y,A)-d(x,A)}{h} \\ &< \frac{|y-a_x^\varepsilon|-|x-a_x^\varepsilon|-\varepsilon}{h} \leq \frac{|x-y|-\varepsilon}{h}. \end{aligned}$$

Taking $\varepsilon \to 0$, $f_h(x)-f_h(y) \leq |x-y|/h$, so f_h is indeed Lipschitz with Lipschitz constant $1/h$. As a result, by Rademachers theorem,

$$|\nabla f_h| \leq \frac{1}{h}$$

whenever the gradient of f_h exists (which is almost everywhere). Recall that $A^h = \{x \in \mathbb{R}^n : d(x,A) < h\}$ is the h-extension of A. For any point p in A, $f_h(p) = 1$ is a local maximum, so if f_h is also differentiable at p, $|\nabla f_h(p)| = 0$. Similarly, $|\nabla f_h(p)| = 0$ for $p \in (A^h)^c$. Therefore,

$$I(\mathbb{E}f_h) \le \mathbb{E}|\nabla f_h| + \mathbb{E}I(f_h) \le \int_{A^h - A} \frac{1}{h} d\gamma_n + \mathbb{E}I(f_h) = \frac{\gamma_n(A^h) - \gamma_n(A)}{h} + \mathbb{E}I(f_h). \tag{11}$$

Remark 3. If here A is closed, then $d(y,A) > 0$ for any $y \notin A$, so $f_h \downarrow \mathbb{1}_A$. By monotone convergence theorem, $\mathbb{E}f_h \downarrow \gamma_n(A)$, and by dominated convergence theorem, $\mathbb{E}I(f_h) \downarrow 0$. We also know that I is increasing on $[0,1/2)$, and decreasing on $[1/2,1]$, so for a small $h > 0$ (small enough such that $\mathbb{E}f_h$ stays on the same side of $1/2$ as $\gamma_n(A)$),

$$\min\{I(\gamma_n(A)), I(\mathbb{E}f_h)\} = \inf_{0<t\le h} I(\mathbb{E}f_t) \le \inf_{0<t\le h} \frac{\gamma_n(A^t) - \gamma_n(A)}{t} + \mathbb{E}I(f_h) \tag{12}$$

Take limit on both sides, we have the differential version of our theorem

$$I(\gamma_n(A)) \le \liminf_{h\to 0^+} \frac{\gamma_n(A^h) - \gamma_n(A)}{h} = \gamma_n^+(A). \tag{13}$$

Now let's finish the proof. By (11), for $\delta > 0$,

$$\min\left\{\varphi(s) : \Phi^{-1}(\gamma_n(A)) \le s \le \Phi^{-1}(\gamma_n(A^\delta))\right\} \le \frac{\gamma_n(A^\delta) - \gamma_n(A)}{\delta} + \mathbb{E}I(f_\delta). \tag{14}$$

Since $I(f_\delta)(x) = 0$ when $x \in A$ or when $x \in (A^\delta)^c$, and I is uniformly bounded by $\varphi(0) < 1$,

$$\mathbb{E}I(f_\delta) \le \gamma_n(A^\delta) - \gamma_n(A).$$

So we have,

$$\min\left\{\varphi(s) : \Phi^{-1}(\gamma_n(A)) \le s \le \Phi^{-1}(\gamma_n(A^\delta))\right\} \le \Big(\gamma_n(A^\delta) - \gamma_n(A)\Big)\Big(1 + \frac{1}{\delta}\Big).$$

Using the same inequality for A^x instead of A, and multiplying both sides by $\delta/(1+\delta)$,

$$\frac{\delta}{1+\delta} \min\left\{\varphi(s) : \Phi^{-1}(\gamma_n(A^x)) \le s \le \Phi^{-1}(\gamma_n(A^{x+\delta}))\right\} \le \gamma_n(A^{x+\delta}) - \gamma_n(A^x). \tag{15}$$

If we consider the non-decreasing function $u(x) = \Phi^{-1}(\gamma_n(A^x))$ for $x \ge 0$, we can rewrite this as

$$\frac{\delta}{1+\delta} \min\left\{\varphi(s) : u(x) \le s \le u(x+\delta)\right\} \le \Phi(u(x+\delta)) - \Phi(u(x)). \tag{16}$$

On the other hand, by the mean value theorem,

$$\Phi(u(x)+\delta) - \Phi(u(x)) \le \delta \max\left\{\varphi(s) : u(x) \le s \le u(x) + \delta\right\}. \tag{17}$$

Using that the derivative φ' is uniformly bounded, in fact $\|\varphi'\|_\infty \le 1$, by the mean value theorem,

$$\max\{\varphi(s) : u(x) \le s \le u(x)+\delta\} \le \varphi(u(x))+\delta \tag{18}$$

and

$$\min\{\varphi(s) : u(x) \le s \le u(x+\delta)\} \ge \varphi(u(x)) - (u(x+\delta)-u(x)). \tag{19}$$

Combining all the inequalities, we get

$$\begin{aligned}\Phi(u(x+\delta)) - \Phi(u(x)+\delta) &\ge \frac{\delta}{1+\delta}\Big[\varphi\big(u(x)\big) - \big(u(x+\delta)-u(x)\big)\Big] - \delta\Big[\varphi\big(u(x)\big)+\delta\Big]\\ &= -\frac{\delta^2}{1+\delta}\varphi\big(u(x)\big) - \frac{\delta}{1+\delta}\big[u(x+\delta)-u(x)\big] - \delta^2\\ &\ge -\frac{\delta^2}{1+\delta} - \frac{\delta}{1+\delta}\big[u(x+\delta)-u(x)\big] - \delta^2.\end{aligned} \tag{20}$$

If $u(x+\delta) \ge u(x)+\delta$ then by the monotonicity of Φ

$$\Phi(u(x+\delta)) \ge \Phi(u(x)+\delta). \tag{21}$$

If $u(x+\delta) \le u(x)+\delta$ then $u(x+\delta)-u(x) \le \delta$ and (20) implies

$$\Phi(u(x+\delta)) - \Phi(u(x)+\delta) \ge -\frac{2\delta^2}{1+\delta} - \delta^2 \ge -3\delta^2. \tag{22}$$

In either case, we showed that

$$\Phi(u(x+\delta)) \ge \Phi(u(x)+\delta) - 3\delta^2. \tag{23}$$

Now, notice that we can suppose that the set A and $h>0$ are such that probabilities $p_0 := \gamma_n(A)$ and $p_1 := \gamma_n(A^h)$ are strictly between 0 and 1, otherwise, there is nothing to prove. Let us take δ small enough so that $3\delta^2 \le p_0/2$, in which case

$$\Phi(u(x)+\delta) - 3\delta^2 \ge \Phi(u(0)) - 3\delta^2 = p_0 - 3\delta^2 \ge \frac{p_0}{2} =: a$$

and, for any x in the interval $[0,h]$,

$$\Phi(u(x)+\delta) \le \Phi(\Phi^{-1}(p_1)+\delta) =: b.$$

If K is the maximum of the derivative of Φ^{-1} on the interval $[a,b]$ then, by the mean value theorem,

$$\Phi^{-1}\big(\Phi(u(x)+\delta) - 3\delta^2\big) \ge \Phi^{-1}\big(\Phi(u(x)+\delta)\big) - 3K\delta^2 \ge u(x)+\delta-3K\delta^2.$$

This means that taking inverse Φ^{-1} of (23), we get

$$u(x+\delta) \geq u(x) + \delta - 3K\delta^2. \tag{24}$$

Now, let $\delta = h/n$ for large $n \geq 1$ and take $x = kh/n$ for $k = 0, 1, \ldots, n-1$. Then

$$u((k+1)h/n) - u(kh/n) \geq h/n - 3Kh^2/n^2. \tag{25}$$

Adding these over all k, we get $u(h) - u(0) \geq h - 3Kh^2/n$. Letting $n \to \infty$ proves that $u(h) \geq u(0) + h$, which is precisely the isoperimetric inequality.

2.2 Proof from Geometric Viewpoint

This proof basically uses measure theory and some geometric propertis of Euclidean space, and is based heavily on Section 11 from Gaussian Random Functions(M.A. Lifeshits). It is sparked by the classical proof of isoperimetric problem for the Euclidean space, and the idea is to construct an operation (called symmetrization here) to send a closed or open subset A to a "nicer" form; i.e. we keep shrinking the surface area of set A with its volume fixed by applying this operation, until we get our desired set. And it turns out that the "best" set we can get is a half space.

The followings are some notations we use in this proof: We use γ_n to denote the n-dimensional Gaussian measure on $\mathbb{R}^n$, and use Φ to denote the cumulative density function of 1-dimensional standard Gaussian distribution as before. And fix $\mathbf{e} \in \mathbb{R}^n$ and $a \in \mathbb{R}$, then define $\Pi(\mathbf{e}, a) = \{\mathbf{x} \in \mathbb{R}^n | \langle \mathbf{x}, \mathbf{e} \rangle > a\}$ as the n-dimensional half space corresponding to vector $\mathbf{e}$ and scalar a. Also we need to modify a notation from the previous proof: we define the h-extension of A as $A^h = \{x \in \mathbb{R}^n : \|x - a\| \leq h$ for some $a \in A\}$ for some $h > 0$. By this modification, we can get a closed extension A^h when A is closed. Moreover, let $\mathscr{D}_h$ denote the closed ball centered at origin with radius h, then $A^h = A + \mathscr{D}_h$.

2.2.1 Symmetrization

We first construct a mapping sending each k-dimensional slice of an open or closed set A in $\mathbb{R}^n$ to a corresponding half space within that slice.

Definition 1. Take $1 \leq k \leq n$ and fix a $(n-k)$-dimensional subspace L of $\mathbb{R}^n$, and pick a unit vector $\mathbf{e}$ orthogonal to L.

For an arbitrary open or closed subset $A \subset \mathbb{R}^n$, we define a subset $A' \subset \mathbb{R}^n$ to be the k-symmetrization of A with respect to L along $\mathbf{e}$, written as $A' = S(L, \mathbf{e})[A]$ or $A' = S[A]$, if for any $\mathbf{x} \in \mathbb{R}^n$, A' satisfies the following:

1. If $\gamma_k\left(\left(\mathbf{x} + L^\perp\right) \cap A\right) = 0$, then
$$(\mathbf{x} + L^\perp) \cap A' = \emptyset.$$
2. If $\gamma_k\left(\left(\mathbf{x} + L^\perp\right) \cap A\right) = 1$, then

$$(\mathbf{x}+L^{\perp})\cap A' = \mathbf{x}+L^{\perp}.$$

3. If $0 < \gamma_k\left((\mathbf{x}+L^{\perp})\cap A\right) < 1$, then

   a. If A is open, then $(\mathbf{x}+L^{\perp})\cap A' = (\mathbf{x}+L^{\perp})\cap \Pi(\mathbf{e},a)$;

   b. If A is closed, then

$$(\mathbf{x}+L^{\perp})\cap A' = (\mathbf{x}+L^{\perp})\cap Closure\ \Pi(\mathbf{e},a);$$

   where a is determined by the equation:

$$\gamma_k\left(\left(\mathbf{x}+L^{\perp}\right)\cap A\right) = \gamma_k\left((\mathbf{x}+L^{\perp})\cap \Pi(\mathbf{e},a)\right).$$

Remark 4. It makes no difference if we only consider $\mathbf{x}\in L$ instead of $\mathbf{x}\in\mathbb{R}^n$ in this definition.

By this construction, for each $\mathbf{x}\in\mathbb{R}^n$, the slice $A\cap(\mathbf{x}+L^{\perp})$ is replaced by a k-dimensional half space with the same volume under Gaussian measure. In particular, if we take $k=n$, $L=\{0\}$, and $\mathbf{e}$ to be any unit vector in $\mathbb{R}^n$, then $S(A)$ would be a half space having the same Gaussian measure as A. And the followings are some properties of symmetrization:

Proposition 3. *The k-symmetrization S has the following properties:*

1. *(Monotonicity.) If $A\subset B$, $S[A]$ and $S[B]$ are defined, then*

$$S[A]\subset S[B].$$

2. *(Lower continuity.) If $\{A_i\}_{i\in\mathbb{N}}$ is a sequence of open sets with $A_i\subset A_{i+1}$ for each i, then*

$$S[\bigcup_{i\in\mathbb{N}} A_i] = \bigcup_{i\in\mathbb{N}} S[A_i].$$

   For the following properties, A is an arbitrary open or closed subset in $\mathbb{R}^n$:

3. *(Consistency with taking complement.) Let A^c be the complement of A, then we have*

$$S(L,\boldsymbol{e})[A^c] = (S(L,-\boldsymbol{e})[A])^c.$$

4. *(Invariance with respect to $(L+\mathscr{L}\{\boldsymbol{e}\})^{\perp}$.) Let $\mathscr{L}\{\boldsymbol{e}\}$ be the linear space spanned by vector $\boldsymbol{e}$, then we have*

$$S[A]+(L+\mathscr{L}\{\boldsymbol{e}\})^{\perp} = S[A].$$

5. *(Semi-invariance with respect to $\mathscr{L}(\boldsymbol{e})$.) For any $c\geq 0$, we have*

$$S[A]+c\boldsymbol{e}\subset S[A].$$

6. *(Invariance with respect to L.) For any $\boldsymbol{l}\in L$, we have*

$$S[A+\boldsymbol{l}] = S[A]+\boldsymbol{l}.$$

Moreover, if M is a subspace of L and $M+A=A$, then we have

$$S[A]+M=S[A].$$

7. *(Measure preserving.)Let B be a Borel set with the Gaussian measure satisfying $B+L^{\perp}=B$, then*
$$\gamma_n(B\cap A)=\gamma_n(B\cap S[A]).$$
In particular,
$$\gamma_n(A)=\gamma_n(S[A]).$$

Proof. 1. Fix $\mathbf{x}\in L$. On the slice $\mathbf{x}+L^{\perp}$, we have

$$A\cap(\mathbf{x}+L^{\perp})\subset B\cap(\mathbf{x}+L^{\perp}),$$

therefore,

$$S[A]\cap(\mathbf{x}+L^{\perp})\subset S[B]\cap(\mathbf{x}+L^{\perp})$$

by definition of symmetrization. Take the union of all slices and we get

$$S[A]\subset S[B].$$

2. Fix $\mathbf{x}\in L$. By lower continuity of γ_k, we have
$$\gamma_k\big(\bigcup_{i\in\mathbb{N}}(A_i\cap(\mathbf{x}+L^{\perp}))\big)=\bigcup_{i\in\mathbb{N}}\gamma_k(A_i\cap(\mathbf{x}+L^{\perp})).$$
This implies
$$S[\bigcup_{i\in\mathbb{N}}A_i]\cap(\mathbf{x}+L^{\perp})=\bigcup_{i\in\mathbb{N}}S[A_i]\cap(\mathbf{x}+L^{\perp}).$$
Take the union of all slices and get
$$S[\bigcup_{i\in\mathbb{N}}A_i]=\bigcup_{i\in\mathbb{N}}S[A_i].$$
3. On each slice $\mathbf{x}+L^{\perp}$, $A\cap(\mathbf{x}+L^{\perp})$ and $A^c\cap(\mathbf{x}+L^{\perp})$ are sent to half spaces over that slice in opposite directions, and they should cover the whole space since their Gaussian measures sum up to 1.
4. It follows from the fact that $\Pi(\mathbf{e},a)\cap(\mathbf{x}+L^{\perp})$ is invariant with respect to $(L+\mathscr{L}\{\mathbf{e}\})^{\perp}$ for all $a\in\mathbb{R}$ and all $\mathbf{x}\in L$.
5. With the fact that $\Pi(\mathbf{e},a)+c\mathbf{e}\subset\Pi(\mathbf{e},a)$ for all $a\in\mathbb{R}$ and all $c\geq 0$, it is easy to check
$$S[A]\cap(\mathbf{x}+L^{\perp})+c\mathbf{e}\subset S[A]\cap(\mathbf{x}+L^{\perp})$$
for all $\mathbf{x}\in L$ and all $c\geq 0$.
6. For each $\mathbf{x}\in L$, it is easy to check
$$\gamma_k(A\cap(\mathbf{x}+L^{\perp}))=\gamma_k((A+\mathbf{l})\cap((\mathbf{x}+\mathbf{l})+L^{\perp})).$$

This implies

$$S[A] \cap (\mathbf{x} + L^{\perp}) + \mathbf{l} = S[A + \mathbf{l}] \cap ((\mathbf{x} + \mathbf{l}) + L^{\perp}).$$

Take union over all $\mathbf{x} \in L$ and we will get

$$S[A + \mathbf{l}] = S[A] + \mathbf{l}.$$

7. Follows directly from Fubini's theorem.

2.2.2 Gaussian isoperimetric inequality

We will prove the Gaussian isoperimetric inequality using the fact that the symmetrization S defined in Section 2.2.1 reduces the surface area.

Theorem 3 (The reduction of the surface area under symmetrization).
Let $S = S(L, \boldsymbol{e})$ be a k-symmetrization as defined in Section 2.2.1.
Then for any closed set $A \in \mathbb{R}^n$, we have

$$S[A^h] \supset S[A]^h. \tag{26}$$

Proof. The proof of Inclusion (26) will be presented in Section 2.2.3.

Before introducing the Gaussian isoperimetric inequality, we need to show that $S[A]$ is a Borel set for any open or closed set A:

Lemma 4. *The symmetrization S translates closed sets to closed sets and open sets to open sets.*

Proof. Suppose that A is closed in $\mathbb{R}^n$. Then the $\frac{1}{n}$-extension $A^{\frac{1}{n}}$ of A is closed for all $n \in \mathbb{N}$, and $A = \bigcap_{n\in\mathbb{N}} A^{\frac{1}{n}}$. Therefore, by properties 2 and 3 from Proposition 3

$$S[A] = (S(L, -\mathbf{e})[A^c])^c = \left(S(L, -\mathbf{e})[\bigcup_{n\in\mathbb{N}} (A^{\frac{1}{n}})^c]\right)^c = \bigcap_{n\in\mathbb{N}} \left(S(L, -\mathbf{e})[(A^{\frac{1}{n}})^c]\right)^c = \bigcap_{n\in\mathbb{N}} \left(S[A^{\frac{1}{n}}]\right).$$

Apply Inclusion (26) to $A^{\frac{1}{n}}$ for all $n \in \mathbb{N}$ and get

$$S[A] = \bigcap_{n\in\mathbb{N}} \left(S[A^{\frac{1}{n}}]\right) \supset \bigcap_{n\in\mathbb{N}} \left(S[A]^{\frac{1}{n}}\right).$$

And $S[A] \subset S[A]^{\frac{1}{n}}$ for all $n \in \mathbb{N}$ implies

$$S[A] \subset \bigcap_{n\in\mathbb{N}} \left(S[A]^{\frac{1}{n}}\right).$$

Therefore,

$$S[A] = \bigcap_{n \in \mathbb{N}} \left(S[A]^{\frac{1}{n}} \right),$$

from which we conclude $S[A]$ is closed.

Suppose that B is open in $\mathbb{R}^n$, then by property 3 from Proposition 3, $S[B] = (S(L, -\mathbf{e})[B^c])^c$ is open in $\mathbb{R}^n$.

Note 1. We need the following results from the previous proof: If H is a half space in $\mathbb{R}^n$, and fix $h > 0$, then

$$\Phi^{-1}(\gamma_n(H^h)) = \Phi^{-1}(\gamma_n(H)) + h.$$

We proved this when proving Proposition 1.

Now we can prove the Gaussian isoperimetric inequality with our modified definition of h-extension of A:

Theorem 4 (Gaussian Isoperimetric Inequality). *Let $A \subset \mathbb{R}^n$ be a Borel set, then for any $h > 0$*

$$\Phi^{-1}(\gamma_n(A^h)) \geq \Phi^{-1}(\gamma_n(A)) + h. \tag{27}$$

Proof. We prove this for a closed set A first:

Fix a n-symmetrization S in $\mathbb{R}^n$, then by Inclusion (26)

$$S[A^h] \supset S[A]^h,$$

which implies

$$\gamma_n(S[A^h]) \geq \gamma_n(S[A]^h).$$

Therefore, by the monotonicity of Φ^{-1}, we have

$$\Phi^{-1}(\gamma_n(A^h)) = \Phi^{-1}\left(\gamma_n(S[A^h])\right) \geq \Phi^{-1}\left(\gamma_n(S[A]^h)\right). \tag{28}$$

Also, $S[A]$ is a half space, then

$$\Phi^{-1}\left(\gamma_n(S[A]^h)\right) = \Phi^{-1}\left(\gamma_n(S[A])\right) + h. \tag{29}$$

Therefore,

$$\Phi^{-1}(\gamma_n(A^h)) \geq \Phi^{-1}\left(\gamma_n(S[A])\right) + h.$$

We proved that the Gaussian isoperimetric inequality for an arbitrary closed subset A. But it is not hard to generalize our result to all Borel sets from all closed sets by regularity of Gaussian measure as shown in the previous proof.

2.2.3 Reduction of surface area under symmetrization

In this section, we will give the proof of Inclusion (26). This proof involves the following four steps:

1. Prove Inclusion (26) when $n = k = 1$.
2. Prove Inclusion (26) when $n \geq k = 1$.
3. Prove Inclusion (26) when $n = 2 \geq k \geq 1$ by showing every 2-symmetrization on $\mathbb{R}^2$ is a limit of compositions of 1-symmetrizations.
4. Prove Inclusion (26) when $n \geq k \geq 3$ by showing the k-symmetrization can be written as a composition of finite many 2-symmetrizations.

Step 1: Symmetrizations on $\mathbb{R}$
There are only two symmetrizations over $\mathbb{R}$, i.e.

$$S_+ = (\{0\}, 1)$$

and

$$S_- = (\{0\}, -1)$$

We will prove that Inclusion (26) for $S = S_-$, and the case S_+ follows by the symmetry of Gaussian measure, i.e. $S_+[A]$ and $S_-[A]$ are symmetric about the origin. We break this proof into the following lemmas:

Lemma 5. *Inclusion (26) holds for all open (or closed) intervals in $\mathbb{R}$.*

Proof. Let $A \subset \mathbb{R}$ be an open (or closed) interval, and let $p = \gamma_1(A)$, then we can define the following family of all open (or closed) intervals with probability p in $\mathbb{R}$:
If A is open,

$$\{A_u = (u, v(u)) |\ u \in [-\infty, \Phi^{-1}(1-p)] \text{ and } v(u) = \Phi^{-1}(p + \Phi(u))\}.$$

If A is closed,

$$\{A_u = [u, v(u)] |\ u \in [-\infty, \Phi^{-1}(1-p)] \text{ and } v(u) = \Phi^{-1}(p + \Phi(u))\}.$$

Easy to check $S[A] = A_{-\infty}$.
Fix $h > 0$. Define a function $p_h : [-\infty, \Phi^{-1}(1-p)] \to \mathbb{R}$ by

$$p_h(u) = \gamma_1[(A_u)^h].$$

Then we want to find the minimum point of p_h. Differentiate p_h and get:

$$
\begin{aligned}
p_h'(u) &= \frac{d}{du}\int_{u-h}^{v(u)+h}\phi(x)dx \\
&= \phi(u-h) - v'(u)\phi(v(u)+h) \\
&= \phi(u-h) - \phi(v(u)+h)\left(\frac{d}{du}\Phi^{-1}(p+\Phi(u))\right) \\
&= \phi(u-h) - \phi(v(u)+h)\frac{\phi(u)}{\phi\left(\Phi^{-1}(p+\Phi(u))\right)} \\
&= \phi(u-h) - \phi(v(u)+h)\frac{\phi(u)}{\phi\left(v(u)\right)} \\
&= \phi(u)\left[\frac{\phi(u-h)}{\phi\left(u\right)} - \frac{\phi(v(u)+h)}{\phi\left(v(u)\right)}\right] \\
&= \phi(u)\left[\frac{\phi(-u+h)}{\phi\left(-u\right)} - \frac{\phi(v(u)+h)}{\phi\left(v(u)\right)}\right] \\
&= \theta(-u) - \theta(v(u)),
\end{aligned}
$$

where

$$
\begin{aligned}
\theta(x) &= \frac{\phi(x+h)}{\phi\left(x\right)} \\
&= exp\left[log(\phi(x+h)) - log(\phi(x))\right] \\
&= exp\left[\int_x^{x+h}(log\phi)'(t)dt\right].
\end{aligned}
$$

One can check by computation that $log(\phi)$ is concave, therefore,

$$
\begin{aligned}
\theta'(x) &= \frac{d}{dx}exp\left[\int_x^{x+h}(log\phi)'(t)dt\right] \\
&= \left[(log\phi)'(x+h) - (log\phi)'(x)\right]\exp\left[\int_x^{x+h}(log\phi)'(t)dt\right] \\
&< 0.
\end{aligned}
$$

θ is strictly decreasing and we know for $p_h(u)$:

1. p_h is increasing when $v < -u$;
2. p_h is decreasing when $v > -u$.

Also $v = -u$ implies $u = \Phi^{-1}\left(1-\frac{p}{2}\right)$. Take $u_0 = \Phi^{-1}\left(1-\frac{p}{2}\right)$, then

1. p_h is increasing when $u < u_0$;
2. p_h is decreasing when $u > u_0$.

By symmetry of Gaussian measure, we know p_h is symmetric about u_0. Hence,

$$
p_h(-\infty) \le p_h(u)
$$

for all u where $p_h(u)$ is defined.Then by the measure preserving property of S in Proposition 3,

$$\gamma_1\left(S[A_u]^h\right)=\gamma_1\left((A_{-\infty})^h\right)=p_h(-\infty)\leq p_h(u)=\gamma_1\left((A_u)^h\right)=\gamma_1\left(S[(A_u)^h]\right).$$

Both $S[A_u]^h$ and $S[(A_u)^h]$ are left rays, therefore,

$$S[A_u]^h\subset S[(A_u)^h].$$

This is just the inclusion (26) restricted to all open(or closed) intervals in $\mathbb{R}$

Lemma 6. *Inclusion (26) holds for finite union of disjoint open (or closed) intervals in* $\mathbb{R}$.

Proof. Again, we only prove Inclusion (26) for $S=S_-$. Assume that $\{A_i\}_{i=1}^{m+1}$ is a collection of disjoint open (or closed) intervals on $\mathbb{R}$ arranged from left to right. We will prove that Inclusion (26) holds for $A=\sqcup_{i=1}^{m+1}A_i$ by induction on m.

The base case m=0 is exactly the previous lemma. Assume that Inclusion (26) holds when $m\leq n-1$ for some $n\in\mathbb{N}$, and then consider the case when $m=n$: Fix $h>0$, and let $A_1,A_2,\ldots,A_{n+1}$ be disjoint open intervals in $\mathbb{R}$. We start by assuming $A_i^h\cap A_{i+1}^h=\emptyset$ for all $1\leq i\leq n$.

Define $J=\sqcup_{i=2}^{n}A_i$. Construct a new set A' by first replacing $A_1,A_2,\ldots,A_n,A_{n+1}$ with $S[A_1],A_2,\ldots,A_n,S_+[A_{n+1}]$ and then taking union, i.e. $A'=S[A_1]\sqcup J\sqcup S_+[A_{n+1}]$.

We know symmetrization preserves Gaussian measure by Proposition 3. S sends A_1 to a left ray and S_+ sends A_{n+1} to a right ray. Therefore, $S[A_1],A_2,\ldots,S_+[A_{n+1}]$ are disjoint open intervals arranged from left to right with the same Gaussian measure with $A_1,A_2,\ldots,A_{n+1}$ respectively. And $S[A_1]^h,A_2^h,\ldots,S_+[A_{n+1}]^h$ are disjoint. Hence,

$$S[A]=S[A_1\sqcup J\sqcup A_{n+1}]=S[S[A_1]\sqcup J\sqcup S[A_{n+1}]]=S[A'].$$

Therefore,

$$S[A]^h=S[A']^h. \tag{30}$$

Also, we can apply Inclusion (26) to A_1 and A_{n+1} since they are both open intervals and get

$$S[A_1^h]\supset S[A_1]^h,$$
$$S[A_{n+1}1^h]\supset S[A_{n+1}]^h.$$

Hence, by measure preserving of S from Proposition 3, we have

$$\begin{aligned}S[A^h]=S[A_1^h\sqcup J^h\sqcup A_{n+1}^h]&=S[S[A_1^h]\sqcup J^h\sqcup S[A_{n+1}^h]]\\&\supset S[S[A_1]^h\sqcup J^h\sqcup S[A_{n+1}]^h]=S[(A')^h],\end{aligned}$$

i.e.

$$S[A^h]\supset S[(A')^h]. \tag{31}$$

Let $I = (A')^h$. It is easy to verify that $(((B^h)^c)^h)^c = B$ holds for any subset $B \subset \mathbb{R}$. Therefore, we can also write $A' = ((I^c)^h)^c$. Using property 2 from Proposition 3, we have $S[((I^c)^h)^c] = S_+[(I^c)^h]^c$. I^c is a union of n disjoint closed intervals, and this means we can apply our induction hypothesis to it. Therefore,

$$S[A'] = S[((I^c)^h)^c] = (S_+[(I^c)^h)]^c \subset (S_+[I^c]^h)^c.$$

Hence,

$$S[A']^h \subset ((S_+[I^c]^h)^c)^h = S_+[I^c]^c = S[I] = S[(A')^h],$$

i.e.

$$S[A']^h \subset S[(A')^h]. \tag{32}$$

Hence, by (30), (31) and (32),

$$S[A]^h = S[A']^h \subset S[(A')^h] \subset S[A]^h.$$

Now we consider the case that $A_i^h \cap A_{i+1}^h \neq \emptyset$ for some i . We define B_i to be the open interval such that

$$B_i^h = A_i^h \cup A_{i+1}^h.$$

Consider the sequence $A_1, A_2, \ldots A_{i-1}, B_i, A_{i+2}, \ldots, A_{n+1}$. And define

$$B = A_1 \sqcup A_2 \sqcup \ldots \sqcup A_{i-1} \sqcup B_i \sqcup A_{i+2} \sqcup \ldots \sqcup A_{n+1}.$$

Then $B^h = A^h$ and $B \supset A$ by our construction. B consists of n disjoint open intervals, therefore, we can apply induction hypothesis to B. Together with property 1 from Proposition 3, we get

$$S[A^h] = S[B^h] \supset S[B]^h \supset S[A]^h$$

The proof for the case $A_1, A_2, \ldots, A_{n+1}$ are disjoint closed intervals is similar.

For an open subset $A \subset \mathbb{R}$, we can write it as a countable union of open intervals. After combining all the overlapping intervals, we can write $A = \bigsqcup_{i \in \mathbb{N}} I_i$ where I_is are disjoint open intervals. Define $A_n = \bigsqcup_{i=1}^n I_i$. Then $\{A_n\}$ is an increasing sequence of open sets with $A_n \to A$ as $n \to \infty$. By lower continuity of S from Proposition 3 and previous lemma, we have

$$\begin{aligned} S[A]^h = S\left[\bigcup_{n\in\mathbb{N}} A_n\right]^h = \left(\bigcup_{n\in\mathbb{N}} S[A_n]\right)^h &= \bigcup_{n\in\mathbb{N}} (S[A_n]^h) \\ &\subset \bigcup_{n\in\mathbb{N}} S[A_n^h] = S\left[\bigcup_{n\in\mathbb{N}} A_n^h\right] = S[A^h]. \end{aligned}$$

Therefore, Inclusion (26) holds for any open set $A \subset \mathbb{R}$.

Fix a closed subset B of $\mathbb{R}$, then $B = \bigcap_{n\in\mathbb{N}} B_n$ for some open sets B_n. By upper continuity of 1-dimensional Gaussian measure, we have $\gamma_1(B) = \lim_{n\to\infty} \gamma_1(B_n)$, which implies $S[B] = \bigcap_{n\in\mathbb{N}} S[B_n]$. Similarly, $S[B^h] = \bigcap_{n\in\mathbb{N}} S[B_n^h]$. Then

$$S[B]^h = \left(\bigcap_{n \in \mathbb{N}} S[B_n] \right)^h \subset \bigcap_{n \in \mathbb{N}} (S[B_n])^h \subset \bigcap_{n \in \mathbb{N}} S[B_n^h] = S[B^h].$$

Therefore, we generalize Inclusion (26) to any closed set $B \subset \mathbb{R}$.

Step 2: 1**-symmetrization on** $\mathbb{R}^n$

The 1-symmetrization on $\mathbb{R}^n$ is of the form $S\left(L = \mathscr{L}\{\mathbf{e}\}^\perp, \mathbf{e}\right)$, where $\mathbf{e}$ is a unit vector in $\mathbb{R}^n$. And define $R_\mathbf{x} = \mathbf{x} + \mathscr{L}\{\mathbf{e}\} = \{\mathbf{x} + r\mathbf{e} | r \in \mathbb{R}\}$. $R_\mathbf{x}$ represents the 1-dimensional slice where we try to replace $A \cap R_\mathbf{x}$ with a corresponding half space. We want to verify Inclusion (26) restricted to each slice.

Fix $\mathbf{x} \in L$, by our definition of S, we have

$$S[A^h] \cap R_\mathbf{x} = S[A^h \cap R_\mathbf{x}].$$

For any $\mathbf{k} \in L$, we have

$$(A \cap R_\mathbf{k})^h \cap R_\mathbf{x} \subset A^h \cap R_\mathbf{x}.$$

Then by monotonicity of S from Proposition 3,

$$S\left[(A \cap R_\mathbf{k})^h \cap R_\mathbf{x}\right] \subset S\left[A^h \cap R_\mathbf{x}\right] = S[A^h] \cap R_\mathbf{x}.$$

Therefore,

$$\bigcup_{\mathbf{k} \in L} S\left[(A \cap R_\mathbf{k})^h \cap R_\mathbf{x}\right] \subset S[A^h] \cap R_\mathbf{x}. \tag{33}$$

Also by definition of S, we have the following

$$S[A]^h \cap R_\mathbf{x} = \left(\bigcup_{\mathbf{k} \in L} S[A \cap R_\mathbf{k}] \right)^h \cap R_\mathbf{x} = \bigcup_{\mathbf{k} \in L} (S[A \cap R_\mathbf{k}])^h \cap R_\mathbf{x}. \tag{34}$$

Let $\mathscr{D}_h$ be the closed ball centered at origin with radius h in $\mathbb{R}^n$, then h-extension A_h of A can be expressed as $A_h = A + \mathscr{D}_h$. Then for each $\mathbf{k} \in L$, we have:

$$(S[A \cap R_\mathbf{k}])^h \cap R_\mathbf{x} = S[A \cap R_\mathbf{k}] + (\mathscr{D}_h \cap R_{\mathbf{x}-\mathbf{k}}),$$

and

$$S\left[(A \cap R_\mathbf{k})^h \cap R_\mathbf{x}\right] = S[A \cap R_\mathbf{k} + (\mathscr{D}_h \cap R_{\mathbf{x}-\mathbf{k}})].$$

When restricted to the 1-dimensional subspace $R_\mathbf{x}$, $A \cap R_\mathbf{k} + (\mathscr{D}_h \cap R_{\mathbf{x}-\mathbf{k}})$ is just an l-extension of $A \cap R_\mathbf{x} + \mathbf{x} - \mathbf{k}$ for some $l > 0$. Therefore, we can apply the Inclusion (26), and get:

$$(S[A \cap R_\mathbf{k}])^h \cap R_\mathbf{x} = S[A \cap R_\mathbf{k}] + (\mathscr{D}_h \cap R_{\mathbf{x}-\mathbf{k}}) \subset S[A \cap R_\mathbf{k} + (\mathscr{D}_h \cap R_{\mathbf{x}-\mathbf{k}})] = S\left[(A \cap R_\mathbf{k})^h \cap R_\mathbf{x}\right].$$

Together with (33) and (34), we have:

$$S[A^h] \cap R_{\mathbf{x}} \supset \bigcup_{\mathbf{k} \in L} S\left[(A \cap R_{\mathbf{k}})^h \cap R_{\mathbf{x}}\right] \supset \bigcup_{\mathbf{k} \in L} S[A \cap R_{\mathbf{k}}]^h \cap R_{\mathbf{x}} = S[A]^h \cap R_{\mathbf{x}},$$

i.e. for any $\mathbf{x} \in L$

$$S[A^h] \cap R_{\mathbf{x}} \supset S[A]^h \cap R_{\mathbf{x}}. \tag{35}$$

We have proved the Inclusion (26) holds for each slice $R_{\mathbf{x}}$. Taking the union over all $\mathbf{x} \in L$, we can conclude that it holds for all 1-symmetrizations on $\mathbb{R}^n$.

Lemma 7. *If Inclusion (26) holds for k-symmetrizations on $\mathbb{R}^k$, then it also holds for k-symmetrizations on $\mathbb{R}^n$, for all $n \geq k$.*

Proof. Use the proof in step 2 by replacing the 1-dimensional slice with k-dimensional slice in $\mathbb{R}^n$.

Step 3: 2**-symmetrizations on** $\mathbb{R}^2$
We set up a sequence of unit vectors in $\mathbb{R}^2$, say $\{\mathbf{e}_n\}_{n=0}^{\infty}$ as following:

$$\mathbf{e}_0 = (0,1),$$

$$\mathbf{e}_n = \left(\cos\left(\frac{3\pi}{2} + \frac{\pi}{2^n}\right), \sin\left(\frac{3\pi}{2} + \frac{\pi}{2^n}\right)\right),\ n \geq 1.$$

By this construction, $\{\mathbf{e}_n\}_{n=0}^{\infty}$ has the following properties:

1. $\lim_{n\to\infty} \mathbf{e}_n = -\mathbf{e}_0$.
2. $\mathbf{e}_n + \mathbf{e}_0 \in \mathscr{L}\{\mathbf{e}_{n+1}\}^{\perp}$.

The above properties can be verified by direct computation.

Now we define a sequence of 1-symmetrization on $\mathbb{R}^2$ corresponding to $\{\mathbf{e}_n\}_{n=0}^{\infty}$ by

$$S_n = S\left(\mathscr{L}\{\mathbf{e}_n\}^{\perp}, \mathbf{e}_n\right).$$

And define

$$Q_n = S_n S_{n-1} \dots S_1 S_0.$$

Then we want to prove that the sequence $\{Q_n\}$ converges to $Q = S(0, \mathbf{e}_1)$.

Lemma 8. *Let c, c' be positive real numbers, and $n \in \mathbb{Z}_{\geq 0}$. For each closed set $A \subset \mathbb{R}^2$, we have*

$$\boldsymbol{x} + c\boldsymbol{e}_0 + c'\boldsymbol{e}_n \in Q_n[A] \tag{36}$$

for all $\boldsymbol{x} \in Q_n[A]$.

Proof. Fix a closed set $A \subset \mathbb{R}^2$. We will prove the lemma by induction on n.

Base case: when $n = 0$, we have

$$Q_0 = S\left(\mathscr{L}\{\mathbf{e}_0\}^{\perp}, \mathbf{e}_0\right),$$

i.e. Q_0 is an 1-symmetrization on $\mathbb{R}^2$ along $\mathbf{e}_0$.

Let c, $c \in \mathbb{R}$ be positive and pick a point $\mathbf{x} \in Q_n[A]$. From Proposition 3, we know symmetrization S is semi-invariant with respect to $\mathscr{L}\{\mathbf{e}_0\}$. Therefore,

$$x + c\mathbf{e}_0 + c'\mathbf{e}_0 \in Q_0[A] + (c + c')\mathbf{e}_0 \subset Q_0[A].$$

Assume that (36) holds for n, then want to show that it also holds for $n+1$: Define a vector

$$\mathbf{h}_n = \mathbf{e}_0 + \mathbf{e}_n.$$

By properties of sequence $\mathbf{e}_n$, we know

$$\mathbf{h}_n \in \mathscr{L}\{\mathbf{e}_{n+1}\}^{\perp}.$$

Define a line segment joining $r\mathbf{e}_0$ and $r\mathbf{e}_n$ for some $r > 0$ by

$$\triangle_{n,r} = \{tr\mathbf{e}_0 + (1-t)r\mathbf{e}_n |\ 0 \le t \le 1\}.$$

Consider the slices on which the S_{n+1} applies, say:

$$R_\alpha = \alpha\mathbf{h}_n + \mathscr{L}\{\mathbf{e}_{n+1}\},$$

where $\alpha \in \mathbb{R}$. We fix such a layer, say R_α and consider all the points contained in $Q_n[A] \cap R_\alpha$. Define

$$\begin{aligned} B_\alpha &= \bigcup_{\mathbf{x} \in Q_n[A] \cap R_\alpha} \{\mathbf{x} + c\mathbf{e}_0 + c'\mathbf{e}_n |\ c, c' > 0\} \\ &= \{\mathbf{x} + c\mathbf{e}_0 + c'\mathbf{e}_n |\ \mathbf{x} \in Q_n[A] \cap R_\alpha \ \mathit{and}\ c, c' > 0\} \end{aligned}$$

$B_\alpha \subset Q_n[A]$ by induction hypothesis. Consider the slice of B_α cut up by R_β with $\beta > \alpha$. We have

$$S_{n+1}\left[R_\beta \cap B_\alpha\right] \subset S_{n+1}\left(R_\beta \cap Q_n[A]\right) = S_{n+1}\left(Q_n[A]\right) \cap R_\beta = Q_{n+1}[A] \cap R_\beta$$

Geometrically, since vector $\mathbf{h}_n$ half the angle between $\mathbf{e}_0$ and $\mathbf{e}_n$, we have

$$R_\beta \cap B_\alpha = Q_n[A] \cap R_\alpha + (\beta - \alpha)\mathbf{h}_n + \triangle_{n,r}$$

for some $r > 0$. And this is just an extension of $Q_n[A] \cap R_\alpha + (\beta - \alpha)\mathbf{h}_n$ when restricted to the 1-dimensional slice R_β. Apply Inclusion (26) to $R_\beta \cap B_\alpha$ within the slice R_β and we have

$$\begin{aligned}
S_{n+1}\left[R_\beta \cap B_\alpha\right] &= S_{n+1}\left[Q_n[A]\cap R_\alpha + (\beta-\alpha)\mathbf{h}_n + \triangle_{n,r}\right] \\
&\supset S_{n+1}\left[Q_n[A]\cap R_\alpha + (\beta-\alpha)\mathbf{h}_n\right] + \triangle_{n,r} \\
&= S_{n+1}\left[Q_n[A]\cap R_\alpha\right] + \triangle_{n,r} + (\beta-\alpha)\mathbf{h}_n \\
&= S_{n+1}\left[Q_n[A]\right]\cap R_\alpha + \triangle_{n,r} + (\beta-\alpha)\mathbf{h}_n \\
&= Q_{n+1}[A]\cap R_\alpha + \triangle_{n,r} + (\beta-\alpha)\mathbf{h}_n \\
&= \{x + c\mathbf{e}_0 + c'\mathbf{e}_n \mid \mathbf{x}\in Q_{n+1}[A]\cap R_\alpha \textit{ and } c,c'>0\}\cap R_\beta
\end{aligned}$$

Therefore,

$$\{\mathbf{x} + c\mathbf{e}_0 + c'\mathbf{e}_n \mid \mathbf{x}\in Q_{n+1}[A]\cap R_\alpha \textit{ and } c,c'>0\}\cap R_\beta \subset Q_{n+1}[A]\cap R_\beta$$

Take union of all slices R_β, we have

$$\{\mathbf{x} + c\mathbf{e}_0 + c'\mathbf{e}_n \mid \mathbf{x}\in Q_{n+1}[A]\cap R_\alpha \textit{ and } c,c'>0\} \subset Q_{n+1}[A]$$

Remark 5. As n approaches infinity, the angle between vector $\mathbf{e}_0$ and $\mathbf{e}_n$ approaches π. This implies the cone $K_n = \{c\mathbf{e}_0 + c'\mathbf{e} \mid c,c'>0\}$ converges to $\Pi(\mathbf{e}_1,0)$. Fix a closed subset $A\subset\mathbb{R}^n$ and a point $\mathbf{x}\in Q[A]$, then by the lemma above, $(\mathbf{x}+\Pi(\mathbf{e}_1,0))\subset Q[A]$. And $\mathbf{x}$ is arbitrarily chosen, which means that Q[A] has to be a half space along direction $\mathbf{e}_1$. The Gaussian measure of Q[A] is the same as that of A, since symmetrization preserves measure. Therefore, $Q[A]$ and $S(\{0\},\mathbf{e}_1)[A]$ are the same half space, i.e. $Q_n[A]$ converges to $S(\{0\},\mathbf{e}_1)[A]$.

Now we prove the inclusion (26) for 2-symmetrization on $\mathbb{R}^2$.

Lemma 9. *For any closed set $A\subset\mathbb{R}^2$ and any R, $\varepsilon>0$, then for all n large enough, the following holds:*

$$(Q_n[A]^\varepsilon \cap \mathscr{D}_R) \supset (Q[A] \cap \mathscr{D}_R)\,, \tag{37}$$

$$(Q[A]^\varepsilon \cap \mathscr{D}_R) \supset (Q_n[A] \cap \mathscr{D}_R)\,. \tag{38}$$

Proof. Define $K_n = \{c\mathbf{e}_0 + c'\mathbf{e} \mid c,c'>0\}$. Then by (36), we know that for all $\mathbf{x}\in Q_n[A]$,

$$\mathbf{x} + K_n \subset Q_n[A].$$

Now we prove (38) by contradiction:
Suppose $\forall n\geq 0$, we can find a point $\mathbf{x}_n\in\mathbb{R}^2$ such that

$$\mathbf{x}_n \in (Q_n[A] \cap \mathscr{D}_R)\cap(Q[A]^\varepsilon\cap\mathscr{D}_R)^c = Q_n[A]\cap(\mathscr{D}_R/Q[A]^\varepsilon)\,.$$

Apply (36) and get

$$\begin{aligned}
\gamma_2(Q_n[A]) &\geq \gamma_2(\mathbf{x}_n + K_n) \\
&\geq \gamma_2\left(\bigcap_{\mathbf{x}\in\mathscr{D}_R/Q[A]^\varepsilon} (\mathbf{x}+K_n)\right).
\end{aligned}$$

$\{K_n\}$ is increasing and converges to half plane, therefore we have

$$\begin{aligned}\lim_{n\to\infty}\left(\bigcap_{\mathbf{x}\in\mathscr{D}_R/Q[A]^\varepsilon}(\mathbf{x}+K_n)\right) &= \bigcup_n\left(\bigcap_{\mathbf{x}\in\mathscr{D}_R/Q[A]^\varepsilon}(\mathbf{x}+K_n)\right)\\ &= \bigcap_{x\in\mathscr{D}_R/Q[A]^\varepsilon}\left(\mathbf{x}+\bigcup_n K_n\right)\\ &\supset Q[A]^\varepsilon.\end{aligned}$$

Take the limit infimum,

$$\liminf_{n\to\infty}\gamma_2(Q_n[A]) \geq \liminf_{n\to\infty}\ \gamma_2\left(\bigcap_{x\in\mathscr{D}_R/Q[A]^\varepsilon}(x+K_n)\right) \geq \gamma_2\left(Q[A]^\varepsilon\right) > \gamma_2\left(Q[A]\right). \tag{39}$$

On the other hand, since the symmetrization preserves the Gaussian measure, we have:

$$\gamma_2\left(Q_n[A]\right) = \gamma_2\left(A\right) = \gamma_2\left(Q[A]\right),$$

which forms a contradiction to (39).

Therefore, (38) holds. And the proof for (37) is similar to this one.

Now we can prove the Inclusion (26) for 2-symmetrization S on $\mathbb{R}^2$:
For each $n\in\mathbb{N}$,

$$\begin{aligned}Q_n[A]^h &= S_n\left(Q_{n-1}[A]\right)^h\\ &\subset S_n\left(Q_{n-1}[A]^h\right)\\ &\vdots\\ &\subset S_nS_{n-1}\dots S_1S_0[A^h]\\ &= Q_n[A^h].\end{aligned}$$

Next, we generalize this to Q from Q_n: Fix a closed set $A\subset\mathbb{R}^2$, $h>0$, a small $\varepsilon>0$ and also a large $R>0$. Then by (37) and (38)

$$\begin{aligned}\left(Q[A]\cap\mathscr{D}_R\right)^h &\subset \left(Q_n[A]^\varepsilon\cap\mathscr{D}_R\right)^h\\ &\subset Q_n[A]^{\varepsilon+h}\cap\mathscr{D}_{R+h}\\ &\subset Q_n[A^{\varepsilon+h}]\cap\mathscr{D}_{R+h}\\ &\subset Q[A^{\varepsilon+h}]^\varepsilon\cap\mathscr{D}_{R+h},\end{aligned}$$

for n large enough.

Take $R\to\infty$, and get

$$Q[A]^h \subset Q[A^{\varepsilon+h}]^\varepsilon.$$

Then take $\varepsilon\to 0$, and get

$$Q[A]^h \subset Q[A^h].$$

Now we prove the Inclusion (26) holds for a particular 2-symmetrization $Q = S[\{0\},(1,0)]$. But the other symmetrizations are just a rotation of Q and we know Gaussian measure is invariant under rotation. Therefore, the Inclusion (26) holds for all 2-symmetrization in $\mathbb{R}^2$.

Remark 6. By Lemma 7, we know that the Inclusion (26) holds for all 2-symmetrizations on $\mathbb{R}^n$, where $n \geq 2$.

Step 4: *k*-symmetrization on $\mathbb{R}^n$

Lemma 10. *Let M_1, M_2 and M_3 be mutually orthogonal subspaces in $\mathbb{R}^n$, and let $\boldsymbol{e}$ be a vetor orthogonal to M_1, M_2 and M_3. And define symmetrizations in $\mathbb{R}^n$:*

$$S_1 = S(M_1 + M_2, \boldsymbol{e}),$$

and

$$S_2 = S(M_2 + M_3, \boldsymbol{e}).$$

Let A be a closed subset of $\mathbb{R}^n$ such that $S_2[A]$ is also closed, then

$$S_1 S_2[A] = S(M_2, \boldsymbol{e})[A].$$

Proof. Define

$$H = (M_1 + M_2 + M_3 + \mathscr{L}\{\mathbf{e}\})^{\perp}.$$

S_1 is invariant under $(M_1 + M_2 + \mathscr{L}\{\mathbf{e}\})^{\perp}$ by property 4 from Proposition 3, therefore,

$$S_1[A] = S_1[A] + (M_1 + M_2 + \mathscr{L}\{\mathbf{e}\})^{\perp} = S_1[A] + H + M_3. \tag{40}$$

Similarly,

$$S_2[A] = S_2[A] + H + M_1. \tag{41}$$

M_1 is a subspace of $M_1 + M_2$, so S_1 is invariant with respect to M_1 by property 6 from Proposition 3 of symmetrization S_1. Moreover, $S_2[A] = S_2[A] + H + M_1$ implies that $S_2[A]$ is invariant under M_1. Hence,

$$S_1 S_2[A] = S_1 S_2[A] + M_1. \tag{42}$$

Therefore, by equations (40), (41) and (42), we have

$$\begin{aligned} S_1 S_2[A] &= S_1 S_2[A] + M_3 + H \\ &= (S_1 S_2[A] + M_1) + M_3 + H \\ &= S_1 S_2[A] + (M_1 + M_3 + H) \\ &= S_1 S_2[A] + (M_2 + \mathscr{L}\{\mathbf{e}\})^{\perp}. \end{aligned}$$

And apply property 4 from Proposition 3 to symmetrization $S(M_2,\mathbf{e})$,

$$S(M_2,\mathbf{e})[A] = S(M_2,\mathbf{e})[A] + (M_2 + \mathscr{L}\{\mathbf{e}\})^{\perp}.$$

Apply property 5 from Proposition 3 to symmetrizations S_1, S_2, and $S(M_2,\mathbf{e})$, we know that both S_1S_2 and $S(M_2,\mathbf{e})$ are both semi-invariant with respect to $c\mathbf{e}$ for $c > 0$. And they are both invariant with respect to $(M_2 + \mathscr{L}\{\mathbf{e}\})^{\perp}$. This implies that inside each slice $R_{\mathbf{x}} = \mathbf{x} + M_2^{\perp}$, $\mathbf{x} \in M_2$, we have $S(M_2,\mathbf{e})[A] \cap R_{\mathbf{x}}$ and $S_1S_2[A] \cap R_{\mathbf{x}}$ are both half planes with the same unit normal vector $\mathbf{e}$.

Let $k = dimM_2^{\perp}$, and now we observe the Gaussian measure of $S(M_2,\mathbf{e})[A] \cap R_{\mathbf{x}}$ and $S_1S_2[A] \cap R_{\mathbf{x}}$ inside the slice $R_{\mathbf{x}}$:

$$\gamma_k(S(M_2,\mathbf{e})[A] \cap R_{\mathbf{x}}) = \gamma_k(A \cap R_{\mathbf{x}}) = \gamma_k(S_1[A] \cap R_{\mathbf{x}}) = \gamma_k(S_1S_2[A] \cap R_{\mathbf{x}}),$$

by measure preserving of symmetrizations.

Having the same measure implies $S(M_2,\mathbf{e})[A] \cap R_{\mathbf{x}} = S_1S_2[A] \cap R_{\mathbf{x}}$, since they are half spaces with the same unit normal vector. Take the union of all slices $R_{\mathbf{x}}$, and we get

$$S(M_2,\mathbf{e})[A] = S_1S_2[A].$$

Lemma 11. *Let $Q = S(L,\boldsymbol{e})$ be a k-symmetrization in $\mathbb{R}^n$, $n \geq 3$, and $k \geq 2$. Then there exist 2-symmetrizations $Q_1,\ Q_2,\dots,\ Q_{k-1}$ such that*

$$Q[A] = Q_1Q_2\dots Q_{k-1}[A], \tag{43}$$

for all closed set A.

Proof. Fix a closed subset A of $\mathbb{R}^n$.

We prove this lemma by induction on k:

Base case $k = 2$ is automatically true: Q itself is a 2-symmetrization.

Assume (43) holds for k. Then we need to show it also holds for $k+1$. Pick a unit vector $\mathbf{u} \in (L + \mathscr{L}\{\mathbf{e},\mathbf{u}\})^{\perp}$. Then consider the following subspace: $M_1 = \mathscr{L}\{\mathbf{u}\}$, $M_2 = L$, and $M_3 = (L + \mathscr{L}\{\mathbf{e},\mathbf{u}\})^{\perp}$.

We construct two symmetrizations as in previous lemma: $S_1 = S(M_1 + M_2,\mathbf{e}) = S(\mathscr{L}\{\mathbf{u}\} + L,\mathbf{e})$ and $S_2 = S(M_2 + M_3,\mathbf{e}) = S((\mathscr{L}\{\mathbf{e},\mathbf{u}\})^{\perp},\mathbf{e})$, where S_1 is a k-symmetrization on $\mathbb{R}^n$. By induction hypothesis, there exists 2-symmetrizations $Q_1,\ Q_2,\dots,\ Q_{k-1}$ such that

$$S_1[A] = Q_1Q_2\dots Q_{k-1}[A].$$

And S_2 is a 2-symmetrization. Set $Q_k = S_2$.

$S[A]$ is closed for all 2-symmetrizations on $\mathbb{R}^n$ by Lemma 4 applied to 2-symmetrizations. Then by Lemma 10, we have

$$Q[A] = S(M_2,e)[A] = S_1S_2[A] = Q_1Q_2\dots Q_{k-1}Q_k[A].$$

Now for any k-symmetrization Q on $\mathbb{R}^n$, ≥ 3, we can find 2-symmetrizations $Q_1,\ Q_2,\ldots,\ Q_{k-1}$ such that

$$Q[A] = Q_1 Q_2 \ldots Q_{k-1}[A],$$

for all closed subset A of $\mathbb{R}^n$.

Therefore, fix a closed set $A \subset \mathbb{R}^n$ and $h > 0$, we have

$$\begin{aligned} Q[A^h] = Q_1 Q_2 \ldots Q_{k-1}[A^h] &\supset Q_1 Q_2 \ldots Q_{k-2}(Q_{k-1}[A]^h) \\ &\vdots \\ &\supset Q_1 Q_2 \ldots Q_{k-1}[A]^h \\ &= Q[A]^h. \end{aligned}$$

This completes the proof for Inclusion (26).

Appendix

Proof of Proposition 2

We will work backwards: at each step we look for a sufficient condition for (4) to hold. Given $a, b \in (0,1)$, let $c = (a+b)/2$, $x = (a-b)/2$, then $a, b \in (0,1)$ if and only if $x \in (-\min(c, 1-c), \min(c, 1-c))$. Let $g(x) = I(x+c)^2 + x^2$. It is clear that (4) is equivalent to

$$\sqrt{g(0)} \leq \frac{1}{2}\sqrt{g(x)} + \frac{1}{2}\sqrt{g(-x)}. \tag{44}$$

Multiplying by 2 and taking the square, we have:

$$4g(0) - (g(x) + g(-x)) \leq 2\sqrt{g(x)g(-x)}. \tag{45}$$

If the left-hand side is negative, then we are done. Otherwise take the square again, after all the cancellation, we have

$$16g(0)^2 + \big(g(x) - g(-x)\big)^2 \leq 8g(0)\big(g(x) + g(-x)\big). \tag{46}$$

Now define $h(x) = g(x) - g(0) = I(c+x)^2 + x^2 - I(c)^2$. We can rewrite (28) as

$$\big(h(x) - h(-x)\big)^2 \leq 8I(c)^2\big(h(x) + h(-x)\big). \tag{47}$$

To prove (29) is true, we will use the following facts.

Lemma 12. *(a) $I \cdot I'' = -1$. (b) The function $(I')^2$ is convex on $(0,1)$*

Proof. (a) and (b) can be derived by directly computing the first and second derivatives. $\square$

Lemma 13. *Let $R(x) = h(x) + h(-x) - 2I'(c)^2x^2$, then $R(x)$ has non-negative second derivative on $(-\min(c, 1-c), \min(c, 1-c))$, and therefore it is convex on $(-\min(c, 1-c), \min(c, 1-c))$.*

Proof. We compute the second derivative of R, by the convexity of $(I')^2$ proven in lemma 4:

$$R'' = 4\Big[\frac{I'(c+x)^2 + I'(c-x)^2}{2}\Big] - I'(c)^2 \geq 0$$

□

Since R is even, and $R(0) = 0$, by lemma 5, we have $R(x) \geq 0$ for all $x \in (-\min(c, 1-c), \min(c, 1-c))$. That is,

$$h(x) + h(-x) \geq 2I'(c)^2x^2,$$

which is equivalent to:

$$8I(c)^2\big(h(x) + h(-x)\big) \geq 16I(c)^2I'(c)^2x^2 \tag{48}$$

Therefore, to prove (29), it suffices to show that the right-hand side of (30) is at least $(h(x) - h(-x))^2$, that is,

$$\left|\frac{h(x) - h(-x)}{x}\right| \leq 4I(c)\big|I'(c)\big| \tag{49}$$

Recall that $h(x) = I(c+x)^2 + x^2 - I(c)^2$, so $h(x) - h(-x) = I(c+x)^2 - I(c-x)^2$, so (31) is equivalent to

$$\left|\frac{I(c+x)^2 - I(c-x)^2}{x}\right| \leq 4I(c)\big|I'(c)\big| \tag{50}$$

Since both I and $|I'|$ are symmetric around $1/2$, we can assume without loss of generality that $c \in (0, 1/2)$ (otherwise we can replace c with $1-c$ and both sides in (32) remain unchanged). Moreover, we can assume $x \geq 0$, because (32) is an even function of x (otherwise we can replace x with $-x$). With these assumptions, all the terms in (32) inside absolute value are positive, so (32) is equivalent to,

$$\frac{I(c+x)^2 - I(c-x)^2}{x} \leq 4I(c)I'(c) \tag{51}$$

under the assumption that $0 \leq x < c \leq 1/2$. Let $u(x) = I(c+x)^2 - I(c-x)^2$. Using lemma 4(a), one can find the second derivative of u to be

$$u''(x) = 2\big(I'(c+x)^2 - I'(c-x)^2\big)$$

It is clear that $(I')^2$ increases on $(0, 1/2)$, decreases on $(1/2, 1)$, and is symmetric around $1/2$, so $I'(c+x)^2 \leq I'(c-x)^2$, $u''(x) \leq 0$. Therefore u is a concave non-negative function on $[0, c]$, and

$$\frac{u(x)}{x} = \int_0^1 u'(xt)dt$$

is non-increasing for $x \in [0, c)$ (to see this, take the derivative on the right-hand side and switch derivative and integral). Finally, using Taylor expansion

$$I(c+x)^2 = I(c)^2 + 2I(c)I'(c)x + O(x^2).$$

For each $x \in (0, c]$,

$$\frac{u(x)}{x} \leq \lim_{t \to 0} \frac{u(t)}{t} = \lim_{t \to 0} \frac{4I(c)I'(c)t + O(t^2)}{t} = 4I(c)I'(c) \tag{52}$$

□

References

1. Bobkov, S.G.: An Isoperimetric Inequality On the Discrete Cube, and An Elemetary Proof of the Isoperimetric Inequality In Gauss Space. The Annals of Probability (1997)
2. Lifshits M.A.: Gaussian Random Functions. Kluwer Academic Publishers (1995)