# Isoperimetry and functional inequalities

Piotr Nayar, Tomasz Tkocz

April 9, 2012

# **1** Isoperimetric inequality

### 1.1 Brunn-Minkowski inequality

**1.1 Theorem.** (BRUNN-MINKOWSKI, '88) If A and B are non-empty compact sets then for all  $\lambda \in [0, 1]$  we have

$$\operatorname{vol}\left((1-\lambda)A + \lambda B\right)^{1/n} \ge (1-\lambda)(\operatorname{vol} A)^{1/n} + \lambda(\operatorname{vol} B)^{1/n}.$$
 (B-M)

Note that if either  $A = \emptyset$  or  $B = \emptyset$ , this inequality does not hold since  $(1-\lambda)A + \lambda B = \emptyset$ . We can use the homogenity of volume to rewrite Brunn-Minkowski inequality in the form

$$\operatorname{vol}(A+B)^{1/n} \ge (\operatorname{vol} A)^{1/n} + (\operatorname{vol} B)^{1/n}.$$
 (1.1)

We can deduce from this inequality the isoperimetric inequality.

**1.2 Theorem.** Among sets with prescribed volume, the Euclidean balls are the one with minimum surface area.

*Proof.* We can assume that C is compact and vol  $C = \text{vol } B_2^n$ . We have

$$\operatorname{vol} \partial C = \liminf_{\varepsilon \to 0} \frac{\operatorname{vol}(C + \varepsilon B_2^n) - \operatorname{vol}(C)}{\varepsilon}$$

By the Brunn-Minkowski inequality we get

$$\operatorname{vol}(C + \varepsilon B_2^n)^{1/n} \ge (\operatorname{vol} C)^{1/n} + \varepsilon (\operatorname{vol} B_2^n)^{1/n},$$

hence

$$\operatorname{vol}(C + \varepsilon B_2^n) \ge (1 + \varepsilon)^n \operatorname{vol} C,$$

 $\mathbf{SO}$ 

$$\operatorname{vol}(\partial C) \ge \liminf_{\varepsilon \to 0} \frac{((1+\varepsilon)^n - 1)\operatorname{vol}(C)}{\varepsilon} = n\operatorname{vol}(C) = n\operatorname{vol}(B_2^n) = \operatorname{vol}(\partial B_2^n).$$

There is an a priori weaker statement of the Brunn-Minkowski inequality. Using A - G mean inequality we get

$$|(1-\lambda)A + \lambda B| \ge |A|^{1-\lambda}|B|^{\lambda}, \qquad \lambda \in [0,1].$$

Note that this inequality is valid for any compact sets A and B (the asumption that A and B are non-empty is no longer needed). We can see that there is no appearance of dimension in this expression.

The strong version of the Brunn-Minkowski inequality tells us that the Lebesgue measure is a  $\frac{1}{n}$ -concave measure. The weaker statement tells us that it is a log-concave measure.

**1.3 Definition.** A measure  $\mu$  on  $\mathbb{R}^n$  is log-concave if for all compact sets A and B we have

$$\mu((1-\lambda)A + \lambda B) \ge \mu(A)^{1-\lambda}\mu(B)^{\lambda}, \qquad \lambda \in [0,1].$$

**1.4 Definition.** The function  $f : \mathbb{R}^n \to \mathbb{R}$  is log-concave if for all  $x, y \in \mathbb{R}^n$  we have

$$f((1-\lambda)x+\lambda y) \ge f(x)^{1-\lambda}f(y)^{\lambda}, \qquad \lambda \in [0,1].$$

Note that this definitions are dimension free.

The weak form of the (B-M) inequality for the Lebesgue measure is equivalent to (B-M) inequality. It is a consequence of the homogenity of the Lebesgue measure. Indeed, if

$$\mu = \frac{\lambda(\operatorname{vol} B)^{1/n}}{(1-\lambda)(\operatorname{vol} A)^{1/n} + \lambda(\operatorname{vol} B)^{1/n}}$$

then

$$\operatorname{vol}\left(\frac{(1-\lambda)A + \lambda B}{(1-\lambda)(\operatorname{vol} A)^{1/n} + \lambda(\operatorname{vol} B)^{1/n}}\right) = \operatorname{vol}\left((1-\mu)\frac{A}{(\operatorname{vol} A)^{1/n}} + \mu\frac{B}{(\operatorname{vol} B)^{1/n}}\right)$$
$$\geq \operatorname{vol}\left(\frac{A}{(\operatorname{vol} A)^{1/n}}\right)^{1-\mu}\left(\frac{B}{(\operatorname{vol} B)^{1/n}}\right)^{\mu} = 1.$$

#### 1.2 Functional version of Brunn-Minkowski inequality

If we take  $f = I_A$ ,  $g = I_B$  and  $m = I_{(1-\lambda)A+\lambda B}$  then (B-M) says that

$$\int m \ge \left(\int f\right)^{1-\lambda} \left(\int g\right)^{\lambda}$$

and obviously m, f, g satisfies

$$m((1-\lambda)x + \lambda y) \ge f(x)^{1-\lambda}g(y)^{\lambda}.$$

We are to prove the following functional version of (B-M) inequality (which will give the proof of (B-M) inequality).

**1.5 Theorem.** (PREKOPA-LEINDLER, '88) Let f, g, m be nonnegative measurable functions on  $\mathbb{R}^n$  and let  $\lambda \in [0, 1]$ . If for all  $x, y \in \mathbb{R}^n$  we have

$$m((1-\lambda)x + \lambda y) \ge f(x)^{1-\lambda}g(y)^{\lambda},$$

then

$$\int_{\mathbb{R}^n} m \ge \left(\int_{\mathbb{R}^n} f\right)^{1-\lambda} \left(\int_{\mathbb{R}^n} g\right)^{\lambda}.$$
(1.2)

*Proof. Step 1.* We start with proving (B-M) inequality in dimension 1. Let A, B be compact sets in  $\mathbb{R}$ . Observe that the operations  $A \to A+v_1$ ,  $B \to B+v_2$  where  $v_1, v_2 \in \mathbb{R}$  does not change the volumes of A, B and  $(1 - \lambda)A + \lambda B$  (adding a number to one of the sets only shifts all of this sets). Therefore we can assume that  $\sup A = \inf B = 0$ . But then, since  $0 \in A$  and  $0 \in B$ , we have

$$(1 - \lambda)A + \lambda B \supset (1 - \lambda)A \cup (\lambda B).$$

But  $(1 - \lambda)A$  and  $(\lambda B)$  are disjoint, up to the one point 0. Therefore

$$|(1-\lambda)A + \lambda B| \ge |(1-\lambda)A| + |\lambda B|,$$

hence we have proved (B-M) in dimension 1.

Step 2. Let us now justify the Prekopa-Leindler inequality in dimension 1. We can assume, considering  $fI_{f\leq M}$  and  $gI_{g\leq M}$  instead of f and g, that f, g are bounded. Note also that this inequality possesses some homogenity. Indeed, if we multiply f, g, m by numbers  $c_f, c_g, c_m$  satisfying

$$c_m = c_f^{1-\lambda} c_g^{\lambda},$$

then the hyphotesis and the thesis do not change. Therefore, taking  $c_f = ||f||_{\infty}^{-1}$ ,  $c_g = ||g||_{\infty}^{-1}$  and  $c_m = ||f||_{\infty}^{-(1-\lambda)} ||g||_{\infty}^{-\lambda}$  we can assume (since we are in the situation when f and g are bounded) that  $||f||_{\infty} = ||g||_{\infty} = 1$ . But then

$$\int_{\mathbb{R}} m = \int_{0}^{+\infty} |\{m \ge s\}| \, \mathrm{d}s,$$
$$\int_{\mathbb{R}} f = \int_{0}^{1} |\{f \ge r\}| \, \mathrm{d}r,$$
$$\int_{\mathbb{R}} g = \int_{0}^{1} |\{g \ge r\}| \, \mathrm{d}r.$$

Note also that if  $x \in \{f \ge r\}$  and  $y \in \{g \ge r\}$  then by the assumption of the theorem we have  $(1 - \lambda)x + \lambda y \in \{m \ge r\}$ . Hence,

$$(1-\lambda)\{f \ge r\} + \lambda\{g \ge r\} \subset \{m \ge r\}.$$

Moreover, the sets  $\{f \ge r\}$  and  $\{g \ge r\}$  are non-empty for  $r \in [0, 1)$ . This is very important since we want to use 1 dimensional (B-M) inequality! We have

$$\begin{split} \int m &= \int_0^{+\infty} |\{m \ge r\}| \, \mathrm{d}r \ge \int_0^1 |\{m \ge r\}| \, \mathrm{d}r \ge \int_0^1 |(1-\lambda)\{f \ge r\} + \lambda\{g \ge r\}| \, \mathrm{d}r \\ &\ge (1-\lambda) \int_0^1 |\{f \ge r\}| \, \mathrm{d}r + \lambda \int_0^1 |\{g \ge r\}| \, \mathrm{d}r = (1-\lambda) \int f + \lambda \int g \\ &\ge \left(\int f\right)^{1-\lambda} \left(\int g\right)^{\lambda}. \end{split}$$

Observe that we have proved

$$\int m \ge (1-\lambda) \int f + \lambda \int g,$$

but this inequality does not have the previous homogenity, hence it requires the assumption  $\|f\|_{\infty} = \|g\|_{\infty} = 1$ .

Step 3 (the inductive step). Suppose our inequality in true in dimension n-1. We will prove it in dimension n. Suppose we have a numbers  $y_0, y_1, y_2 \in \mathbb{R}$  satisfying  $y_0 = (1 - \lambda)y_1 + \lambda y_2$ . Define  $m_{y_0}, f_{y_1}, g_{y_2} : \mathbb{R}^{n-1} \to \mathbb{R}_+$  by

$$m_{y_0}(x) = m(y_0, x), \quad f_{y_1}(x) = f(y_1, x), \quad g_{y_2}(x) = (y_2, x),$$

where  $x \in \mathbb{R}^{n-1}$ . Note that since  $y_0 = (1 - \lambda)y_1 + \lambda y_2$  we have

$$m_{y_0}((1-\lambda)x_1 + \lambda x_2) = m((1-\lambda)y_1 + \lambda y_2, (1-\lambda)x_1 + \lambda x_2)$$
  

$$\geq f(y_1, x_1)^{1-\lambda}g(y_2, x_2)^{\lambda} = f_{y_1}(x_1)^{1-\lambda}g_{y_2}(x_2)^{\lambda},$$

hence  $m_{y_0}, f_{y_1}$  and  $g_{y_2}$  satisfies the assumption of the (n-1)-dimensional Prekopa-Leindler inequality. Therefore we have

$$\int_{\mathbb{R}^{n-1}} m_{y_0} \ge \left(\int_{\mathbb{R}^{n-1}} f_{y_1}\right)^{1-\lambda} \left(\int_{\mathbb{R}^{n-1}} g_{y_2}\right)^{\lambda}.$$

Step 4. Define new functions  $M, F, G : \mathbb{R} \to \mathbb{R}_+$ 

$$M(y_0) = \int_{\mathbb{R}^{n-1}} m_{y_0}, \quad F(y_1) = \int_{\mathbb{R}^{n-1}} f_{y_1}, \quad G(y_2) = \int_{\mathbb{R}^{n-1}} g_{y_2}$$

We have seen (the above inequality) that when  $y_0 = (1 - \lambda)y_1 + \lambda y_2$  then there holds

$$M((1-\lambda)y_1+\lambda y_2) \ge F(y_1)^{1-\lambda}G(y_2)^{\lambda}.$$

Hence, by 1-dimensional (P-L) inequality we get

$$\int_{\mathbb{R}} M \ge \left(\int_{\mathbb{R}} F\right)^{1-\lambda} \left(\int_{\mathbb{R}} G\right)^{\lambda}.$$

But

so we shown that

$$\int_{\mathbb{R}} M = \int_{\mathbb{R}^n} m, \quad \int_{\mathbb{R}} F = \int_{\mathbb{R}^n} f, \quad \int_{\mathbb{R}} G = \int_{\mathbb{R}^n} g,$$
$$\int_{\mathbb{R}^n} m \ge \left( \int_{\mathbb{R}^n} f \right)^{1-\lambda} \left( \int_{\mathbb{R}^n} g \right)^{\lambda}.$$

	п.
	н

# 1.3 Isoperimetric problem

The Brunn-Minkowski inequality yields the isoperimetric inequality for the Lebesgue measure on  $\mathbb{R}^n$ . Indeed, suppose we have a compact set  $A \subset \mathbb{R}^n$  and let B be a Euclidean ball of the radius  $r_A$  such that |B| = |A|. Then from the Brunn-Minkowski inequality we have

$$|A_{\varepsilon}|^{1/n} = |A + \varepsilon B_2^n|^{1/n} \ge |A|^{1/n} + |\varepsilon B_2^n|^{1/n}$$
  
=  $|B_2^n|^{1/n} r_A + |B_2^n|^{1/n} \varepsilon = |B + \varepsilon B_2^n|^{1/n} = |B_{\varepsilon}|^{1/n}$ 

It means that

$$|A_{\varepsilon}| \ge (r_A + \varepsilon)^n |B_2^n|$$

and therefore

$$|\partial A| \ge nr_a^{n-1}|B_2^n| = n\left(\frac{|A|}{|B_2^n|}\right)^{\frac{n-1}{n}} = n|B_2^n|^{1/n}|A|^{\frac{n-1}{n}}.$$

Now we can state an isoperimetric problem.

**Isoperimetric problem** Let  $(\Omega, d)$  be a metric space and let  $\mu$  be a Borel measure on  $\Omega$ . Let  $\alpha > 0$  and  $\varepsilon > 0$ . We set

$$A_{\varepsilon} = \{ x \in \Omega \mid d(x, A) \le \varepsilon \}.$$

What are the sets  $A \subset \Omega$  of the measure  $\alpha$  such that

$$\mu(A_{\varepsilon}) = \inf_{\mu(B)=\alpha} \mu(B_{\varepsilon}).$$

This problem is very difficult in general.

# **1.4** Co-area formula and $L_1$ Sobolev inequality

Let (X, d) be a metric space. Define its gradient

$$|\nabla f(x)| = \limsup_{d(x,y) \to 0^+} \frac{|f(x) - f(y)|}{d(x,y)}.$$

If x is isolated then we set  $|\nabla f(x)| = 0$ . Let  $\mu$  be a Borel probability measure on X. Then we have the following lemma

**1.6 Lemma** (co-area formula). If  $f: X \to \mathbb{R}$  is a Lipschitz function then

$$\int_{X} |\nabla f(x)| \, \mathrm{d}\mu(x) \ge \int_{-\infty}^{+\infty} \mu^+ \left( \{ x \in X : \ f(x) > t \} \right) \, \mathrm{d}t.$$

*Proof.* We can assume that f > 0. Let  $A_t = \{x \in X : f(x) > t\}$  and define

$$f_h(x) = \sup_{d(x,y) < h} f(y).$$

Then  $\{x \in X : f_h(x) > t\} = A_t^h$ . Since

$$|\nabla f(x)| \ge \limsup_{y \to x} \frac{f(y) - f(x)}{d(x, y)} = \limsup_{h \to 0^+} \frac{f_h(x) - f(x)}{h},$$

we have

$$\begin{split} \int_{X} |\nabla f(x)| \, \mathrm{d}\mu(x) &\geq \int_{X} \limsup_{h \to 0^{+}} \frac{f_h - f}{h} \, \mathrm{d}\mu \geq \limsup_{h \to 0^{+}} \int_{X} \frac{f_h - f}{h} \, \mathrm{d}\mu \\ &= \limsup_{h \to 0^{+}} \int_{0}^{+\infty} \frac{\mu(A_t^h) - \mu(A_t)}{h} \, \mathrm{d}t \geq \liminf_{h \to 0^{+}} \int_{0}^{+\infty} \frac{\mu(A_t^h) - \mu(A_t)}{h} \, \mathrm{d}t \\ &\geq \int_{0}^{+\infty} \liminf_{h \to 0^{+}} \frac{\mu(A_t^h) - \mu(A_t)}{h} \, \mathrm{d}t = \int_{0}^{+\infty} \mu^+(A_t) \, \mathrm{d}t. \end{split}$$

**1.7 Theorem.** Let  $f : \mathbb{R}^n \to \mathbb{R}$  be locally Lipschitz function with compact support with  $|\nabla f| \in L_1(\mathbb{R}^n)$ . Then

$$||f||_{\frac{n}{n-1}} \le \frac{1}{n|B_2^n|^{1/n}} \int_{\mathbb{R}^n} |\nabla f|.$$

*Proof.* We know that for any Borel set of finite measure we have

$$|A^+| \ge n |B_2^n|^{1/n} |A|^{\frac{n-1}{n}}.$$

Observe that  $||f(x)| - |f(y)|| \le |f(x) - f(y)|$ . Therefore, taking |f| instead of f we can assume that f is nonnegative. Hence, by co-area formula we have

$$\int_{\mathbb{R}^n} |\nabla f| \ge \int_0^{+\infty} |\{f > t\}^+| \, \mathrm{d}t \ge \int_0^{+\infty} n |B_2^n|^{1/n} |\{f > t\}|^{\frac{n-1}{n}} \, \mathrm{d}t.$$

Note that the sets  $\{f > t\}$  are of finite measure since f is compactly supported. Now it suffices to note that

$$\int_{0}^{+\infty} |\{f > t\}|^{\frac{n-1}{n}} dt = \int_{0}^{+\infty} \left( \int_{\mathbb{R}^{n}} \left( \mathbf{1}_{\{f > t\}} \right)^{\frac{n}{n-1}} \right)^{\frac{n-1}{n}} dt$$
$$= \int_{0}^{+\infty} \left\| \mathbf{1}_{\{f > t\}} \right\|_{\frac{n}{n-1}} dt \ge \left\| \int_{0}^{+\infty} \mathbf{1}_{\{f > t\}} dt \right\|_{\frac{n}{n-1}} = \left\| f \right\|_{\frac{n}{n-1}}.$$

Using this inequality with the function  $f_{\varepsilon}(x) = \max\{1 - \frac{d(x,A)}{\varepsilon}, 0\}$  and taking  $\varepsilon \to 0$  we recover the classical isoperimetric inequality.

#### 1.5 Cheeger inequality

We consider a triple  $(M, d, \mu)$ . Recall that

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

where  $A^h = \{x \in X : \exists_{a \in A} \ d(x, a) < h\}$ . W say that  $\mu$  satisfies Cheeger inequality with constant h > 0 if

$$\mu^+(A) \ge h \min\{\mu(A), 1 - \mu(A)\}.$$

We show that this is equivalent to

$$h \int_{M} |f - \operatorname{Med}_{\mu} f| \, \mathrm{d}\mu \leq \int_{M} |\nabla f| \, \mathrm{d}\mu,$$

so the best constants in this inequalities are the same. Indeed, using the standard approximation of the indicator function of a set A and applying the above inequality we arrive at

$$\mu^+(A) \ge h \min\{\mu(A), 1 - \mu(A)\}$$

since  $\mu(A) \ge \frac{1}{2}$  implies  $\operatorname{Med}_{\mu} \mathbf{1}_{A} = 1$  and  $\mu(A) \le \frac{1}{2}$  implies  $\operatorname{Med}_{\mu} \mathbf{1}_{A} = 0$ . We prove that the Cheeger inequality implies the functional version. We can assume that  $\operatorname{Med}_{\mu} f = 0$ .

We have

$$\int |\nabla f| \, \mathrm{d}\mu \ge \int_{-\infty}^{+\infty} \mu^+ \left( \{x : \ f(x) > t\} \right) \, \mathrm{d}t \ge h \int \min\{\mu(f > t), 1 - \mu(f > t)\} \, \mathrm{d}t$$
$$= h \int_{-\infty}^{0} 1 - \mu(f > t) \, \mathrm{d}t + h \int_{0}^{+\infty} \mu(f > t) \, \mathrm{d}t$$
$$= h \int_{0}^{+\infty} \mu(f \le -t) \, \mathrm{d}t + h \int_{0}^{+\infty} \mu(f > t) \, \mathrm{d}t = h \int_{0}^{+\infty} \mu(|f| > t) \, \mathrm{d}t$$
$$= h \int |f| \, \mathrm{d}\mu.$$

Now we consider a situation when we have o product structure  $(M^n, d_n, \mu^n)$ , where  $d_n^2(x, y) = \sum_{i=1}^n d(x_i, y_i)^2$ . We assume that

$$|\nabla f(x)|^2 = \sum_{i=1}^n |\nabla_{x_i} f(x)|^2$$

Moreover, let  $\nu$  be the exponential distribution on  $\mathbb{R}$ , i.e. the measure with density  $\frac{1}{2}e^{-|x|}$ . Our goal is to prove the following theorem

**1.8 Theorem.** For a triple  $(M^n, d_n, \mu^n)$  we have

$$h_{\mu^n} \ge \frac{1}{2\sqrt{6}} h_{\mu}$$

If  $\mu = \nu$  then we get  $h_{\nu^n} \ge \frac{1}{2\sqrt{6}}$ , since  $h_{\nu} = 1$ .

No we prove that  $h_{\nu} = 1$ .

**1.9 Theorem.** If  $\phi(x) = \frac{1}{2}e^{-|x|}$ ,  $\Phi(x) = \int_{-\infty}^{x} \phi(s) ds = \nu((-\infty, x])$  then  $\nu(A) = \Phi(a)$  implies  $\nu(A^u) \ge \Phi(a+u)$ .

*Proof.* Since  $A^h = (A^{\varepsilon})^{h-\varepsilon}$  we can assume that A is open. Open sets in  $\mathbb{R}$  are the sums of open intervals, which can be approximated by finite sums of closed intervals from inside. We can construct a sequence of sets  $A_n$  which are finite sums of intervals such that  $A \supset A_n$  and  $\nu(A_n) \rightarrow \nu(A)$  as  $n \rightarrow \infty$ . If we can prove out theorem for  $A_n$  then

$$\nu(A^u) \ge \nu(A^u_n) \ge \Phi(a_n + u) \to \Phi(a + u),$$

where  $a_n$  is such that  $\nu(A_n) = \Phi(a_n)$ . Therefore we assume that A is a finite sum of disjoint closed intervals.

The proof is based on the induction on the number n of bounded intervals in A. If n = 0 then A can be a half line or a sum of two half lines. In the first case the proposition is trivial. In the second case we have  $A = (-\infty, a] \cup [b, \infty)$  where a < b. We want to prove that  $\Phi(a) + 1 - \Phi(b) = \Phi(c)$  implies  $\Phi(a + u) + 1 - \Phi(b - u) \ge \Phi(c + u)$ . Since  $1 - \Phi(x) = \Phi(-x)$ , we have to prove that  $\Phi(x_0) + \Phi(y_0) = \Phi(z)$  implies  $\Phi(x_0 + u) + \Phi(y_0 + u) \ge \Phi(z + u)$ . Fix  $z \in \mathbb{R}$  and assume, without loss of generality, that  $x_0 \le y_0$ . If  $\Phi(x) + \Phi(y(x)) = \Phi(z)$  and  $\Phi(x_0) + \Phi(y_0) = \Phi(z)$  then  $y(x_0) = y_0$  and if  $x < x_0$  then  $y > y_0 \ge x_0 > x$ . Differentiating the constraint we arrive at  $\phi(y)y' = -\phi(x)$ . Note that

$$\frac{\mathrm{d}}{\mathrm{d}x} \left( \Phi(x+u) + \Phi(y+u) \right) = \phi(y+u)y' + \phi(x+u) = \phi(y+u)\frac{-\phi(x)}{\phi(y)} + \phi(x+u)$$

the sign of this expression is the same as the sign of

$$-e^{-|y+u|-|x|} + e^{-|x+u|-|y|}$$

Note that if  $x \to -\infty$  then  $y \to z$ . Therefore it is enough to show that

 $-e^{-|y+u|-|x|} + e^{-|x+u|-|y|} > 0$ 

and hence, to show that  $|x + u| - |x| \le |y + u| - |y|$ . But this is clear since x < y and  $s \to |s + u| - |s|$  is nondecreasing.

Now we consider the case when we have n + 1 bounded intervals and  $n \ge 0$ . Assume that  $A = B \cup I$ , where I = [u, v] and B consists of at most n bounded intervals. If  $B^u \cap I^u \ne \emptyset$  then there exists an interval J such that  $J^u \cap I^u \ne \emptyset$ . If we take  $\operatorname{conv}(I, J)$  instead of I and J then we have a set A' with at most n bounded intervals such that  $\nu((A')^u) = \nu(A^u)$  and  $\nu(A') > \nu(A)$ . From the induction hypothesis we have  $\nu(A') = \Phi(a')$  implies  $\nu((A')^u) \ge \Phi(a'+u)$ . Therefore

$$\nu(A^u) = \nu((A')^u) \ge \Phi(a'+u) \ge \Phi(a+u),$$

since  $\nu(A') > \nu(A)$  implies a' > a.

If  $B^u \cap I^u = \emptyset$  then we show that we can move I to obtain the situation  $B^u \cap I^u \neq \emptyset$ or to make the interval I infinite. More precisely, we show that if  $u + v \leq 0$  then we can move I to the left (in such a way that the measure is preserved). Of course, by symmetry, if  $u + v \geq 0$  then we can move I to the right. If is enough to show that if v < w,  $v + w \leq 0$  and v' < v, w' < w are such that  $\nu([v,w]) = \nu([v',w'])$ then  $\nu([v - u, w + u]) \geq \nu([v' - u, w' + u])$ . Let  $\nu([x, w(x)]) = \Phi(w(x)) - \Phi(x)$  be constant and  $x + w(x) \leq 0$ . We have  $\phi(w(x))w'(x) = \phi(x)$ . It suffices to show that  $x \mapsto \Phi(w(x) + u) - \Phi(x - u)$  is nondecreasing. It is equivalent to

$$\phi(w(x) + u)w'(x) \ge \phi(x - u)$$

and therefore, equivalent to

$$\frac{1}{\phi(w(x))}\left(\phi(x)\phi(w(x)+u)-\phi(w(x))\phi(x-u)\right) \ge 0$$

We are to show that

$$|w(x) + u| + |x| \le |w(x)| + |x - u|.$$

We have  $x \leq 0$  and  $x - u \leq 0$  therefore it suffices to prove  $|w(x) + u| \leq |w(x)| + u$ , which is the triangle inequality.

This theorem yields  $h_{\nu} = 1$ . Indeed, if  $\nu(A) = \Phi(a)$  then we have

$$\nu^{+}(A) = \liminf_{h \to 0^{+}} \frac{\nu(A^{h}) - \nu(A)}{h} \ge \liminf_{h \to 0^{+}} \frac{\Phi(a+h) - \Phi(a)}{h} = \Phi'(a) = \Phi'\left(\Phi^{-1}(\nu(A))\right).$$

If  $\nu(A) \leq \frac{1}{2}$  then  $\nu(A) = \Phi(a) = \frac{1}{2}e^a$  for  $a \leq 0$  and therefore

$$\Phi'\left(\Phi^{-1}(\nu(A))\right) = \Phi'\left(\Phi^{-1}(\Phi(a))\right) = \Phi'(a) = \frac{1}{2}e^{-|a|} = \nu(A).$$

If  $\nu(A) \ge \frac{1}{2}$  then  $\nu(A) = \Phi(a) = 1 - \frac{1}{2}e^{-a}$  for  $a \ge 0$  and hence

$$\Phi'\left(\Phi^{-1}(\nu(A))\right) = \Phi'\left(\Phi^{-1}(\Phi(a))\right) = \Phi'(a) = \frac{1}{2}e^{-|a|} = 1 - \nu(A)$$

We can provide a simple description of the isoperimetric constant for a general measure on the real line.

**1.10 Theorem.** Suppose  $\mu$  is a Borel measure on  $\mathbb{R}$  which is not a Dirac delta. Let  $F(x) = \mu((-\infty, x])$  and let p be a density of the absolutely continuous part of  $\mu$ . Set  $a = \inf\{x : F(x) > 0\}, b = \sup\{x : F(x) < 1\}$ . Let

$$K_{\mu} = \operatorname{essinf}_{a < x < b} \left( \frac{p(x)}{\min\{F(x), 1 - F(x)\}} \right).$$

Let U be a nondecreasin, left-continuous function which transports the exponential measure  $\nu$  onto  $\mu$ . Then

$$h_{\mu} = K_{\mu} = \frac{1}{\|U\|_{Lip}}.$$

*Proof.* For simplicity we give a proof in the case of measures with positive densities on the real line. For Borel A we have  $\mu(A) = \nu(U^{-1}(A))$ . Hence

$$F(a) = \mu((-\infty, a]) = \nu(U^{-1}((-\infty, a])) = \nu((-\infty, U^{-1}(a)]) = F_{\nu}(U^{-1}(a)).$$

Therefore  $U(a) = F^{-1}(F_{\nu}(a))$  and if  $V = U^{-1}$  then  $V(a) = F_{\nu}^{-1}(F(a))$ . We have

$$V'(a) = \frac{F(a)}{F'_{\nu}(F^{-1}_{\nu}(F(a)))} = \frac{p(a)}{\min\{F(a), 1 - F(a)\}} \ge K_{\mu}.$$

Therefore  $V(y) - V(x) \ge K_{\mu}(y-x)$ , so  $y-x \ge K_{\nu}(U(y) - U(x))$ . This implies  $||U||_{Lip} \le 1/K_{\mu}$ .

Now let  $f = g \circ U$ ,  $f' = (g' \circ U)U'$ , hence  $|f'| \leq ||U||_{Lip} |g'(U)|$ . Let  $\xi \sim \nu$ . Then  $\eta = U(\xi) \sim \mu$ . When Med  $f(\xi) = 0$  then  $\mathbb{E}|f(\xi)| \leq \mathbb{E}|f'(\xi)|$ , because we know that  $h_{\nu} = 1$ . We have  $f(\xi) = g(\eta)$ . Therefore Med  $g(\eta) = 0$  implies

$$\mathbb{E}|g(\eta)| \le \mathbb{E}|(g \circ U)'(\xi)| \le \|U\|_{Lip} \mathbb{E}|g'(U(\xi))| = \|U\|_{Lip} \mathbb{E}|g'(\eta)| \le \frac{1}{K_{\mu}} \mathbb{E}|g'(\eta)|.$$

This implies  $K_{\mu} \leq h_{\mu}$ . The opposite inequality is obvious. Moreover, from  $\mathbb{E}|g(\eta)| \leq ||U||_{Lip} \mathbb{E}|g'(\eta)|$  we deduce  $1/||U||_{Lip} \leq h_{\mu}$ . Since  $K_{\mu} = h_{\mu}$  and  $||U||_{Lip} \leq 1/K_{\mu}$  the proposition follows.

#### **1.6** Generalized Poincare inequality

We prove the following lemma

**1.11 Lemma.** Let  $\Psi : \mathbb{R} \to \mathbb{R}$  be a Young function (even, convex,  $\Psi(0) = 0$ ). If  $h_{\mu} > 0$  then for all locally Lipschitz functions  $f : M \to \mathbb{R}$  with  $\operatorname{Med}_{\mu} f = 0$  we have

$$\|f\|_{\Psi} \le \frac{c_{\Psi}}{h_{\mu}} \|\nabla f\|_{\Psi}$$

and

$$\mathbb{E}\Psi(f) \le \mathbb{E}\Psi\left(\frac{c_{\Psi}}{h_{\mu}}|\nabla f|\right),$$

where  $c_{\Psi} = \sup_{x>0} \frac{x\Psi'(x)}{\Psi(x)}$ ,

$$\|f\|_{\Psi} = \inf\{\lambda > 0 : \mathbb{E}\Psi(f/\lambda) \le 1\} < \infty$$

and  $\|\nabla f\|_{\Psi} = \||\nabla f|\|_{\Psi}$ .

We need a simple lemma

**1.12 Lemma** (Hölder type inequality). Let f, g be measurable on  $(\Omega, \mathcal{F}, \mu)$  and let  $\Psi$ :  $\mathbb{R} \to \mathbb{R}$  be differentiable and convex. Assume that  $\mathbb{E}N(g) \leq \mathbb{E}N(f)$ . Then  $\mathbb{E}N'(f)g \leq \mathbb{E}N'(f)f$ , provided that this expectations exist.

Proof. We have

$$\mathbb{E}N((1-t)f + tg) \le (1-t)\mathbb{E}N(f) + t\mathbb{E}N(g).$$

Differentiating this inequality at t = 0 we get

$$\mathbb{E}N'(f)(g-f) \le \mathbb{E}N(g) - \mathbb{E}N(f) \le 0.$$

**1.13 Remark.** If we take  $N(x) = |x|^p$ ,  $f = u^{1/(p-1)}$ , g = v, q = p/(p-1) and assume that  $||u||_q = ||v||_p$  then we obtain  $\mathbb{E}uv \leq 1$ , hence we recover classical Hölder inequality.

Proof of lemma. We can assume that  $\Psi$  is differentiable,  $\Psi'(0) = 0$ . We can also assume that  $||f||_{\Psi} = 1$ , hence  $\mathbb{E}\Psi(f) = 1$  and that f is bounded. Let  $f_1 = f_+$  and  $f_2 = f_-$ , so that  $f = f_1 - f_2$ . We have  $\operatorname{Med}_{\mu}(f_1) = \operatorname{Med}_{\mu}(f_2) = 0$ , since  $\operatorname{Med}_{\mu}(f) = 0$  and therefore also  $\operatorname{Med}_{\mu}(N(f_1)) = \operatorname{Med}_{\mu}(N(f_2)) = 0$ . We have

$$h_{\mu}\mathbb{E}\Psi(f_{1}) \leq \mathbb{E}\Psi'(f_{1})|\nabla f_{1}| = \mathbb{E}\Psi(|f|)|\nabla f|\mathbf{1}_{\{f>0\}},$$
  
$$h_{\mu}\mathbb{E}\Psi(f_{2}) \leq \mathbb{E}\Psi'(f_{2})|\nabla f_{2}| = \mathbb{E}\Psi(|f|)|\nabla f|\mathbf{1}_{\{f<0\}}.$$

Therefore

$$h_{\mu}\mathbb{E}\Psi(f) \leq \mathbb{E}\Psi(|f|)|\nabla f|.$$

We want to apply lemma to the functions |f| and  $g = \frac{|\nabla f|}{\|\nabla f\|_{\Psi}}$ . Note that  $\mathbb{E}\Psi(f) = 1$ and  $\mathbb{E}\Psi(g) = 1$  by the definition of the norm  $\|\nabla f\|_{\Psi}$ . We arrive at

$$h_{\mu}\mathbb{E}\Psi(f) \leq \mathbb{E}\Psi(|f|)|\nabla f| = \|\nabla f\|_{\Psi} \mathbb{E}\Psi'(|f|)g$$
  
$$\leq \|\nabla f\|_{\Psi} \mathbb{E}\Psi'(|f|)|f| \leq c_{\Psi} \|\nabla f\|_{\Psi} \mathbb{E}\Psi(f).$$

We conclude that  $h_{\mu} \leq c_{\Psi} \|\nabla f\|_{\Psi}$  and since  $\|f\|_{\Psi} = 1$  we have

$$\|f\|_{\Psi} \le \frac{c_{\Psi}}{h_{\mu}} \|\nabla f\|_{\Psi}.$$

To prove the second part take  $\Psi_{\alpha} = \Psi/\alpha$  (note that  $c_{\Psi} = c_{\Psi_{\alpha}}$ ) and observe that from the first part if  $||f||_{\Psi_{\alpha}} \ge 1$  then  $||c_{\Psi}\nabla f/h_{\mu}||_{\Psi_{\alpha}} \ge 1$ . Therefore if  $\mathbb{E}\Psi(f) \ge \alpha$  then  $\mathbb{E}\Psi(c_{\Psi}|\nabla f|/h_{\mu}) \ge \alpha$ . The proposition follows.

**1.14 Corollary.** If we takie  $\Psi(x) = |x|$  then we recover the inequality

$$h\mathbb{E}|f - \operatorname{Med}_{\mu} f| \le \mathbb{E}|\nabla f|.$$

If we take  $\Psi(x) = x^2$  then we have  $c_{\Psi} = 2$  and we obtain the Poincaré inequality

$$\mathbb{E}|f - \mathbb{E}f|^2 \le \mathbb{E}|f - \operatorname{Med}_{\mu} f| \le \frac{4}{h_{\mu}^2} \mathbb{E}|\nabla f|^2.$$

It means that the optimal constant C in the Poincare inequality

$$C\mathbb{E}|f - \mathbb{E}f|^2 \le \mathbb{E}|\nabla f|^2$$

satisfies  $C \ge \frac{h_{\mu}^2}{4}$ .

### 1.7 Bobkov-type inequality

**1.15 Theorem.** For  $p \in [0,1]$  we take I(p) = 4p(1-p) and let  $C = 4\sqrt{6}/h_{\mu}$ . Then for all Lipschitz functions  $f: M^n \to [0,1]$  we have

$$I\left(\int f \,\mathrm{d}\mu^n\right) \leq \int \sqrt{I(f)^2 + C^2 |\nabla f|^2} \,\mathrm{d}\mu^n.$$

Note that we have

$$I(\int f) - \int I(f) = 4(\int f)(1 - \int f) - 4 \int f(1 - f) = \operatorname{Var}(f).$$

Therefore Bobkov inequality implies, since  $\sqrt{a^2 + b^2} \leq a + b$  for  $a, b \geq 0$ , that

$$4\operatorname{Var}_{\mu^n} f \le C \int |\nabla f| \,\mathrm{d}\mu^n.$$

Moreover, Bobkov inequality implies, by the standard approximation argument, that

$$(\mu^n)^+(A) \ge \frac{4}{C}\mu^n(A)(1-\mu^n(A)) \ge \frac{2}{C}\min\{\mu^n(A), 1-\mu^n(A)\},\$$

which means that  $h_{\mu^n} \ge h_{\mu}/2\sqrt{6}$ . That was our goal.

In the next section we prove that Bobkov inequality possesses the tensorization property. Therefore it suffices to prove it in dimension n = 1.

*Proof in the case* n = 1. Rewrite our inequality in the form

$$4\operatorname{Var}_{\mu}(f) \leq \int \left(\sqrt{I(f)^2 + C^2 |\nabla f|^2} - I(f)\right) \,\mathrm{d}\mu.$$

Note that  $0 \leq I(f) \leq 1$  and  $u \mapsto \sqrt{u^2 + v^2} - u$  is nonincreasing. Therefore it suffices to prove

$$4\operatorname{Var}_{\mu}(f) \leq \int \left(\sqrt{1+C^2|\nabla f|^2} - 1\right) \,\mathrm{d}\mu.$$

If  $m = Med_{\mu}(f)$  then f - m has values in [-1, 1]. Note that when  $|t| \leq 1$  then

$$4t^2 \le \frac{24t^2}{\sqrt{1+24t^2}+1} = \sqrt{1+24t^2}-1.$$

Take the Young function  $\Psi(t) = \sqrt{1+t^2} - 1$ . We have

$$4 \operatorname{Var}_{\mu}(f) \le 4 \int (f - m)^2 \, \mathrm{d}\mu \le \int \left(\sqrt{1 + 24(f - m)^2} - 1\right) \, \mathrm{d}\mu$$
$$= \int \Psi(\sqrt{24}(f - m)) \, \mathrm{d}\mu.$$

We compute

$$\frac{t\Psi'(t)}{\Psi(t)} = t \cdot \frac{t}{\sqrt{1+t^2}} \cdot \frac{1}{\sqrt{1+t^2}-1} = \frac{\sqrt{1+t^2}+1}{\sqrt{1+t^2}} = 1 + \frac{1}{\sqrt{1+t^2}}.$$

Therefore  $c_{\Psi} = 2$  and lemma implies

$$\int \Psi(\sqrt{24}(f-m)) \,\mathrm{d}\mu \le \int \Psi\left(\frac{2\sqrt{24}}{h_{\mu}}|\nabla f|\right) \,\mathrm{d}\mu = \int \left(\sqrt{1+C^2}|\nabla f|^2 - 1\right) \,\mathrm{d}\mu.$$

**1.16 Theorem.** Let  $\nu$  be the exponential distribution and let  $\Phi$  be its distribution function. If  $\nu^n(A) = \Phi(a)$  then

$$\nu^n(A^h) \ge \Phi\left(a + \frac{h}{2\sqrt{6}}\right).$$

*Proof.* Let a(h) be such that  $\nu^n(A^h) = \nu((-\infty, a(h)])$ . We would like to prove that  $a(h) \ge a + \frac{h}{2\sqrt{6}}$ . We have

$$\frac{\mathrm{d}}{\mathrm{d}h}\nu^{n}(A^{h})\Big|_{h=h_{0}} = a'(h_{0})\frac{\mathrm{d}}{\mathrm{d}a}\nu((-\infty,a])\Big|_{a=a(h_{0})} = a'(h_{0})\min\{\nu^{n}(A^{h_{0}}), 1-\nu^{n}(A^{h_{0}})\}$$

Moreover, using  $h_{\nu^n} \ge 1/2\sqrt{6}$  we obtain

$$\frac{\mathrm{d}}{\mathrm{d}h}\nu^{n}(A^{h})\Big|_{h=h_{0}} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\nu^{n}(A^{h_{0}+\varepsilon})\Big|_{\varepsilon=0} \ge \frac{1}{2\sqrt{6}}\min\{\nu^{n}(A^{h_{0}}), 1-\nu^{n}(A^{h_{0}})\}.$$

It means that  $a'(h_0) \ge 1/2\sqrt{6}$  and therefore  $a(h) \ge a + h/2\sqrt{6}$ .

# 1.8 Tensorization property of Bobkov type inequalities

**1.17 Lemma.** Suppose for all Lipschitz functions  $f: M \to [0, 1]$  we have

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

Then is  $n \ge 1$  then for all Lipschitz functions  $f: M^n \to [0,1]$  we have

$$I\left(\int f \,\mathrm{d}\mu^n\right) \leq \int \sqrt{I(f)^2 + |\nabla f|^2} \,\mathrm{d}\mu^n.$$

*Proof.* Induction. Let  $f: M^{n+1} \to [0,1]$ . For  $x \in M^n$  and  $x_{n+1} \in M$  take

$$u(x, x_{n+1}) = \sqrt{I(f(x, x_{n+1}))^2 + |\nabla_x f|^2}, \qquad v(x, x_{n+1}) = |\nabla_{x_{n+1}} f|,$$
$$a(x_{n+1}) = \int f(x, x_{n+1}) \, \mathrm{d}\mu^n(x).$$

We have triangle inequality

$$\int \sqrt{u(x, x_{n+1})^2 + v(x, x_{n+1}^2)} \, d\mu^n(x) \ge \sqrt{\left(\int u(x, x_{n+1}) \, d\mu^n(x)\right)^2 + \left(\int v(x, x_{n+1}) \, d\mu^n(x)\right)^2}.$$

Therefore, by the induction hypothesis we have

$$\int \sqrt{I(f)^{2} + |\nabla f|^{2}} \, \mathrm{d}\mu^{n+1} = \int \int \sqrt{I(f(x, x_{n+1}))^{2} + |\nabla_{x} f|^{2} + |\nabla_{x_{n+1}} f|^{2}} \, \mathrm{d}\mu^{n} \, \mathrm{d}\mu(x_{n+1})$$

$$\geq \int \sqrt{\left(\int u(x, x_{n+1}) \, \mathrm{d}\mu^{n}(x)\right)^{2} + \left(\int v(x, x_{n+1}) \, \mathrm{d}\mu^{n}(x)\right)^{2}} \, \mathrm{d}\mu(x_{n+1})$$

$$\geq \int \sqrt{\left(I(a(x_{n+1}))\right)^{2} + \left(\int v(x, x_{n+1}) \, \mathrm{d}\mu^{n}(x)\right)^{2}} \, \mathrm{d}\mu(x_{n+1})$$
We also have

We also have

$$\int v(x, x_{n+1}) \,\mathrm{d}\mu^n(x) \ge |a(x_{n+1})|$$

and (using again induction hypothesis for n = 1) we arrive at

$$\int \sqrt{I(f)^{2} + |\nabla f|^{2}} \, \mathrm{d}\mu^{n+1} \ge \int \sqrt{\left(I(a(x_{n+1}))\right)^{2} + |\nabla_{x_{n+1}}a(x_{n+1})|^{2}} \, \mathrm{d}\mu(x_{n+1})$$
$$\ge I\left(\int a(x_{n+1}) \, \mathrm{d}\mu(x_{n+1})\right) = I\left(\int f \, \mathrm{d}\mu^{n+1}\right).$$

#### Two point Bobkov inequality 1.9

Let  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ ,  $\Phi(x) = \int_{-\infty}^x \varphi(s) \, ds$  and  $I(x) = \varphi(\Phi^{-1}(x))$ . In this section we prove the following two-point Bobkov inequality

**1.18 Theorem.** For all  $0 \le a, b \le 1$  we have

$$I\left(\frac{a+b}{2}\right) \le \frac{1}{2}\sqrt{I(a)^2 + \left|\frac{a-b}{2}\right|^2} + \frac{1}{2}\sqrt{I(b)^2 + \left|\frac{a-b}{2}\right|^2}.$$

We will frequently use some simple identities involving functions  $\varphi$ ,  $\Phi$  and I. We have  $\varphi(x) = \varphi(-x)$  and  $\Phi(x) + \Phi(-x) = 1$ . Moreover, we have I(1-p) = I(p). Indeed, if we take  $x = \Phi^{-1}(p)$  for some  $p \in [0, 1]$  then  $\Phi(x) + \Phi(-x) = 1$  implies  $\Phi(-\Phi^{-1}(p)) = 1 - p$  and therefore

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

Let  $\mu = \frac{1}{2}\delta_{\{1\}} + \frac{1}{2}\delta_{\{-1\}}$ . Take  $f : \{-1, 1\} \to [0, 1]$ . Let us set f(-1) = a, f(1) = b. Then two-point Bobkov inequality is equivalent to

$$I\left(\int f \,\mathrm{d}\mu\right) \leq \int I(f)^2 + |\nabla f|^2 \,\mathrm{d}\mu$$

where  $(\nabla f)(x) = \frac{f(x)-f(-x)}{2}$ . Note that  $|(\nabla f)(1)| = |(\nabla f)(-1)| = \left|\frac{a-b}{2}\right|$ . Therefore, if we prove two-point Bobkov inequality then by tensorization properties of Bobkov-type inequalities we obtain the following theorem

**1.19 Theorem.** Let  $f : \{-1,1\}^n \to [0,1]$  and let  $\mu = \frac{1}{2}\delta_{\{1\}} + \frac{1}{2}\delta_{\{-1\}}$ . Take  $\mu^n = \mu^{\otimes n}$ . Then we have

$$I\left(\int f \,\mathrm{d}\mu^n\right) \leq \int I(f)^2 + |\nabla f|^2 \,\mathrm{d}\mu^n,$$

where

$$|\nabla f|^2 = \sum_{i=1}^n |\nabla_{x_i} f|^2 = \sum_{i=1}^n \frac{1}{4} \left( f(x) - f(\sigma_i(x)) \right)^2.$$

Here  $\sigma_i((x_1,\ldots,x_i,\ldots,x_n)) = (x_1,\ldots,-x_i,\ldots,x_n).$ 

Now, let  $\gamma_1$  be the one dimensional standard Gaussian measure, namely, measure with density  $\phi$ . Let  $f : \mathbb{R} \to [0, 1]$  be a bounded function with bounded first and second derivatives. Define  $f_n : \{-1, 1\}^n \to [0, 1]$  by

$$f_n(x_1,\ldots,x_n) = f\left(\frac{x_1+\ldots+x_n}{\sqrt{n}}\right).$$

Note that by Central Limit Theorem we have

$$\lim_{n \to \infty} \int f_n \, \mathrm{d}\mu^n = \int f \, \mathrm{d}\gamma_1$$

Moreover,

$$\begin{aligned} |\nabla f_n|^2(x) &= \frac{1}{4} \sum_{i=1}^n \left( f\left(\frac{x_1 + \dots + x_n}{\sqrt{n}}\right) - f\left(\frac{x_1 + \dots + x_n}{\sqrt{n}} - \frac{2x_i}{\sqrt{n}}\right) \right)^2 \\ &= \frac{1}{4} \sum_{i=1}^n \left| f'\left(\frac{x_1 + \dots + x_n}{\sqrt{n}}\right) \right|^2 \frac{4x_i^2}{n} + O(1/n) \\ &= \left| f'\left(\frac{x_1 + \dots + x_n}{\sqrt{n}}\right) \right|^2 + O(1/n). \end{aligned}$$

Therefore also

$$\lim_{n \to \infty} \int \sqrt{I(f_n)^2 + |\nabla f_n|^2} \, \mathrm{d}\mu^n = \int \sqrt{I(f)^2 + |f'|^2} \, \mathrm{d}\gamma_1.$$

Hence we obtain

$$I\left(\int f \,\mathrm{d}\gamma_1\right) \leq \int \sqrt{I(f)^2 + |f'|^2} \,\mathrm{d}\gamma_1.$$

for all locally Lipschitz functions (we have to use appropriate approximation). Using again tensorization we get

**1.20 Theorem.** Let  $f : \mathbb{R}^n \to [0, 1]$  be locally Lipschitz and let  $\gamma_n = \gamma_1^{\otimes n}$  be the standard Gaussian measure on  $\mathbb{R}^n$ . Then we have

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

Now let A be a Borel set. Using standard approximation of  $\mathbf{1}_A$  and a fact that I(0) = I(1) = 0 we obtain the celebrated Gaussian isoperimetry in the infinitesimal form.

**1.21 Theorem.** Let A be a Borel set with  $\gamma_n(A) < \infty$ . Then

$$\gamma_n^+(A) \ge I(\gamma_n(A))$$

It means that if  $\gamma_n(A) = \gamma_n((-\infty, h] \times \mathbb{R}^{n-1}) = \Phi(h)$  then

$$\gamma_n^+(A) \ge \gamma_n^+((-\infty,h] \times \mathbb{R}^{n-1}) = \varphi(h) = \varphi(\Phi^{-1}(\gamma_n(A))) = I(\gamma_n(A)).$$

Therefore half spaces  $(-\infty, h] \times \mathbb{R}^{n-1}$  are the extremal sets in the Gaussian isoperimetry.

In fact half spaces are optimal in the strong version of Gaussian isoperimetry.

**1.22 Theorem.** Let A be a Borel set in  $\mathbb{R}^n$  with  $\gamma_n(A) < \infty$  and let  $B_h = (-\infty, h] \times \mathbb{R}^{n-1}$  be such that  $\gamma_n(A) = \gamma_n(B_h) = \Phi(h)$ . Then we have

$$\gamma_n(A^u) \ge \gamma_n(B_h^u) = \Phi(h+u) = \Phi\left(\Phi^{-1}(\gamma_n(A)) + u\right).$$

*Proof.* Let  $f:[0,\infty) \to \mathbb{R}$  be such that

$$\gamma_n(A^u) = \Phi\left(\Phi^{-1}(\gamma_n(A)) + f(u)\right).$$

Clearly f(0) = 0. We would like to prove that  $f(u) \ge u$ . We have

$$\frac{\mathrm{d}}{\mathrm{d}u}\gamma_n(A^u)\Big|_{u=u_0} = \varphi\left(\Phi^{-1}(\gamma_n(A)) + f(u_0)\right)f'(u_0) = I(\gamma_n(A^{u_0}))f'(u_0).$$

Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}u}\gamma_n(A^u)\Big|_{u=u_0} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\gamma_n(A^{u_0+\varepsilon})\Big|_{\varepsilon=0} = \gamma_n^+(A^{u_0}) \ge I(\gamma_n(A^{u_0})).$$

It follows that  $f'(u_0) \ge 1$  and therefore  $f(u) \ge u + f(0) = u$ .

**1.23 Remark.** Suppose  $J: [0,1] \to \mathbb{R}$  satisfies two-point Bobkov inequality

$$J\left(\frac{a+b}{2}\right) \le \frac{1}{2}\sqrt{J(a)^2 + \left|\frac{a-b}{2}\right|^2} + \frac{1}{2}\sqrt{J(b)^2 + \left|\frac{a-b}{2}\right|^2}, \quad a, b \in [0,1].$$

Then we can prove (using above arguments) that  $\gamma_n^+(A) \ge J(\gamma_n(A))$  for all Borel  $A \subset \mathbb{R}^n$ . Let A be a half space of Gaussian measure p. Then  $I(p) = \gamma_n^+(A) \ge J(\gamma_n(A)) = J(p)$ . Hence  $I \ge J$  and therefore I is maximal among all functions satisfying two-point Bobkov inequality.

**1.24 Remark.** For  $a, b \ge 0$  we have  $\sqrt{a^2 + b^2} \le a + b$ . Hence, we have the weaker form of Gaussian Bobkov inequality

$$I\left(\int f \, \mathrm{d}\gamma_n\right) - \int I(f) \, \mathrm{d}\gamma_n \leq \int |\nabla f| \, \mathrm{d}\gamma_n.$$

This inequality also implies Gaussian isoperimetry.

**1.25 Remark.** We can try to deduce Gaussian Bobkov inequality in dimension n from the inequality  $\gamma_{n+1}^+(A) \geq I(\gamma_{n+1}(A))$  in dimension n + 1. However, it is not clear if argument that we will give is rigorous enough to deduce Bobkov inequality for all smooth functions.

For a function  $f : \mathbb{R}^n \to \mathbb{R}$  take

$$A = \{ (x, y) : x \in \mathbb{R}^n, y \in \mathbb{R}, \ \Phi(y) < f(x) \}.$$

Let  $g = \Phi^{-1} \circ f$ . We have

$$\gamma_{n+1}(A) = \int_{\mathbb{R}^n} \int_{-\infty}^{g(x)} \mathrm{d}\gamma_1(y) \,\mathrm{d}\gamma_n(x) = \int_{\mathbb{R}^n} \Phi(g(x)) \,\mathrm{d}\gamma_n(x) = \int_{\mathbb{R}^n} f \,\mathrm{d}\gamma_n.$$

Moreover,

$$\int \sqrt{I(f)^2 + |\nabla f|^2} \, \mathrm{d}\gamma_n = \int \sqrt{I(\Phi(g))^2 + |\nabla \Phi(g)|^2} \, \mathrm{d}\gamma_n = \int \sqrt{\varphi(g)^2 + \Phi'(g)|\nabla g|^2} \, \mathrm{d}\gamma_n$$
$$= \int \varphi(g)\sqrt{1 + |\nabla g|^2} \, \mathrm{d}\gamma_n = \int_{\mathbb{R}^n} \varphi_n(x)\varphi(g(x))\sqrt{1 + |\nabla g|^2} \, \mathrm{d}x$$
$$= \gamma_{n+1}^+(A).$$

Hence  $\gamma_{n+1}^+(A) \ge I(\gamma_{n+1}(A))$  implies

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

It means that the inequality  $\gamma_n^+(A) \ge I(\gamma_n(A))$  is essentially two dimensional. If we can prove it for n = 2 the we can deduce one dimensional Bobkov inequality and then tensorizing it, obtain general Gaussian Bobkov inequality. Then  $\gamma_n^+(A) \ge I(\gamma_n(A))$  follows for all  $n \ge 1$ .

Proof of two-point Bobkov inequality. Let  $x = \frac{a-b}{2}$ ,  $c = \frac{a+b}{2}$ . Hence, a = x+c, b = c-x. Now  $a, b \in [0, 1]$  means that  $0 \le x+c \le 1$  and  $0 \le c-x \le 1$ . It is equivalent to  $-c \le x \le 1-c$ ,  $-1+c \le x \le c$  and therefore equivalent to  $x \in \Delta_c = (-\min\{c, 1-c\}, \min\{c, 1-c\})$ . Take  $g(x) = I(x+c)^2 + x^2$ . We can rewrite our inequality in the following form,

$$\sqrt{g(0)} \le \frac{1}{2}\sqrt{g(x)} + \frac{1}{2}\sqrt{g(-x)}, \quad x \in \Delta_c.$$

Squaring we obtain

$$4g(0) - (g(x) + g(-x)) \le 2\sqrt{g(x)g(-x)}$$

Squaring this inequality again we can see that it is enough to prove

$$16g(0)^{2} + (g(x) - g(-x))^{2} \le 8g(0)(g(x) + g(-x))$$

Let h(x) = g(x) - g(0). Rewriting the above inequality i terms of h we obtain

$$(h(x) - h(-x))^2 \le 8g(0)^2(h(x) + h(-x)) = 8I(c)^2(h(x) + h(-x)).$$

We need a lemma

#### 1.26 Lemma. We have

- a)  $I \cdot I'' = -1$ ,
- b)  $(I')^2$  is convex on (0, 1),
- c)  $R(x) = h(x) + h(-x) 2I'(c)^2 x^2$  is convex on  $\Delta_c$ .

*Proof.* a) We have  $I = \varphi \circ \Phi^{-1}$ . Therefore

$$I' = (\varphi' \circ \Phi^{-1}) \frac{1}{\Phi' \circ \Phi^{-1}} = -\Phi^{-1} (\varphi \circ \Phi^{-1}) \frac{1}{\Phi' \circ \Phi^{-1}} = -\Phi^{-1}.$$

Hence  $I'' = -\frac{1}{\Phi' \circ \Phi^{-1}} = -\frac{1}{I}$ . b) We have  $(I'^2)' = 2I' \cdot I'' = -2I'/I$ , hence

$$(I'^2)'' = -2\frac{I''I - I'^2}{I^2} = \frac{2(1+I'^2)}{I^2} \ge 0.$$

c) Recall that  $h(x) = I(x+c)^2 + x^2 - I(c)^2$ , hence

$$R(x) = I(x+c)^{2} + I(-x+c)^{2} + 2x^{2} - 2I'(c)^{2}x^{2} - 2I(c)^{2},$$

$$R'(x) = 2I(x+c)I'(x+c) - 2I(-x+c)I'(-x+c) + 4x - 4I'(c)^{2}x$$

and

$$R''(x) = 2I'(x+c)^2 + 2I'(-x+c)^2 + 2I(x+c)I''(x+c) + 2I(-x+c)I''(-x+c) - 4I'(c)^2 = 4\left(\frac{I'(x+c)^2 + I'(-x+c)^2}{2} - I'(c)^2\right) \ge 0,$$

since  $(I')^2$  is convex.

We continue the proof. The function R is even and convex. Therefore  $R(x) \ge R(0) = 2h(0) = 0$ . We get

$$h(x) + h(-x) \ge 2I'(c)^2 x^2.$$

Clearly,

$$(h(x) - h(-x))^2 \le 8I(c)^2(h(x) + h(-x))$$

will follow from the inequality

$$(h(x) + h(-x))^{2} \le 16I(c)^{2}I'(c)^{2}x^{2},$$

which is equivalent to

$$\left|\frac{h(x) - h(-x)}{x}\right| \le 4I(c)|I'(c)|$$

and therefore, to

$$\left|\frac{I(x+c)^2 - I(-x+c)^2}{x}\right| \le 4I(c)|I'(c)|.$$

Note that I(c) = I(1-c), I'(c) = -I'(1-c). It follows that

$$\left|I(x+(1-c))^{2}-I(-x+(1-c))^{2}\right| = \left|I(-x+c)^{2}-I(x+c)^{2}\right|.$$

We also have  $\Delta_c = \Delta_{1-c}$ . We conclude that the above inequality does not change when we replace c by 1-c. We can therefore assume that  $c \in [0, 1/2]$ . Then  $\Delta_c = (-c, c)$ . Observe that the left hand side is an even function of  $x \in \Delta_c$ . We can therefore assume that  $0 < x \le c \le 1/2$ . Since I is increasing on [0, 1/2] and decreasing on [1/2, 1] and symmetric around 1/2 we have  $I(c + x) \ge I(c - x)$ . Therefore we have to prove

$$\frac{I(x+c)^2 - I(-x+c)^2}{x} \le 4I(c)|I'(c)|.$$

Consider a function  $u(x) = I(x+c)^2 - I(-x+c)^2$ . WE have

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

and

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

The function  $I'^2$  is convex and symmetric around 1/2. It follows that  $I'(c+x)^2 \leq I'(c-x)^2$ . Therefore  $u''(x) \leq 0$ . We have u(0) = 0 and u is concave on [0,c]. Hence u(x)/x is nonincreasing on [0,c]. To prove the inequality  $u(x)/x \leq 4I(c)|I'(c)|$  it suffices to prove it in the limit  $x \to 0^+$ . We have

$$I(x+c)^{2} = I(c)^{2} + 2I(c)I'(c)x + o(x)$$

and therefore

$$u(x) = 4I(c)I'(c)x + o(x).$$

It follows that

$$\lim_{x \to 0^+} \frac{u(x)}{x} = 4I(c)I'(c).$$

# 1.10 Log-Sobolev inequalities

Let  $\mu$  be a Borel probability measure on a metric space (X, d). For a positive function g on X define the *entropy* 

$$\operatorname{Ent}_{\mu}(f) = \int_{X} f \ln f \, \mathrm{d}\mu - \left(\int f \, \mathrm{d}\mu\right) \ln \left(\int f \, \mathrm{d}\mu\right).$$

We say that measure  $\mu$  satisfies log-Sobolev inequality with constant C if

$$\operatorname{Ent}_{\mu}(f^2) \le C \int_X |\nabla f|^2 \, \mathrm{d}\mu.$$

To prove tensorization property for the log-Sobolev inequality we need a simple lemma

#### 1.27 Lemma. We have

Ent 
$$\mu(f) = \sup \left\{ \int fg \, d\mu : \int e^g \, d\mu \leq 1 \right\}.$$

*Proof.* One can easily check that for  $\lambda > 0$  we have  $\operatorname{Ent}_{\mu}(\lambda f) = \lambda \operatorname{Ent}_{\mu}(f)$ . Therefore, by the homogenity we can assume that  $\int f \, d\mu = 1$ . Then  $\operatorname{Ent}_{\mu}(f) = \int f \ln f \, d\mu$ . Note that for u > 0 and  $v \in \mathbb{R}$  we have

$$uv \le u \ln u - u + e^v.$$

Indeed, the function  $v \mapsto e^v - uv$  attains its minimum in the point  $v = \ln u$  (compute the derivative). If  $\int e^g d\mu \leq 1$  we get

$$\int fg \, \mathrm{d}\mu \leq \int f \ln f - f + e^g \, \mathrm{d}\mu \leq \int f \ln f - f + 1 \, \mathrm{d}\mu = \int f \ln f \, \mathrm{d}\mu = \mathrm{Ent}_{\mu}(f).$$

Hence

Ent 
$$\mu(f) = \sup \left\{ \int fg \, \mathrm{d}\mu : \int e^g \, \mathrm{d}\mu \leq 1 \right\}.$$

To obtain the converse it suffices to take  $g = \ln f$ .

Now we can prove the subadditivity of the entropy.

**1.28 Lemma.** Let  $\mu_1, \ldots, \mu_n$  be probability measures on  $X_1, \ldots, X_n$ . Take the measure  $\mu = \mu_1 \otimes \ldots \otimes \mu_n$  on  $X = X_1 \times \ldots \times X_n$ . For  $f : X \to (0, \infty)$  we have

$$\operatorname{Ent}_{\mu}(f) \leq \sum_{i=1}^{n} \int \operatorname{Ent}_{\mu_{i}}(f) \, \mathrm{d}\mu_{i}$$

Here  $\operatorname{Ent}_{\mu_i}(f)$  is the entropy of the function  $X_i \ni x_i \mapsto f(x_1, \ldots, x_i, \ldots, x_n)$  where variables other than  $x_i$  are fixed.

*Proof.* Let  $g: X \to \mathbb{R}$  be such that  $\int_X g \, d\mu \leq 1$ . Take

$$g^{i}(x_{1},...,x_{n}) = \ln\left(\frac{\int e^{g(x_{1},...,x_{n})d_{\mu_{1}(x_{1})}...d_{\mu_{i-1}(x_{i-1})}}}{\int e^{g(x_{1},...,x_{n})d_{\mu_{1}(x_{1})}...d_{\mu_{i}(x_{i})}}}\right)$$

We have

$$\sum_{i=1}^{n} g^{i} = \ln(e^{g}) - \ln\left(\int e^{g} \,\mathrm{d}\mu\right) \ge g.$$

Note that

$$\int e^{g^i} d\mu_i = \int \frac{\int e^g d_{\mu_1} \dots d_{\mu_{i-1}}}{\int e^g dd_{\mu_1} \dots d_{\mu_i}} d\mu_i = 1.$$

Hence,

$$\int fg \, \mathrm{d}\mu \leq \sum_{i=1}^n \int fg^i \, \mathrm{d}\mu = \sum_{i=1}^n \int \int fg^i \, \mathrm{d}\mu_i \, \mathrm{d}\mu \leq \sum_{i=1}^n \int \mathrm{Ent}_{\mu_i}(f) \, \mathrm{d}\mu.$$

We finish the proof taking supremum over all functions g with  $\int e^g d\mu \leq 1$ .

**1.29 Theorem.** Take  $(X_i, d_i, \mu_i)_{i=1,\dots,n}$ ,  $X = X_1 \times \dots \times X_n$ ,  $\mu = \mu_1 \otimes \dots \otimes \mu_n$  and assume that for  $f : X \to \mathbb{R}$  we have  $|\nabla f|^2 = \sum_{i=1}^n |\nabla_i f|^2$ . Suppose  $\mu_i$  satisfies log-Sobolev inequality with constant  $C_i$ . Then the measure  $\mu$  on X satisfies log-Sobolev inequality with constant  $C = \max_{1 \le i \le n} C_i$ .

*Proof.* We have

$$\operatorname{Ent}_{\mu}(f^{2}) \leq \sum_{i=1}^{n} \int \operatorname{Ent}_{\mu_{i}}(f^{2}) \, \mathrm{d}\mu \leq \sum_{i=1}^{n} C_{i} \int \int |\nabla_{i}f|^{2} \, \mathrm{d}\mu_{i} \, \mathrm{d}\mu \leq C \int |\nabla f|^{2} \, \mathrm{d}\mu.$$

Now we prove log-Sobolev inequality for the discrete cube  $\{-1, 1\}^n$ .

**1.30 Theorem.** Let  $f : \{-1,1\}^n \to (0,\infty)$ . Let  $|\nabla f|^2 = \frac{1}{4} \sum_{i=1}^n (f(x) - f(\sigma_i(x)))^2$ . Take  $\mu = \left(\frac{1}{2}\delta_{\{-1\}} + \frac{1}{2}\delta_{\{1\}}\right)^{\otimes n}$ . Then

$$\operatorname{Ent}_{\mu}(f^2) \le 2 \int |\nabla f|^2 \, \mathrm{d}\mu$$

*Proof.* Because of the tensorization property of log-Sobolev inequality it suffices to prove the theorem in the case n = 1. By homogenity we can assume that  $\int f^2 d\mu = (f(1)^2 + f(-1)^2)/2 = 1$ . Clearly, there exists  $t \in [-1, 1]$  such that  $f(1)^2 = 1 + t$ ,  $f(-1)^2 = 1 - t$ . We have  $||f(1)| - |f(-1)|| \le |f(1) - f(-1)|$ , therefore we can assume that  $f \ge 0$ . Hence

$$|\nabla f|^2 = \frac{1}{4} \left( \sqrt{1+t} - \sqrt{1-t} \right)^2 = \frac{1}{2} - \frac{1}{2} \sqrt{1-t^2}.$$

We also have

$$\operatorname{Ent}_{\mu}(f^2) = \frac{1+t}{2}\ln(1+t) + \frac{1-t}{2}\ln(1-t).$$

We would like to prove

$$1 - \sqrt{1 - t^2} \ge \frac{1 + t}{2} \ln(1 + t) + \frac{1 - t}{2} \ln(1 - t).$$

Define

$$\alpha(t) = 1 - \sqrt{1 - t^2} - \frac{1 + t}{2} \ln(1 + t) - \frac{1 - t}{2} \ln(1 - t).$$

The function  $\alpha$  is even, therefore it suffices to prove  $\alpha(t) \ge 0$  for  $t \ge 0$ . Note that f(0) = 0. It suffices to prove that

$$\alpha'(t) = \frac{t}{\sqrt{1-t^2}} - \frac{1}{2}\ln(1+t) + \frac{1}{2}\ln(1-t) \ge 0.$$

Again f'(0) = 0 and it suffices to observe that

$$\alpha''(t) = \frac{\sqrt{1-t^2} + \frac{t^2}{\sqrt{1-t^2}}}{1-t^2} - \frac{1}{2}\frac{1}{1+t} - \frac{1}{2}\frac{1}{1-t}$$
$$= \frac{1}{1-t^2}\left(\frac{t^2}{\sqrt{1-t^2}} - \sqrt{1-t^2} - 1\right) = \frac{1}{1-t^2}\left(\frac{t^2}{\sqrt{1-t^2}} - \frac{t^2}{1+\sqrt{1-t^2}}\right) \ge 0.$$

Using Central Limit Theorem we can deduce, like in the case of Gaussian Bobkov inequality, the log-Sobolev inequality for the Gaussian measure.

**1.31 Theorem.** Let  $\gamma_n$  be a standard Gaussian measure on  $\mathbb{R}^n$ . Then for Lipschitz function  $f : \mathbb{R}^n \to (0, \infty)$  we have

$$\operatorname{Ent}_{\gamma_n}(f^2) \le 2 \int |\nabla f|^2 \, \mathrm{d}\gamma_n.$$

# 2 Semigroups approach

#### 2.1 Discrete cube

Consider a triple  $(\Sigma_n, P(\Sigma_n), \mu^n)$ , where  $\mu^n = \left(\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}\right)^{\otimes n}$  and  $P(\Sigma_n)$  is a  $\sigma$ -algebra of all subsets of  $\Sigma_n$ . The set  $\Sigma_n$  is called a *discrete cube*. Clearly our triple is a probability space, therefore for a function  $f: \Sigma_n \to \mathbb{R}$  we have expectation

$$\mathbb{E}f = \int_{\Sigma_n} f \, \mathrm{d}\mu^n = 2^{-n} \sum_{x \in \Sigma_n} f(x).$$

We also set  $||f||_p = (\mathbb{E}|f|^p)^{1/p}$  and  $||f||_{\infty} = \sup_{x \in \Sigma_n} |f(x)|.$ 

We can equip  $\Sigma_n$  with the Hamming metric

$$d(x,y) = |\{i \in [n] : x_i \neq y_i\}| = \frac{1}{2} ||x - y||_1$$

where  $||x||_1 = \sum_{i=1}^n |x_i|$ . Let  $[n] = \{1, 2, \dots, n\}$ .

We can also consider a Hilbert space  $\mathcal{H}_n(\Sigma_n) = \mathcal{H}_n(\Sigma_n, P(\Sigma_n), \mu^n)$  of all functions  $f: \Sigma_n \to \mathbb{R}$  with a scalar product

$$\langle f,g \rangle = \mathbb{E} fg = 2^{-n} \sum_{i=1}^{n} f(x)g(x)$$

For  $x \in \Sigma_n$  take a function  $\mathbf{1}_x : \Sigma_n \to \mathbb{R}$  given by  $\mathbf{1}_x(y) = \delta_{x,y}$  (Kronecker delta). Clearly, the collection  $(\mathbf{1}_x)_{x \in \Sigma_n}$  forms a basis in  $\mathcal{H}_n(\Sigma_n)$  and dim  $\mathcal{H}_n = 2^n$ .

The discrete cube possesses also a group structure  $(\Sigma_n, V)$ , where  $(x, y) \in V$  if and only if d(x, y) = 1. Note also that  $\Sigma_n$  is a locally compact group with multiplication given by

$$x \cdot y = (x_1 y_1, \dots, x_n y_n).$$

The measure  $\mu^n$  is a Haar measure on a group  $\Sigma_n$ .

### 2.2 The Gaussian space

Take a probability space  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \gamma_n)$ , where  $\mathcal{B}(\mathbb{R}^n)$  is a  $\sigma$ -algebra of all Borel sets in  $\mathbb{R}^n$ . Here  $\gamma_n$  is a standard Gaussian measure  $\mathbb{R}^n$ , which is the product of standard normal distributions, i.e. measures with density  $\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$ . For  $f : \mathbb{R}^n \to \mathbb{R}$  we also have an expectation

$$\mathbb{E}f = \int_{\mathbb{R}^n} f \, \mathrm{d}\gamma_n = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-\|x\|_2/2} \, \mathrm{d}x,$$

where  $||x||_2^2 = \sum_{i=1}^n x_i^2$ . Of course we assume that this expression is well-defined.

Clearly,  $\mathbb{R}^n$  has a standard Euclidean metric structure. It possesses also a group structure ( $\mathbb{R}^n$ , +), but it is the Lebesgue measure that is the Haar measure on this group, not the measure  $\gamma_n$ .

We introduce the Hilbert structure  $L^2(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \gamma_n))$  with a scalar product

$$\langle f,g\rangle = \mathbb{E}fg = \int_{\mathbb{R}^n} f(x)g(x) \,\mathrm{d}\gamma_n(x).$$

#### 2.3 Walsh functions

We introduce a set of functions  $(w_S)_{S \subset [n]}, w_S : \Sigma_n \to \{-1, 1\}$ , given by

$$w_S(x) = \prod_{i \in S} x_i, \qquad w_{\emptyset} \equiv 1.$$

The functions  $(r_i)_{i=1}^n$ ,  $r_i = w_{\{i\}}$  are called the Rademacher functions. Since  $r_i$  are defined on the probability space, they form a sequence of random variables. One can check that these random variables are independent.

Clearly,  $\mathbb{E}w_S = 0$  when  $S \neq \emptyset$  and  $\mathbb{E}w_{\emptyset} = 1$ . Moreover, for  $S, T \subset [n]$  we have

$$(w_S \cdot w_T)(x) = \prod_{i \in S} x_i \prod_{i \in T} w_T = \prod_{i \in S \Delta T} x_i \prod_{i \in S \cap T} x_i = \prod_{i \in S \Delta T} x_i = w_{S\Delta T}(x),$$

hence  $w_S \cdot w_T = w_{S\Delta T}$ . Therefore

$$\langle w_S, w_T \rangle = \mathbb{E} w_S w_T = \mathbb{E} w_{S\Delta T} = \delta_{S,T}.$$

We have proved that  $(w_S)_{S \subset [n]}$  is an orthonormal sequence of functions in  $L^2(\Sigma_n)$ . Since it is a subset of cardinality  $2^n$ , it forms a basis in  $L^2(\Sigma_n)$ . It follows that a function  $f: \Sigma_n \to \mathbb{R}$  admits an unique expansion

$$f = \sum_{S \subset [n]} \langle f, w_S \rangle w_S.$$

It can be also seen by an elementary argument. Indeed, we have

$$\mathbf{1}_{x}(y) = \prod_{i=1}^{n} \frac{1 + x_{i}y_{i}}{2} = 2^{-n} \sum_{S \subset [n]} w_{S}(x)w_{S}(y).$$

Hence,

$$f(x) = \sum_{y \in \Sigma_n} f(y) \mathbf{1}_y(x) = 2^{-n} \sum_{S \subset [n]} \left( \sum_{y \in \Sigma_n} f(y) w_S(y) \right) w_S(x) = \sum_{S \subset [n]} \langle f, w_S \rangle w_S(x)$$

Clearly,  $w_S : \Sigma_n \to \{-1, 1\}$  is a homomorphism and therefore it is a character on  $\Sigma_n$ . Let  $\hat{\Sigma}_n$  be the dual group of  $\Sigma_n$ , namely the group of all characters on  $\sigma_n$ . One can prove that  $\hat{\Sigma}_n = (w_S)_{S \subset [n]}$  and  $\sigma_n$  is isomorphic to  $\hat{\Sigma}_n$  by the isomorphism

$$x \in \Sigma_n \mapsto S = \{i \in [n] : x_i = -1\}.$$

We also have the Fourier transform

$$\hat{f}(S) = \hat{f}(w_S) = \int_{\Sigma_n} w_S(x) f(x) \, \mathrm{d}\mu^n(x) = \langle f, w_S \rangle.$$

Note also that we have

$$\mathbb{E}f = \mathbb{E}f \cdot \mathbf{1} = \mathbb{E}f \cdot w_{\emptyset} = \langle f, w_{\emptyset} \rangle = \hat{f}(\emptyset)$$

and

$$\mathbb{E}f^2 = \langle f, f \rangle = \left\langle \sum_{S \subset [n]} \hat{f}(S) w_S, \sum_{T \subset [n]} \hat{f}(T) w_T \right\rangle = \sum_{S, T \subset [n]} \hat{f}(S) \hat{f}(T) \langle w_S, w_T \rangle = \sum_{S \subset [n]} \hat{f}(S)^2.$$

### 2.4 Hermite polynomials in the Gaussian space

The sequence of polynomials  $(H_m)_{m\geq 0}$  obtained by the Gram-Schmidt process involving the monomials  $(x^m)_{m\geq 0}$  and the scalar product of the space  $L_2(\mathbb{R}, \gamma_1)$  is called the sequence of *Hermite polynomials*. That is,

$$H_0 = 1,$$
  

$$H_m = x^m - \sum_{i=0}^{m-1} \langle x^m, H_i \rangle \frac{H_i}{\|H_i\|^2}.$$
(2.1)

There are other equivalent definitions. For instance,

$$H_m(x) = (-1)^n e^{x^2/2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} (e^{-x^2/2}), \qquad (2.2)$$

or via the generating function

$$e^{xt-t^2/2} = \sum_{m\ge 0} \frac{H_m(x)}{m!} t^m.$$
(2.3)

This one is particularly useful for deriving the coefficients of  $H_m$ . Indeed,

$$e^{xt-t^2/2} = \left(\sum_{m\geq 0} \frac{x^m}{m!} t^m\right) \left(\sum_{m\geq 0} \frac{\cos(m\pi/2)t^m}{2^{m/2}(m/2)!}\right).$$

Thus

$$H_m(x) = m! \sum_{k=0}^m \frac{\cos(k\pi/2)}{2^{k/2}(k/2)!} \frac{x^{m-k}}{(m-k)!} = m! \sum_{0 \le k \le m/2} \frac{(-1)^k}{2^k \cdot k!} \frac{x^{m-2k}}{(m-2k)!}.$$
 (2.4)

What makes the Hermite polynomials significant is the fact that they form a basis of  $L_2(\mathbb{R}, \gamma_1)$ .

For  $m \in \mathbb{N}^n$  we define  $H_m : \mathbb{R}^n \to \mathbb{R}$  by

$$H_m(x) = \prod_{i=1}^n H_{m_i}(x_i)$$

One can prove that the collection  $(H_m)_{m \in \mathbb{N}^n}$  forms an orthonormal basis for  $L^2(\mathbb{R}^n, \gamma_n)$ .

# 2.5 Poisson parity process and continuous time random walk

Let  $N(t)_{t \in [0,\infty)}$  be the standard Poisson process. It is an iteger-valued Markov process with independent Poissonian increments, namely

$$N(0) = 0, \quad N(t) - N(s) \sim N(t-s) \sim \text{Pois} (t-s).$$

Recall that if  $Y \sim \text{Pois}(t-s)$ , then

$$\mathbb{P}(Y=k) = e^{-(t-s)} \frac{(t-s)^k}{k!}.$$

With probability one the trajectories of the process  $t \mapsto N(t)$  are non-decreasing integervalued functions. Define  $X(t) = (-1)^{N(t/2)}$ . One can prove that  $(X(t))_{t \in [0,\infty)}$  is also a Markov process. Take  $t > s \ge 0$ . Using the fact that the Poisson process has independent increments we compute

$$\begin{split} \mathbb{P}\left(X(t) = 1 | X(s) = 1\right) &= \mathbb{P}\left(X(t) = -1 | X(s) = -1\right) \\ &= \mathbb{P}\left(N(t/2) - N(s/2) \text{ is even } | N(s/2) \text{ is even }\right) \\ &= \mathbb{P}\left(N(t/2) - N(s/2) \text{ is even}\right) = \mathbb{P}\left(N((t-s)/2) \text{ is even}\right) \\ &= \operatorname{Pois}\left(\frac{t-s}{2}\right)(\{0,2,4,\ldots\}) \\ &= e^{-(t-s)}\sum_{k=0}^{\infty} \frac{(t-s)^{2k}}{(2k)!} = e^{-(t-s)/2} \frac{e^{(t-s)/2} + e^{-(t-s)/2}}{2} \\ &= (1+e^{-(t-s)})/2. \end{split}$$

Also

$$\mathbb{P}(X(t) = -1|X(s) = 1) = \mathbb{P}(X(t) = 1|X(s) = -1) = (1 - e^{-(t-s)})/2.$$

The process  $(X(t))_{t \in [0,\infty)}$  is time and space homogenous. It can be constructed via abstract argument by checking that the transition probabilities

$$p_{t,s}(x,y) = p_{t-s}(x,y)$$

that we have just computed satisfy the Chapman-Kolmogorov equations

$$p_{u-s}(x,z) = \sum_{y \in \{-1,1\}} p_{t-s}(x,y) p_{u-t}(y,z)$$

for all  $u > t > s \ge 0$  and  $x, z \in \{-1, 1\}$ .

Now we can construct a continuous time random walk on  $\Sigma_n$ . Let  $(X_1(t))_{t \in [0,\infty)}$ ,  $(X_2(t))_{t \in [0,\infty)}$ , ...,  $(X_n(t))_{t \in [0,\infty)}$  be independent copies of the process  $(X(t))_{t \in [0,\infty)}$ . For a given  $v \in \Sigma_n$  we set

$$\mathbf{X}^{v}(t) = (v_1 X_1(t), v_2 X_2(t), \dots, v_n X_n(t))$$

Clearly we have  $\mathbf{X}^{v}(0) = v$  and therefore v is a starting point of this process.

For a given  $f: \Sigma_n \to \mathbb{R}$  define a family of linear operators

$$(\mathcal{P}_t f)(v) = \mathbb{E}f(\mathbf{X}^v(t)) = \sum_{x \in \Sigma_n} p_t(v, x)f(x)$$

We have the following properties

- $\mathcal{P}_t \mathbf{1} = \mathbf{1}$  (it is called the invariance of  $\mu^n$  on  $\Sigma_n$ ),
- $f \ge 0$  a.s. implies  $\mathcal{P}_t f \ge 0$  a.s. (positivity preserving),
- $\mathcal{P}_{t+s} = \mathcal{P}_t \circ \mathcal{P}_s, \ \mathcal{P}_0 = \text{Id.}$

The third property indicates that the family  $(\mathcal{P}_t)_{t\geq 0}$  is a semigroup of operators. A semigroup, indexed by a time parameter  $t \in [0, \infty)$ , of linear operators satisfying the above three conditions is called Markovian.

Having a Markovian semigroup  $(\mathcal{P}_t)$  of operators on  $L^2(\Omega, \mu)$  one can try to define a Markov process reproducing this semigroup. We do not want to give precise formulations here. The idea is to take  $q_t(x, y) = (\mathcal{P}_t \mathbf{1}_y)(x)$  and, using the semigroup property, to prove that  $q_t(x, y)$  satisfies Chapman-Kolmogrov equations.

Now we want to look at the action of our semigroup on the Walsh functions. We have

$$(\mathcal{P}_t w_S)(v) = \mathbb{E} w_S(\mathbf{X}^v(t)) = \mathbb{E} \prod_{i \in S} v_i X_i(t) = \prod_{i \in S} v_i \cdot \prod_{i \in S} \mathbb{E} X_i(t)$$
$$= w_S(v) \left(\frac{1 + e^{-t}}{2} + \frac{1 - e^{-t}}{2}\right)^{|S|} = e^{-|S|t} w_S(v).$$

Hence  $\mathcal{P}_t w_S = e^{-|S|t} w_S$ . Therefore, if  $f = \sum_{S \subset [n]} a_S w_S$  then

$$\mathcal{P}_t f = \sum_{S \subset [n]} e^{-|S|t} a_S w_S.$$

Now one can see that indeed  $\mathcal{P}_{t+s} = \mathcal{P}_t \circ \mathcal{P}_s$ . Note also that  $\mathcal{P}_t$  is a multiplier is a Fourier representation.

**2.1 Remark.** The continuous time random walk on  $\Sigma_n$  can be constructed in another way that through the Poisson parity process. For  $\lambda \in (0, 1/2]$  we can consider a lazy random walk  $(Y_k)_{k \in \mathbb{N}}$  on  $\Sigma_n$  started from the point  $v \in \Sigma_n$ , namely a Markov chain with the transition probabilities  $\mathbb{P}(Y_{k+1} = x | Y_k = y) = \lambda/n$  when d(x, y) = 1 and  $\mathbb{P}(Y_{k+1} = x | Y_k = x) = 1 - \lambda$ . Of course  $Y_k = Y_k^{v,\lambda}$ . Let  $f_k(x) = \mathbb{P}(Y_k = x)$ . Take an operator

$$(Kf)(x) = \frac{1}{n} \sum_{Y:d(x,y)=1} f(y).$$

One can see that  $f_{k+1} = (\lambda K + (1 - \lambda) \operatorname{Id}) f_k$ . Hence,

$$f_k = f_k^{v,\lambda} = (\lambda K + (1 - \lambda) \operatorname{Id})^k f_0.$$

Define a discrete semigroup

$$(\mathcal{P}_k f)(v) = \mathbb{E}f(Y_k^{v,\lambda}) = \sum_{x \in \Sigma_n} \mathbb{P}\left(Y_k = x | Y_0 = v\right) f(x) = \sum_{x \in \Sigma_n} f_k(x) f(x).$$

Note that

$$Kw_S = (n - |S|)w_S - |S|w_S,$$

hence

$$(\lambda K + (1 - \lambda) \operatorname{Id}) w_S = (1 - \lambda) w_S + \lambda \left(1 - \frac{2|S|}{n}\right) w_S = \left(1 - \frac{2\lambda|S|}{n}\right) w_S.$$

We have

$$f_0(x) = \prod_{i=1}^n \frac{1 + v_i x_i}{2} = 2^{-n} \sum_{S \subset [n]} w_S(v) w_S(x).$$

Hence

$$f_k(x) = (\lambda K + (1 - \lambda) \operatorname{Id})^k f_0 = 2^{-n} (\lambda K + (1 - \lambda) \operatorname{Id})^k \left( \sum_{S \subset [n]} w_S(v) w_S(x) \right)$$
$$= 2^{-n} \sum_{S \subset [n]} w_S(v) \left( 1 - \frac{2\lambda |S|}{n} \right)^k w_S(x).$$

If  $f = \sum_{S \subset [n]} a_S w_S$  then

$$(\mathcal{P}_t f)(v) = \sum_{S \subset [n]} a_S \sum_{x \in \Sigma_n} f_k(x) w_S(x) = 2^{-n} \sum_{S, T \subset [n]} \sum_{x \in \Sigma_n} a_S w_S(v) w_T(x) w_S(x) \left(1 - \frac{2\lambda |S|}{n}\right)^k$$
$$= \sum_{S, T \subset [n]} a_S \left(1 - \frac{2\lambda |S|}{n}\right)^k \langle w_S, w_T \rangle w_S(v) = \sum_{S \subset [n]} a_S \left(1 - \frac{2\lambda |S|}{n}\right)^k w_S(v)$$

,

Now observe that

$$\lim_{\lambda \to 0^+} \mathcal{P}_{\lceil nt/2\lambda \rceil} f = \sum_{S \subset [n]} a_S e^{-|S|t} w_S.$$

Therefore our lazy random walk converges, as  $\lambda \to 0^+$  (and after appropriate time scaling), to the continuous time random walk.

#### 2.6 Ornstein-Uhlenbeck process

Consider a one dimensional stochastic equation

$$dX_t = -X_t \mathrm{d}t + \sqrt{2} \, \mathrm{d}W_t, \qquad X_0 = x_0.$$

where  $(W_t)_{t\geq 0}$  is a standard Wiener process on  $\mathbb{R}$ . The solution of this equation

$$X_t = x_0 e^{-t} + \sqrt{2} \int_0^t e^{s-t} \, \mathrm{d}W_s$$

can we written in the form

$$X_t = x_0 e^{-t} + e^{-t} W_{e^{2t} - 1}.$$

Clearly,

$$X_t = X_t^{x_0} \sim x_0 e^{-t} + \sqrt{1 - e^{-2t}}g,$$

where g is the standard normal random variable. The process is stationary, Gaussian, and Markovian. In fact it is the only nontrivial process that satisfies these three conditions, up to linear transformations of the space and time variables. Its transition kernel has the form  $f(x_i) = \frac{1}{2} e^{-\frac{1}{2}x_i}$ 

$$p_t(x,y) = \frac{1}{\sqrt{2\pi \left(1 - e^{-2t}\right)}} \exp\left(-\frac{(y - xe^{-t})^2}{2(1 - e^{-2t})}\right).$$

If we consider independent processes

$$(X_1(t))_{t\geq 0}, (X_2(t))_{t\geq 0}, \dots, (X_n(t))_{t\geq 0}$$

starting from points  $x_1, \ldots, x_n \in \mathbb{R}$  then for a function  $f : \mathbb{R}^n \to \mathbb{R}$  of moderate growth and  $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$  we can define

$$(\mathcal{P}_t f)(x) = \mathbb{E}f(X_t^x) = \int_{\mathbb{R}} f(xe^{-t} + \sqrt{1 - e^{-2t}}y) \, \mathrm{d}\gamma_n(y).$$

This is co-called Ornstein-Uhlenbeck semigroup of operators. One can check that it possesses the properties

- $\mathcal{P}_t \mathbf{1} = \mathbf{1}$ ,
- $f \ge 0$  a.s. implies  $\mathcal{P}_t f \ge 0$  a.s.,
- $\mathcal{P}_{t+s} = \mathcal{P}_t \circ \mathcal{P}_s, \ \mathcal{P}_0 = \text{Id.}$

Therefore it is a Markov semigroup of linear operators.

# 2.7 Generators

Having a Markov semigroup  $(\mathcal{P}_t)_{t\geq 0}$  we can define its generator

$$Lf = \lim_{t \to 0^+} \frac{\mathcal{P}_t f - f}{t} = \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{P}_t f \Big|_{t=0}$$

for functions for which this derivative exists. Note that  $L\mathbf{1} \equiv 0$ , since  $\mathcal{P}_0 = \mathrm{Id}$ .

We remember that in the case of discrete cube we had  $\mathcal{P}_t w_S = e^{-|S|t} w_S$ , therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_t w_S = -|S|e^{-|S|t}w_S = -|S|\mathcal{P}_t w_S$$

and

$$Lw_S = -|S|w_S.$$

Hence, since L is linear, we have

$$L\left(\sum_{S\subset[n]}a_Sw_S\right) = -\sum_{S\subset[n]}|S|a_Sw_S.$$

It also follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_t f = L\mathcal{P}_t f = \mathcal{P}_t L f.$$

The operators  $\mathcal{P}_t$  and L clearly commute since they are both multipliers. The above identity is general and follows from the semigroup property. Indeed,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_t f = \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{P}_{t+s} f\Big|_{s=0} = \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{P}_s(\mathcal{P}_t f)\Big|_{s=0} = L\mathcal{P}_t f$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_t f = \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{P}_{t+s} f\Big|_{s=0} = \frac{\mathrm{d}}{\mathrm{d}s}\mathcal{P}_t(\mathcal{P}_s f)\Big|_{s=0} = \mathcal{P}_t\left(\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{P}_s f\Big|_{s=0}\right) = \mathcal{P}_t L f.$$

For  $f = \sum_{S \subset [n]} a_S w_S$  we also have

$$\lim_{t \to \infty} \mathcal{P}_t f = \lim_{t \to \infty} \sum_{S \subset [n]} a_S e^{-|S|t} w_S = a_{\emptyset} w_{\emptyset} = a_{\emptyset} = \mathbb{E} f.$$

Define an operator

$$(\tilde{L}f)(x) = \frac{1}{2} \sum_{y:d(x,y)=1} (f(y) - f(x)).$$

We check that  $Lw_S = -|S|w_S$  and therefore  $\tilde{L} = L$ .

By writing the Fourier expansion of  $f, g: \Sigma_n \to \mathbb{R}$  one can see that the operators  $\mathcal{P}_t$ and L are symmetric in the sense that

$$\mathbb{E}(f\mathcal{P}_tg) = \mathbb{E}(g\mathcal{P}_tf), \quad \mathbb{E}(fLg) = \mathbb{E}(gLf).$$

Using the scalar product in  $L^2(\Sigma_n)$  we can write

$$\langle f, \mathcal{P}_t g \rangle = \langle \mathcal{P}_t f, g \rangle, \quad \langle f, Lg \rangle = \langle Lf, g \rangle.$$

If a Markov semigroup possesses these properties, it is called symmetric.

Now we would like to find a generator for the Ornstein-Uhlenbeck semigroup. We have

$$\begin{split} \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{P}_t f)(x) &= \frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) \,\mathrm{d}\gamma_n(y) \\ &= \int_{\mathbb{R}^n} \sum_{i=1}^n \frac{\partial f}{\partial x_i} \left( e^{-t}x + \sqrt{1 - e^{-2t}}y \right) \left( -e^{-t}x_i + \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}}y_i \right) \,\mathrm{d}\gamma_n(y) \\ &= -\nabla((\mathcal{P}_t f)(x)) \cdot x \\ &\quad + \int_{\mathbb{R}^n} \sum_{i=1}^n \frac{\partial f}{\partial x_i} \left( e^{-t}x + \sqrt{1 - e^{-2t}}y \right) \frac{e^{-2t}}{\sqrt{1 - e^{-2t}}} \frac{\partial}{\partial y_i} \frac{-1}{(2\pi)^{n/2}} e^{-||y||_2^2} \,\mathrm{d}y \\ &= -\nabla((\mathcal{P}_t f)(x)) \cdot x + \int_{\mathbb{R}^n} \sum_{i=1}^n \frac{\partial f^2}{\partial x_i^2} \left( e^{-t}x + \sqrt{1 - e^{-2t}}y \right) e^{-2t} \,\mathrm{d}\gamma_n(y) \\ &= -\nabla((\mathcal{P}_t f)(x)) \cdot x + \Delta((\mathcal{P}_t f)(x)). \end{split}$$

Hence

$$(Lf)(x) = (\Delta f)(x) - (\nabla f(x)) \cdot x.$$

Now we would like to investigate how the operator L and  $\mathcal{P}_t$  act on Hermite polynomials.

#### 2.1 Proposition. In one dimensional case we have

(i)  $H'_m = mH_{m-1} \text{ for } m \ge 1,$ 

(ii)  $LH_m = -mH_m$  for  $m \ge 0$ ,

(iii) 
$$\mathcal{P}_t H_m = e^{-tm} H_m$$
 for  $m \ge 0$ .

Moreover, in multidimensional case for any  $m \in \mathbb{N}^n$ 

(iv) 
$$LH_m = -|m|H_m$$
,

(v)  $P_t H_m = e^{-t|m|} H_m$ .

*Proof.* (i) It follows by formula (2.4) that the leading monomial in  $H_m$  is equal to  $x^m$ . It implies that  $H'_m - mH_{m-1}$  is a polynomial of degree m - 2. Therefore it suffices to check that for every  $i \leq m - 2$  we have  $\langle H'_m - mH_{m-1}, H_i \rangle = 0$ . Integrating by parts we obtain

$$\langle H'_m - mH_{m-1}, H_i \rangle = \langle H'_m, H_i \rangle = \int H'_m(x)H_i(x)e^{-x^2/2}\frac{\mathrm{d}x}{\sqrt{2\pi}}$$
$$= -\int H_m(x)(H'_i(x) - xH_i(x))\,\mathrm{d}\gamma = 0,$$

as the polynomial  $H'_i(x) - xH_i(x)$  is of degree  $i + 1 \le m - 1$ .

(ii) Since the polynomial  $LH_m + mH_m = H''_m - xH'_m + mH_m$  is of degree m-1 we conclude observing that for any  $i \le m-1$ 

$$\langle LH_m + mH_m, H_i \rangle = \langle LH_m, H_i \rangle = \langle H_m, LH_i \rangle = 0,$$

as  $LH_i$  is of degree *i*.

(*iii*) Note that point (*ii*) yields

$$LP_tH_m = P_tLH_m = -mP_tH_m,$$

so that  $P_t H_m$  is an eigenvector of L. As a consequence there is a number  $\lambda(t)$  such that  $P_t H_m = \lambda(t) H_m$ . We get

$$\lambda'(t)H_m = \partial_t P_t H_m = L P_t H_m = -m\lambda(t)H_m$$

Moreover,

$$H_m = P_0 H_m = \lambda(0) H_m.$$

Thus  $\lambda$  solves the Cauchy problem

$$\begin{cases} \dot{\lambda} &= -m\lambda \\ \lambda(0) &= 1 \end{cases},$$

so  $\lambda(t) = e^{-tm}$ .

	_	_	_	-
н				

(*iv*) Since  $H_m(x) = \prod_{i=1}^n H_{m_i}(x_i)$  we have

$$\Delta H_m = \sum_{i=1}^n \partial_{x_i x_i} \left( \prod_{j=1}^n H_{m_j}(x_j) \right) = \sum_{i=1}^n H''_{m_i}(x_i) \prod_{j \neq i} H_{m_j}(x_j),$$
$$x \cdot \nabla H_m(x) = \sum_{i=1}^n x_i \partial_{x_i} \left( \prod_{j=1}^n H_{m_j}(x_j) \right) = \sum_{i=1}^n H'_{m_i}(x_i) \prod_{j \neq i} H_{m_j}(x_j),$$

thus, by the one dimensional result concerning L,

$$LH_m = (\Delta - x \cdot \nabla) H_m = \sum_{i=1}^n \left( H''_{m_i}(x_i) - x_i H'_{m_i}(x_i) \right) \prod_{j \neq i} H_{m_j}(x_j)$$
$$= \sum_{i=1}^m -m_i H_m(x) = -|m| H_m.$$

(v) The second formula follows by its one dimensional counterpart applied to  $H_{m_i}(x_i)$  as the operator  $P_t$  acts on each coordinate  $x_i$  independently because the Gaussian measure is a product measure.

## 2.8 Contractivity

Suppose we have a symmetric Markov semigroup. Then

$$\mathbb{E}\mathcal{P}_t f = \mathbb{E}\mathbf{1}(\mathcal{P}_t f) = \mathbb{E}(\mathcal{P}_t \mathbf{1})f = \mathbb{E}\mathbf{1}f = \mathbb{E}f$$

for every  $t \ge 0$ . Therefore the semigroup preserves expectation.

Now we prove that the semigroup  $(\mathcal{P}_t)_{t\geq 0}$  is contactive in  $L_p$  for every  $p\geq 1$ , i.e.

$$\|\mathcal{P}_t f\|_p \le \|f\|_p, \quad t \ge 0, p \ge 1.$$

**2.2 Lemma.** Let  $\Phi : \mathbb{R} \to \mathbb{R}$  be a convex function. Then for every  $t \ge 0$  we have

$$\mathbb{E}\Phi(\mathcal{P}_t f) \le \mathbb{E}\Phi(f).$$

Taking  $\Phi(x) = |x|^p$  for  $p \ge 1$  we obtain  $\|\mathcal{P}_t f\|_p \le \|f\|_p$ .

*Proof.* Every convex function is a supremum of its supporting affine functions, i.e.

$$\Phi(x) = \sup_{\alpha} (a_{\alpha}x + b_{\alpha})$$

for some real numbers  $(a_{\alpha})_{\alpha \in \mathbb{R}}$ ,  $(b_{\alpha})_{\alpha \in \mathbb{R}}$ . Note that  $f \geq g$  implies  $\mathcal{P}_t(f-g) = \mathcal{P}_t f - \mathcal{P}_t g \geq 0$  (positivity preserving). Thus,

$$\mathcal{P}_t(\Phi(f)) \ge \mathcal{P}_t(a_\alpha f + b_\alpha) = a_\alpha \mathcal{P}_t f + b_\alpha.$$

Therefore, taking supremum of the right hand side,

$$\mathcal{P}_t(\Phi(f)) \ge \Phi(\mathcal{P}_t f).$$

Using the fact that our semigroup preserves expectation we obtain

$$\mathbb{E}\Phi(f) = \mathbb{E}\mathcal{P}_t(\Phi(f)) \ge \mathbb{E}\Phi(\mathcal{P}_t f)$$

Note also that if  $-m \leq f \leq m$  for some real number m then

$$-m = \mathcal{P}_t(-m) \le \mathcal{P}_t f \le \mathcal{P}_t m = m,$$

thus  $\|\mathcal{P}_t f\|_{\infty} \leq \|f\|_{\infty}$ .

### 2.9 Dirichlet forms and integration by parts

Consider the case of a discrete cube  $\Sigma_n$  and a semigroup of a continuous time random walk. Take a bilinear form

$$\mathcal{E}(f,g) = \mathbb{E}f(-L)g.$$

Note that because of the symmetry of the semigroup we have  $\mathcal{E}(f,g) = \mathcal{E}(g,f)$ . We prove that the bilinear form  $\mathcal{E}$  is nonnegative, i.e. we have the following lemma

#### **2.3 Lemma.** We have $\mathcal{E}(f, f) \geq 0$ .

*Proof.* Take  $\Psi(t) = \mathbb{E}(\mathcal{P}_t f)^2$ . We have

$$\Psi'(t) = 2\mathbb{E}\left((\mathcal{P}_t f)\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_t f\right) = 2\mathbb{E}(\mathcal{P}_t f)(L\mathcal{P}_t f).$$

Hence,

$$\Psi'(0^+) = 2\mathbb{E}(fLf) = -2\mathcal{E}(f, f).$$

Because of the contraction we have

$$\Psi(t) = \|\mathcal{P}_t f\|_2^2 \le \|f\|_2^2 = \|\mathcal{P}_0 f\|_2^2 = \Psi(0).$$

Thus  $\Psi'(0^+) \leq 0$  and therefore  $\mathcal{E}(f, f) \geq 0$ .

The same proof works for a large class of symmetric Markov processes, for example in the case of the Ornstein-Uhlenbeck semigroup for appropriate class of functions. In our examples we can show the above fact by a direct calculation.

## (a) discrete cube

In this case we have

$$(Lf)(x) = \frac{1}{2} \sum_{y: \ d(x,y)=1} \left( f(y) - f(x) \right).$$

Therefore

$$\mathcal{E}(f,g) = -\mathbb{E}fLg = -2^{-n-1} \sum_{x \in \Sigma_n} \sum_{y: \ d(x,y)=1} f(x)(g(y) - g(x))$$
$$= 2^{-n-1} \sum_{(x,y): \ d(x,y)=1} \left( f(x)g(x) - f(x)g(y) \right).$$

Replacing indexes x and y we obtain also

$$\mathcal{E}(f,g) = 2^{-n-1} \sum_{(x,y): \ d(x,y)=1} \left( f(y)g(y) - f(y)g(x) \right).$$

Thus, adding these two equalities together we obtain

$$\begin{aligned} \mathcal{E}(f,g) &= 2^{-n-2} \sum_{(x,y):\ d(x,y)=1} \left( f(x)g(x) - f(x)g(y) + f(y)g(y) - f(y)g(x) \right) \\ &= 2^{-n} \sum_{(x,y):\ d(x,y)=1} \left( \frac{f(x) - f(y)}{2} \right) \left( \frac{g(x) - g(y)}{2} \right). \end{aligned}$$

If we take  $\nabla f = (\nabla_1 f, \dots, \nabla_n f)$ , where

$$(\nabla_i f)(x) = \frac{f(x) - f(\sigma_i(x))}{2}, \quad \sigma_i(x_1, \dots, x_i, \dots, x_n) = (x_1, \dots, -x_i, \dots, x_n),$$

then we obtain

$$\mathcal{E}(f,g) = \int_{\Sigma_n} \nabla f \cdot \nabla g \, \mathrm{d}\mu^n.$$

Moreover,

$$\mathcal{E}(f,f) = \int_{\Sigma_n} |\nabla f|^2 \, \mathrm{d}\mu^n \ge 0.$$

The equality

$$\int f Lg \, \mathrm{d}\mu^n = -\int \nabla f \cdot \nabla g \, \mathrm{d}\mu^n$$

can be interpreted as an integration by parts formula. The operator L is often called the *discrete Laplace operator*.

#### (b) Gaussian space

In this case we have

$$(Lf)(x) = \Delta f(x) - x \cdot \nabla f(x)$$

Thus,

$$\begin{aligned} \mathcal{E}(f,g) &= -\int fLg \, \mathrm{d}\gamma_n = -\int_{\mathbb{R}^n} f(x) \left(\Delta g(x) - x \cdot \nabla g(x)\right) \, \mathrm{d}\gamma_n(x) \\ &= -\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} f(x) \left(\Delta g(x) - x \cdot \nabla g(x)\right) e^{-\|x\|_2^2/2} \, \mathrm{d}x \\ &= -\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \nabla \left(f(x) e^{-\|x\|_2^2/2}\right) \cdot \nabla g \, \mathrm{d}x + \int_{\mathbb{R}^n} f(x) \left(x \cdot \nabla g(x)\right) \, \mathrm{d}\gamma_n(x) \\ &= -\int \nabla f \cdot \nabla g \, \mathrm{d}\gamma_n - \int_{\mathbb{R}^n} f(x) \left(x \cdot \nabla g(x)\right) \, \mathrm{d}\gamma_n(x) + \int_{\mathbb{R}^n} f(x) \left(x \cdot \nabla g(x)\right) \, \mathrm{d}\gamma_n(x) \\ &= -\int \nabla f \cdot \nabla g \, \mathrm{d}\gamma_n. \end{aligned}$$

Thus

$$\int f Lg \, \mathrm{d}\gamma_n = -\int \nabla f \cdot \nabla g \, \mathrm{d}\gamma_n.$$

This formula is called the Gaussian integration by parts. Clearly we have

$$\mathcal{E}(f,f) = \int |\nabla f|^2 \, \mathrm{d}\gamma_n \ge 0.$$

The quantity  $\mathcal{E}(f, f)$  is sometimes called the *energy functional* since it is analogous to the kinetic energy in classical mechanics. We adopt the notation  $\mathcal{E}(f) = \mathcal{E}(f, f)$ .

The following lemma states that energy is sable under Lipschitz maps.

**2.4 Lemma.** Let  $\Psi : \mathbb{R} \to \mathbb{R}$  be *C*-Lipschitz. Then for a symmetric Markov semigroup we have  $\mathcal{E}(\Psi(f)) \leq C\mathcal{E}(f)$ . In particular, if  $\Psi(a) = |a|$  then we obtain

$$\mathbb{E}|f|(-L)|f| = \mathcal{E}(|f|) \le \mathcal{E}(f) = \mathbb{E}f(-L)f.$$

*Proof.* Let  $\mathbb{X}^{x}(\cdot)$  be a process associated with our semigroup. Applying the inequality

$$(\Psi(a) - \Psi(b))^2 \le C^2(a-b)^2$$

with a = f and  $b = f(\mathbb{X}^x(t)), t \ge 0$  we obtain

$$\Psi(f(x))^{2} - 2\Psi(f(x))\Psi(f(\mathbf{X}^{x}(t))) + \Psi(f(\mathbf{X}^{x}(t)))^{2}$$
  
$$\leq C^{2} \left( f(x)^{2} - 2f(x)f(\mathbf{X}^{x}(t)) + f(\mathbf{X}^{x}(t))^{2} \right)$$

By taking expectation of both sides we obtain

$$\Psi(f)^2 - 2\Psi(f)\mathcal{P}_t(\Psi(f)) + \mathcal{P}_t(\Psi(f)^2) \le C^2 \left(f^2 - 2f\mathcal{P}_t f + \mathcal{P}_t(f^2)\right).$$

Again, we take expectation and arrive at

$$0 \ge \alpha(t) = \mathbb{E}\Psi(f)^2 - 2\mathbb{E}(\Psi(f)\mathcal{P}_t\Psi(f)) + \mathbb{E}\mathcal{P}_t(\Psi(f)^2) - C^2\mathbb{E}f^2 + 2C^2\mathbb{E}(f\mathcal{P}_tf) - C^2\mathbb{E}\mathcal{P}_t(f^2) = 2\mathbb{E}\Psi(f)^2 - 2\mathbb{E}(\Psi(f)\mathcal{P}_t\Psi(f)) - 2C^2\mathbb{E}f^2 + 2C^2\mathbb{E}(f\mathcal{P}_tf).$$

Since  $\mathcal{P}_0 = \text{Id}$  we have  $\alpha(0) = 0$  and therefore  $\alpha'(0^+) \leq 0$ , namely

$$-2\mathbb{E}\Psi(f)L\Psi(f) + 2C^2\mathbb{E}fLf \le 0.$$

# 2.10 Poincare inequality

**2.5 Definition.** We say that a probability Borel measure  $\nu$  on  $\mathbb{R}^n$  satisfies Poincare inequality with constant C if for every  $C^1$  function  $f : \mathbb{R}^n \to \mathbb{R}$  such that  $\int_{\mathbb{R}^n} f \, d\nu < \infty$  we have

$$\int_{\mathbb{R}^n} f^2 \,\mathrm{d}\nu - \left(\int_{\mathbb{R}^n} f \,\mathrm{d}\nu\right)^2 \le C \int_{\mathbb{R}^n} |\nabla f|^2 \,\mathrm{d}\nu.$$

It means that

$$\operatorname{Var}_{\nu}(f) \leq C \int |\nabla f|^2 \, \mathrm{d}\nu,$$

therefore it is also called the variance-energy inequality.

In the case on the discrete cube one can take the energy functional  $\mathcal{E}(f, f) = \mathbb{E}f(-L)f$ instead on  $|\nabla f|^2$  or adopt the notation of the discrete gradient. We prove the following theorem

**2.6 Theorem.** Let  $f: \Sigma_n \to \mathbb{R}$ . Then

$$\mathbb{E}f^2 - (\mathbb{E}f)^2 \le -\mathbb{E}(fLf).$$

Moreover, if f is even, i.e. f(-x) = f(x) for all  $x \in \Sigma_n$ , then

$$\mathbb{E}f^2 - (\mathbb{E}f)^2 \le \frac{1}{2} - \mathbb{E}(fLf).$$

*Proof.* Expand a given function  $f : \Sigma_n \to \mathbb{R}$  in the Fourier series  $f = \sum_{S \subset [n]} a_s w_S$ . Recall that

$$Lf = -\sum_{S \subset [n]} |S| a_s w_S, \quad \mathbb{E}f^2 = \sum_{S \subset [n]} a_s^2, \quad \mathbb{E}f = a_{\emptyset}.$$

Thus,

$$\mathbb{E}f^2 - (\mathbb{E}f)^2 = \sum_{S \subset [n], S \neq \emptyset} a_s^2 \le \sum_{S \subset [n], S \neq \emptyset} |S| a_s^2 = \sum_{S \subset [n]} |S| a_s^2 = -\mathbb{E}(fLf).$$

To prove the second assertion, note that if f is even then for all  $S \subset [n]$  with |S| odd the function  $w_S$  is odd and we have

$$a_s = \langle f, w_S \rangle = \mathbb{E}(fw_S) = 0.$$

In particular  $a_{\{i\}} = 0$  for all  $i \in [n]$ . Thus

$$\begin{split} \mathbb{E}f^2 - (\mathbb{E}f)^2 &= \sum_{S \subset [n], S \neq \emptyset} a_s^2 = \sum_{S \subset [n], |S| \ge 2} a_s^2 \le \sum_{S \subset [n], |S| \ge 2} \frac{|S|}{2} a_s^2 \\ &= \sum_{S \subset [n], S \neq \emptyset} \frac{|S|}{2} a_s^2 = -\frac{1}{2} \mathbb{E}(fLf). \end{split}$$

We can give the same proof in the Gaussian case by replacing Walsh functions by Hermite polynomials and using the fact that for the generator of the Ornstein-Uhlenbeck semigroup we have  $LH_m = -|m|H_m$ , where  $m \in \mathbb{N}^n$ . Therefore we can state the following theorem

**2.7 Theorem.** For every  $C^1$  function  $f : \mathbb{R}^n \to \mathbb{R}$  such that  $\int_{\mathbb{R}^n} f \, \mathrm{d}\gamma_n < \infty$  we have

$$\operatorname{Var}_{\gamma_n}(f) \leq \int |\nabla f|^2 \, \mathrm{d}\gamma_n.$$

Moreover, if f is even then

$$\operatorname{Var}_{\gamma_n}(f) \leq \frac{1}{2} \int |\nabla f|^2 \, \mathrm{d}\gamma_n.$$

Poincare inequalities are also called *spectral gap inequalities*. As we have seen in the above proof, they hold because there is a gap in the spectrum of  $\sigma(-L)$  between 0 and  $\sigma(-L|_{f:\mathbb{E}f=0})$ .

#### 2.11 Khinchine-Kahane inequality

Consider a sequence of independent symmetric Bernoulli random variables  $r_1, r_2, \ldots, r_n$ and a sequence of real numbers  $a_1, a_2, \ldots, a_n$ . The classical Khinchine inequality states that for p, q > 0 there exists a constant  $C_{p,q}$  independent of n and of a choice of  $a_1, \ldots, a_n$ such that

$$\left(\mathbb{E}\left|\sum_{i=1}^{n} a_{i}r_{i}\right|^{p}\right)^{1/p} \leq C_{p,q} \left(\mathbb{E}\left|\sum_{i=1}^{n} a_{i}r_{i}\right|^{q}\right)^{1/q}$$

It was shown by Khinchine in (?). One would like to find the optimal constants  $C_{p,q}^{(K)}$  in this inequality. Clearly, if  $0 then <math>C_{p,q}^{(K)} = 1$ . However, the case 0 < q < p is nontrivial. Khinchine himself found the constants  $C_{p,2}^{(K)}$  for p even. In 1960 Whittle found  $C_{p,2}^{(K)}$  for all  $p \geq 3$ . In 1976 Szarek proved that  $C_{2,1}^{(K)} = \sqrt{2}$ . Finally, U. Haagerup found  $C_{p,2}^{(K)}$  for p > 2 and  $C_{2,q}^{(K)}$  for 0 < q < 2.

For p > 2 we always have  $C_{p,2}^{(K)} = \frac{\gamma_p}{\gamma_2}$ , where  $\gamma_p = ||g||_p = (\mathbb{E}|g|^p)^{1/p}$ , where g is the standard normal random variable. In the case of constants  $C_{2,q}^{(K)}$  there exists  $p_0 \in (1,2)$  such that

$$C_{2,q}^{(K)} = \begin{cases} \frac{\gamma_2}{\gamma_q} & p_0 < q < 2\\ 2^{\frac{1}{q} - \frac{1}{2}} & 0 < q \le p_0 \end{cases}$$

and  $p_0$  is the solution of the equation  $\gamma_2/\gamma_{p_0} = 2^{\frac{1}{p_0} - \frac{1}{2}}$ .

The constants  $\gamma_p/\gamma_2$  for p > 2 and  $\gamma_2/\gamma_q$  for  $p_0 < q < 2$  are achieved asymptotically by taking  $a_1 = a_2 = \ldots = a_n = 1/\sqrt{n}$  and letting  $n \to \infty$ . This follows from the Central Limit Theorem. On the other hand, the constant  $2^{\frac{1}{q}-\frac{1}{2}}$  is achieved when  $a_1 = a_2 = 1$ and  $a_i = 0$  for  $i \ge 3$ .

Recently it was shown, see (?), that  $C_{p,q}^{(K)} = \gamma_p / \gamma_q$  for all even p > q > 0. Moreover, for p > q > 1 we have  $C_{p,q}^{(K)} \leq \sqrt{\frac{p-1}{q-1}}$ . Later on we will obtain this result by using some hypercontractivity arguments. Note also that by CLT we always have  $C_{p,q}^{(K)} \geq \gamma_p / \gamma_q$ . Since  $\gamma_p \approx \sqrt{p}$  for large values of p, the constant  $\sqrt{\frac{p-1}{q-1}}$  is asymptotically of good order.

One can consider a general case of Khinchine inequality in an arbitrary normed space. Let  $(F, \|\cdot\|)$  be a normed space and let  $v_1, v_2, \ldots, v_n \in F$ . J.P. Kahane proved that for p > q > 0 there exists a constant  $C_{p,q}$  such that

$$\left(\mathbb{E}\left\|\sum_{i=1}^{n} r_{i} v_{i}\right\|^{p}\right)^{1/p} \leq C_{p,q} \left(\mathbb{E}\left\|\sum_{i=1}^{n} r_{i} v_{i}\right\|^{q}\right)^{1/q}$$

The constant  $C_{p,q}$  is independent of a space F, the number n and  $v_1, v_2, \ldots, v_n$ . This is so called Khinchine-Kahane inequality. Let us denote by  $C_{p,q}^{(K-K)}$  the optimal constant in this inequality. The optimal constants are not known in general. It is also not known if they coincide with the constants  $C_{p,q}^{(K)}$ . Clearly we have  $C_{p,q}^{(K-K)} \ge C_{p,q}^{(K)}$  even for a fixed normed space, since we can take  $v_i = a_i v$ , where  $v \in F$  is fixed vector. In this chapter we prove the following theorems

**2.8 Theorem.** We have  $C_{2,1}^{(K-K)} = C_{2,1}^{(K)} = \sqrt{2}$ .

**2.9 Theorem.** We have  $C_{4,2}^{(K-K)} = C_{4,2}^{(K)} = \sqrt[4]{3}$ .

In the proofs we use the fact that functions  $w_{\{i\}}: \Sigma_n \to \{-1, 1\}$  forms a sequence of independent symmetric Bernoulli random variables. Hence, for a function  $f: \Sigma_n \to \mathbb{R}$  we have

$$\mathbb{E}f(r_1,\ldots,r_n) = \int_{\Sigma_n} f(x) \, \mathrm{d}\mu^n(x).$$

Proof of Theorem 2.8. Consider a function  $H: \Sigma_n \to \mathbb{R}$  given by

$$H(x) = \left\| \sum_{i=1}^n x_i v_i \right\|.$$

Note that clearly  $H \ge 0$  and H is even, i.e. H(-x) = H(x) for all  $x \in \Sigma_n$ . By the triangle inequality we have

$$\sum_{y: d(x,y)=1} H(y) = \|-x_1v_1 + x_2v_2 + \ldots + x_nv_n\| + \|x_1v_1 - x_2v_2 + \ldots + x_nv_n\| + \ldots + \|x_1v_1 + x_2v_2 + \ldots - x_nv_n\| \\ \ge (n-2) \|x_1v_1 + x_2v_2 + \ldots + x_nv_n\|.$$

Thus

$$(-L)H(x) = \frac{1}{2} \sum_{y: \ d(x,y)=1} (H(x) - H(y)) \le \frac{n}{2}H(x) - \frac{n-2}{2}H(x) = H(x).$$

Therefore, using Poincare inequality for even functions we obtain

$$\mathbb{E}H^2 - (\mathbb{E}H)^2 \le \frac{1}{2}\mathbb{E}(H(-L)H) \le \frac{1}{2}\mathbb{E}H^2.$$

We arrive at  $\mathbb{E}H^2 \leq 2(\mathbb{E}H)^2$  which is exactly

$$\left(\mathbb{E}\left\|\sum_{i=1}^{n} r_{i} v_{i}\right\|^{2}\right)^{1/2} \leq \sqrt{2} \cdot \mathbb{E}\left\|\sum_{i=1}^{n} r_{i} v_{i}\right\|.$$

The constant  $\sqrt{2}$  is achieved in the case when  $v_1 = v_2 \neq 0$  and  $v_i = 0$  for  $i \geq 3$ .

To prove Theorem 2.9 we need the following proposition.

**2.10 Proposition.** [Stroock-Varopoulos inequality] For any p > 1 and  $f : \Sigma_n \to [0, \infty)$  there is

$$\mathbb{E}\left(f^{p/2}(-L)(f^{p/2})\right) \le \frac{p^2}{4(p-1)} \mathbb{E}\left(f^{p-1}(-L)f\right).$$

The same inequality holds for any generator of a symmetric Markov semigroup (under some technical assumptions on the function f).

We need a simple lemma.

**2.11 Lemma.** For any p > 1 and  $a, b \ge 0$  the following inequality holds

$$(p-2)^{2}(a^{p}+b^{p}) - p^{2}(a^{p-1}b+ab^{p-1}) + 8(p-1)a^{p/2}b^{p/2} \ge 0.$$

Proof of Lemma 2.11. Because of the homogenity and symmetry in a and b we can assume that  $a \ge 1$  and b = 1. Thus it suffices to prove that

$$u(t) = (p-2)^{2}t^{p} - p^{2}t^{p-1} + 8(p-1)t^{p/2} - p^{2}t + (p-2)^{2} \ge 0, \quad t \ge 1.$$

We have

$$u(1) = (p-2)^2 - p^2 + 8(p-1) - p^2 + (p-2)^2 = 0.$$

Therefore it suffices to prove the inequality

$$u'(t) = p(p-2)^{2}t^{p-1} - p^{2}(p-1)t^{p-2} + 4p(p-1)t^{\frac{p}{2}-1} - p^{2} \ge 0, \quad t \ge 1.$$

Again

$$u'(1) = p(p-2)^2 - p^2(p-1) + 4p(p-1) - p^2 = p((p-2)^2 - p(p-1) + 4(p-1) - p)$$
  
=  $p(p^2 - 4p + 4 - p^2 + p + 4p - 4 - p) = 0.$ 

Thus, it suffices to prove that

$$\begin{split} u''(t) &= p(p-1)(p-2)^2 t^{p-2} - p^2(p-1)(p-2)t^{p-3} + 2p(p-1)(p-2)t^{\frac{p}{2}-2} \\ &= 2p(p-1)(2-p)t^{p-2} \left(\frac{2-p}{2} + \frac{p}{2}t^{-1} - t^{-\frac{p}{2}}\right) \\ &= p^2(p-1)(p-2)t^{p-2} \left(\frac{p-2}{p} + \frac{2}{p}t^{-p/2} - t^{-1}\right). \end{split}$$

If  $p \ge 2$  then by the inequality  $\alpha a + (1 - \alpha)b \ge a^{\alpha}b^{1-\alpha}$ ,  $\alpha \in [0, 1]$ , a, b > 0 (concavity of ln) we have

$$\frac{p-2}{2} \cdot 1 + \frac{2}{p}t^{-p/2} \ge 1^{\frac{p-2}{p}}(t^{-\frac{p}{2}})^{\frac{2}{p}} = t^{-1}.$$

Therefore in this case

$$u''(t) = p^{2}(p-1)(p-2)t^{p-2}\left(\frac{p-2}{p} + \frac{2}{p}t^{-p/2} - t^{-1}\right) \ge 0.$$

If  $p \in (1,2)$  it suffices to show that

$$\frac{2-p}{2} + \frac{p}{2}t^{-1} - t^{-\frac{p}{2}} \ge 0.$$

We have

$$\frac{2-p}{2} \cdot 1 + \frac{p}{2}t^{-1} \ge 1^{\frac{2-p}{2}}(t^{-1})^{\frac{p}{2}} = t^{-p/2}.$$

Proof of Stroock-Varopoulos inequality. Take a nonnegative  $f : \Sigma_n \to \mathbb{R}$ . From the Lemma 2.11 we have

$$(p-2)^2(a^p+f^p) - p^2(a^{p-1}f + af^{p-1}) + 8(p-1)a^{p/2}f^{p/2} \ge 0.$$

Since  $\mathcal{P}_t$  preserves positivity, we have

$$(p-2)^2(a^p + \mathcal{P}_t(f^p)) - p^2(a^{p-1}\mathcal{P}_t(f) + a\mathcal{P}_t(f^{p-1})) + 8(p-1)a^{p/2}\mathcal{P}_t(f^{p/2}) \ge 0.$$

Setting a = f and taking expectation we obtain

$$(p-2)^2 \left(\mathbb{E}f^p + \mathbb{E}\mathcal{P}_t(f^p)\right) - p^2 \left(\mathbb{E}(f^{p-1}\mathcal{P}_t f) + \mathbb{E}(f\mathcal{P}_t(f^{p-1}))\right) + 8(p-1)\mathbb{E}\left(f^{p/2}\mathcal{P}_t(f^{p/2})\right) \ge 0.$$

Note that

$$\mathbb{E}\mathcal{P}_t(f^p) = \mathbb{E}f^p, \qquad \mathbb{E}f\mathcal{P}_t(f^{p-1}) = \mathbb{E}(\mathcal{P}_t f)f^{p-1}.$$

Hence

$$\beta(t) = 2(p-2)^2 \mathbb{E}f^p - 2p^2 \mathbb{E}(f^{p-1}\mathcal{P}_t f) + 8(p-1)\mathbb{E}\left(f^{p/2}\mathcal{P}_t(f^{p/2})\right) \ge 0.$$

Since  $\mathcal{P}_{-}f = f$ , we have

$$\beta(0) = \left(2(p-2)^2 - 2p^2 + 8(p-1)\right)\mathbb{E}f^p = 0.$$

Thus  $\beta'(0^+) \ge 0$ . But  $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_t f = Lf$  and therefore

$$\beta'(0^+) = -2p^2 \mathbb{E}(f^{p-1}Lf) + 8(p-1)\mathbb{E}\left(f^{p/2}Lf^{p/2}\right) \ge 0.$$

We arrive at

$$\mathbb{E}\left(f^{p/2}(-L)(f^{p/2})\right) \le \frac{p^2}{4(p-1)} \mathbb{E}\left(f^{p-1}(-L)f\right).$$

-		
L		
L		

Note that it the case of Ornstein-Uhlenbeck semigroup there is equality in the Stroock-Varopoulos inequality. Indeed, we have  $\int fLg d\gamma_n = -\int \nabla f \cdot \nabla g d\gamma_n$ , hence we obtain

$$\mathbb{E}\left(f^{p/2}(-L)(f^{p/2})\right) = \mathbb{E}\nabla(f^{p/2}) \cdot \nabla(f^{p/2}) = \frac{p^2}{4}\mathbb{E}f^{p-2}|\nabla f|^2$$

and

$$\frac{p^2}{4(p-1)}\mathbb{E}\left(f^{p-1}(-L)f\right) = \frac{p^2}{4(p-1)}\mathbb{E}\nabla(f^{p-1})\cdot\nabla f = \frac{p^2}{4}\mathbb{E}f^{p-2}|\nabla f|^2.$$

Proof of Theorem 2.9. As in the proof of Theorem 2.8 we take  $H(x) = \|\sum_{i=1}^{n} x_i v_i\|$ . The function  $H^2$  is even, therefore

$$\mathbb{E}H^4 - (\mathbb{E}H^2)^2 \le -\frac{1}{2}\mathbb{E}\left(H^2L(H^2)\right)$$

Taking p = 4 and f = H in the Stroock-Varopoulos inequality we obtain

$$-\mathbb{E}\left(H^2L(H^2)\right) \le \frac{4}{3}\mathbb{E}(H^3LH) \le \frac{4}{3}\mathbb{E}H^4,$$

since  $-LH \leq H$ . Hence

$$\mathbb{E}H^4 - (\mathbb{E}H^2)^2 \le \frac{1}{2} \cdot \frac{4}{3} \mathbb{E}H^4.$$

Thus  $\frac{1}{3}\mathbb{E}H^4 \leq (\mathbb{E}H^2)^2$ , i.e.

$$\left(\mathbb{E}\left\|\sum_{i=1}^{n}r_{i}v_{i}\right\|^{4}\right)^{1/4} \leq \sqrt[4]{3}\left(\mathbb{E}\left\|\sum_{i=1}^{n}r_{i}v_{i}\right\|^{2}\right)^{1/2}.$$

# 2.12 Hypercontractivity and Log-Sobolev inequalities

We begin this paragraph with a definition.

**2.12 Definition.** We say that a Markov semigroup with an invariant measure  $\mu$  an a generator L satisfies Log-Sobolev inequality with a constant C if

$$\mathbb{E}_{\mu}(f^2 \ln f^2) - (\mathbb{E}_{\mu}f^2) \ln(\mathbb{E}\mu f^2)) \le C\mathbb{E}_{\mu}(f(-L)f),$$

namely

$$\operatorname{Ent}_{\mu}(f^2) \le C\mathcal{E}(f, f).$$

We have already proved that the Ornstein-Uhlenbeck semigroup and the continuous time random walk on  $\Sigma_n$  satisfy Log-Sobolev inequality with constant 2. Note that this is because in these two cases we have

$$\mathcal{E}(f,f) = \int |\nabla f|^2 \, \mathrm{d}\mu$$

We prove that Log-Sobolev inequality is equivalent to the so-called contractivity property.

**2.13 Theorem.** For a symmetric semigroup with a generator L the following statements are equivalent

(i) (Log-Sobolev inequality) for every  $f: \Omega \to \mathbb{R}$  satisfying suitable technical assumptions

$$\mathbb{E}(f^2 \ln f^2) - (\mathbb{E}f^2) \ln(\mathbb{E}f^2) \le C\mathbb{E}(f(-L)f),$$

(ii) (hypercontractivity) for every p > q > 1 and  $f : \Omega \to \mathbb{R}$  we have

$$\left\|\mathcal{P}_t f\right\|_p \le \left\|f\right\|_q$$

for  $t \geq \frac{C}{4} \ln \frac{p-1}{q-1}$ .

*Proof.* Assume that we have (i). Take  $\phi_q : [q, \infty) \to \mathbb{R}$  given by

$$\phi_q(p) = \ln \left\| \mathcal{P}_{t(p)} f \right\|_p = \frac{1}{p} \ln \mathbb{E} \left| \mathcal{P}_{t(p)} f \right|^p$$

where  $t(p) = \frac{C}{4} \ln \frac{p-1}{q-1}$ . It suffices to show that  $\left\| \mathcal{P}_{t(p)} f \right\|_p \le \|f\|_q$ . Indeed, if t > t(p) then we obtain

$$\left\|\mathcal{P}_{t}f\right\|_{p} = \left\|\mathcal{P}_{t(p)+t-t(p)}f\right\|_{p} \le \left\|\mathcal{P}_{t-t(p)}f\right\|_{q} \le \left\|f\right\|_{q},$$

since  $\mathcal{P}_{t-t(p)}$  is a contraction in  $L^q$ .

To prove that  $\|\mathcal{P}_{t(p)}f\|_p \leq \|f\|_q$  we can assume that f i nonnegative. Indeed, the inequality  $-|f| \leq f \leq |f|$  implies (positivity preserving) that  $-\mathcal{P}_t|f| \leq \mathcal{P}_t f \leq \mathcal{P}_t|f|$ , hence  $|\mathcal{P}_t f| \leq \mathcal{P}_t|f|$ . Therefore  $\|\mathcal{P}_{t(p)}f\|_p \leq \|\mathcal{P}_{t(p)}|f|\|_p$ .

Take a nonnegative f. Since t(q) = 0, the inequality  $\|\mathcal{P}_{t(p)}f\|_p \leq \|f\|_q$  is equivalent to  $\phi_q(p) \leq \phi_q(q)$ . Hence, it suffices to show that the function  $[q, \infty) \ni p \mapsto \phi_q(p)$  is nonincreasing. Set  $\mathcal{P}_{t(p)}f = f_p$ . We have

$$\frac{\mathrm{d}}{\mathrm{d}p}\phi_q(p) = \frac{1}{p} \frac{\mathbb{E}\frac{\mathrm{d}}{\mathrm{d}p}(f_p^p)}{\mathbb{E}f_p^p} - \frac{1}{p^2}\ln\mathbb{E}f_p^p$$

and

$$\frac{\mathrm{d}}{\mathrm{d}p}f_p^p = \frac{\mathrm{d}}{\mathrm{d}p}\left(\mathcal{P}_{t(p)}f\right)^p = \frac{\mathrm{d}}{\mathrm{d}p}e^{p\ln(\mathcal{P}_{t(p)}f)} = e^{p\ln(\mathcal{P}_{t(p)}f)}\left(\ln(\mathcal{P}_{t(p)}f) + p\frac{L\mathcal{P}_{t(p)}f}{\mathcal{P}_{t(p)}f}\right) \cdot \frac{\mathrm{d}t(p)}{\mathrm{d}p}$$
$$= f_p^p\ln f_p + f_p^{p-1}p(Lf_p)\frac{C}{4}\ln\frac{1}{p-1}.$$

Thus,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}p}\phi_q(p) &= \frac{1}{p} \cdot \frac{\mathbb{E}f_p^p \ln f_p}{\mathbb{E}f_p^p} + \frac{C}{4} \frac{1}{p-1} \cdot \frac{\mathbb{E}f_p^{p-1}Lf_p}{\mathbb{E}f_p^p} - \frac{1}{p^2} \ln \mathbb{E}f_p^p \\ &= \frac{1}{p^2 \mathbb{E}f_p^p} \left( \left( \mathbb{E}f_p^p \ln(f_p^p) - (\mathbb{E}f_p^p) \ln(\mathbb{E}f_p^p) \right) + \frac{Cp}{4(p-1)} \mathbb{E}(f_p^{p-1}Lf_p) \right) \\ &= \frac{1}{p^2 \mathbb{E}f_p^p} \left( \mathrm{Ent}(f_p^p) + \frac{Cp}{4(p-1)} \mathbb{E}(f_p^{p-1}Lf_p) \right). \end{aligned}$$

We would like to prove

$$\operatorname{Ent}(f_p^p) \le \frac{Cp^2}{4(p-1)} \mathbb{E}(f_p^{p-1}(-L)f_p).$$

Taking  $f = f_p^{p/2}$  in the Log-Sobolev inequality and using Stroock-Varopoulos inequality we obtain

$$\operatorname{Ent}(f_p^p) \le C\mathbb{E}\left(f_p^{p/2}(-L)f_p^{p/2}\right) \le \frac{Cp^2}{4(p-1)}\mathbb{E}(f_p^{p-1}(-L)f_p)$$

To prove that *(ii)* implies *(i)* observe that for a nonnegative function f the inequality  $\left\| \mathcal{P}_{t(p)} f \right\|_p \leq \left\| f \right\|_q$  implies that  $\frac{\mathrm{d}}{\mathrm{d}p} \left\| \mathcal{P}_{t(p)} f \right\|_p \Big|_{p=q} \leq 0$ , which is equivalent to

$$\operatorname{Ent}(f^q) \le \frac{Cq^2}{4(q-1)} \mathbb{E}(f^{q-1}(-L)f).$$

Now it suffices to take q = 2 to obtain Log-Sobolev inequality for nonnegative functions. If f is not necessarily nonnegative then we have

$$\operatorname{Ent}(f^2) = \operatorname{Ent}(|f|^2) \le C\mathbb{E}|f|(-L)|f| \le C\mathbb{E}f(-L)f$$

because of the energy stability lemma.

Since the Ornstein-Uhlenbeck semigroup and the continuous time random walk on  $\Sigma_n$  satisfy Log-Sobolev inequality with constant 2, we have proved the following theorem.

**2.14 Theorem.** Let  $(\mathcal{P}_t)_{t\geq 0}$  be the Ornstein-Uhlenbeck semigroup or the continuous time random walk on  $\Sigma_n$ . Then for every p > q > 1 and  $t \geq \frac{1}{2} \ln \frac{p-1}{q-1}$  we have

$$\left\|\mathcal{P}_t f\right\|_p \le \left\|f\right\|_q.$$

As an application of the hypercontractivity we prove the following theorem.

**2.15 Theorem** (Khinchin-Kahane inequality). Let  $(F, \|\cdot\|)$  be a normed space and let  $v_1, \ldots, v_n \in F$ . Then for p > q > 1 we have

$$\left(\mathbb{E}\left\|\sum_{i=1}^{n}r_{i}v_{i}\right\|\right)^{1/p} \leq \sqrt{\frac{p-1}{q-1}}\left(\mathbb{E}\left\|\sum_{i=1}^{n}r_{i}v_{i}\right\|\right)^{1/q}.$$

*Proof.* Let  $H(x) = \|\sum_{i=1}^{n} x_i v_i\|, H : \Sigma_n \to [0, \infty)$ . We have proved that  $(-L)H \leq H$ . Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_t H = L\mathcal{P}_t H = -\mathcal{P}_t L H \ge -\mathcal{P}_t H.$$

Therefore  $\mathcal{P}_t H \ge e^{-t} \mathcal{P}_0 H = e^{-t} H$ . Take  $t = \frac{1}{2} \ln \frac{p-1}{q-1}$ . By the hypercontractivity of  $\mathcal{P}_t$  we obtain  $\sqrt{q-1}$ 

$$\sqrt{\frac{q-1}{p-1}} \|H\|_p = e^{-t} \|H\|_p \le \|\mathcal{P}_t H\|_p \le \|H\|_q.$$

# References

- [Ball] K. Ball, Ellipsoids of maximal volume in convex bodies, Geom. Dedicata 41 (2) (1992), 241–250.
- [Bec] W. Beckner, Inequalities in Fourier analysis, Ann. of Math. 102(1) (1975), 159– 182.
- [Bob1] S. Bobkov, Localization proof of the isoperimetric Bakry-Ledoux inequality and some applications, *Teor. Veroyatnost. i Primenen.* 47 (2002), no. 2, pp. 340–346. Translation in: *Theory Probab. Appl.* 47 (2003), no. 2, pp. 308–314.
- [Bob2] S. Bobkov, Remarks on the growth of  $L_p$ -norms of polynomials, *Geometric aspects of functional analysis*, pp. 27–35, Lecture Notes in Math., 1745, Springer, Berlin, 2000.
- [BH] S. Bobkov, C. Houdré, Isoperimetric constants for product probability measures, Ann. Probab. 25 (1997), no. 1, 184–205.
- [Bon] A. Bonami. Étude des coefficients de Fourier des fonctions de  $L_p(G)$ , Ann. Inst. Fourier **20** (1970), 335–402.
- [Bor1] Ch. Borell, The Brunn-Minkowski inequality in Gauss space, Invent. Math. 30 (1975), 207–216.
- [Bor2] Ch. Borell, The Ehrhard inequality, C. R. Math. Acad. Sci. Paris 337 (2003), 663–666.
- [Ehr] A. Ehrhard, Symetrisation dans l'espace de Gauss, Math. Scand. 53 (1983), 281– 301.
- [Gro] L. Gross, Logarithmic Sobolev inequalities, Amer. J. Math. 97 (1975), 1061–1083.
- [John] F. John, Extremum problems with inequalities as subsidiary conditions, pp. 187–204 in Studies and essays presented to R. Courant on his 60th birthday (Jan. 8, 1948), Interscience, New York, 1948.
- [Lat] R. Latała, A note on the Ehrhard inequality, Studia Math. 118 (1996), 169–174.
- [Led] M. Ledoux, The geometry of Markov diffusion generators, Ann. Fac. Sci. Toulouse IX (2000), 305-366.
- [Fr] M. Fradelizi, Concentration inequalities for s-concave measures of dilations of Borel sets and applications, *Electron. J. Probab.*, Vol. 14 (2009), 2068–2090.

- [Nel1] E. Nelson, A quartic interaction in two dimensions, In Mathematical Theory of Elementary Particles, MIT Press (1966), 69–73.
- [Nel2] E. Nelson, Construction of quantum fields from Markoff fields, J. Functional Analysis 12 (1973), 97–112.
- [Pi] G. Pisier, The volume of convex bodies and Banach space geometry, Cambridge University Press, 1999.
- [ST] V. Sudakov, B. Tsirelson, Extremal properties of half-spaces for spherically invariant measures (in Russian), Zap. Nauchn. Sem. L.O.M.I. 41 (1974), 14–24.