Notes on the course given by M. Fradelizi and O. Guédon about Convex Geometry

Tomasz Tkocz

March 13, 2012

Abstract

These notes contain the material which has seemed interesting and new to the author. They are not intended to provide a systematic course on the subject. Rather, the notes touch loosely connected and selected parts, which might hopefully entertain the reader.

The course was given at the Université Paris-Est, January - March 2012.

Part I
Matthieu's lectures

1 Lecture I & II — Gaussian Isoperimetry

The material here concerns the isoperimetry problem both for the Lebesgue measure and the Gaussian measure. We emphasize the analytic approach via functional inequalities.

1.1 L_1-Sobolev inequality via classical isoperimetry

We say that a function $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is *locally Lipschitz* if the quantity

$$|\nabla f(x)| = \overline{\lim}_{y\to x}\frac{|f(y)-f(x)|}{|y-x|} \tag{1.1}$$

is bounded on $\mathbb{R}^n$. Note that if f is a C^1 function then

$$|\nabla f(x)| = \left[\sum_{i=1}^{n}|f_{x_i}(x)|^2\right]^{1/2}.$$

1.1 Proposition (Co-area inequality)**.** *Let $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}_+$ be a locally Lipschitz function. Then*

$$\int_{\mathbb{R}^n} |\nabla f(x)|\, \mathrm{d}x \geq \int_{\mathbb{R}} |\partial\{f > t\}|\, \mathrm{d}t. \tag{1.2}$$

Proof (following [BH, Lemma 3.2]). We assume that f is bounded. For a point $x \in \mathbb{R}^n$ and a positive number h let us define

$$f_h(x) = \sup_{|y-x|<h} f(y).$$

Then, noticing that $\{f_h > t\} = \{f > t\}_h$, we have

$$\int_{\mathbb{R}^n} f_h = \int_0^\infty |\{f_h > t\}|\, \mathrm{d}t = \int_0^\infty |\{f > t\}_h|\, \mathrm{d}t.$$

Hence,

$$\int_{\mathbb{R}^n} \frac{f_h - f}{h} = \int_0^\infty \frac{|\{f > t\}_h| - |\{f > t\}|}{h}\, \mathrm{d}t.$$

But,

$$\overline{\lim}_{h\to 0} \frac{f_h(x) - f(x)}{h} \leq \overline{\lim}_{y\to x} \frac{f(y) - f(x)}{|y-x|} \leq |\nabla f(x)|,$$

so

$$\begin{aligned}
\int_{\mathbb{R}^n} |\nabla f| &\geq \int_{\mathbb{R}^n} \overline{\lim}_{h\to 0} \frac{f_h(x) - f(x)}{h}\, \mathrm{d}x \\
&\geq \overline{\lim}_{h\to 0} \int_{\mathbb{R}^n} \frac{f_h(x) - f(x)}{h}\, \mathrm{d}x \\
&\geq \underline{\lim}_{h\to 0} \int_0^\infty \frac{|\{f > t\}_h| - |\{f > t\}|}{h}\, \mathrm{d}t \\
&\geq \int_0^\infty |\partial\{f > t\}|\, \mathrm{d}t,
\end{aligned}$$

where Fatou's lemma is used twice. □

1.2 Theorem. *Let $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a compactly supported locally Lipschitz function for which $|\nabla f| \in L_1(\mathbb{R}^n)$. Then*

$$\|f\|_{\frac{n}{n-1}} \leq \frac{1}{n|B_2^n|^{1/n}} \int_{\mathbb{R}^n} |\nabla f|. \tag{1.3}$$

Proof. We restrict ourselves to the case of nonnegative functions. Then by Proposition 1.1 and the classical isoperimetry inequality we get the estimations

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

(the assumption guaranties that the sets $\{f > t\}$ are of finite measure so we are enable to use the isoperimetry). A game with indicator function and an application of the continuous version of the triangle inequality enable us to finish the proof, for

$$\begin{aligned}\int_0^\infty |\{f > t\}|^{\frac{n-1}{n}}\,\mathrm{d}t &= \int_0^\infty \left(\int_{\mathbb{R}^n} (\mathbf{1}_{\{f>t\}})^{\frac{n}{n-1}}\right)^{\frac{n-1}{n}} \mathrm{d}t \\ &= \int_0^\infty \|\mathbf{1}_{\{f>t\}}\|_{\frac{n}{n-1}}\,\mathrm{d}t \geq \left\| \int_0^\infty \mathbf{1}_{\{f>t\}}\,\mathrm{d}t \right\| \\ &= \|f\|_{\frac{n}{n-1}}.\end{aligned}$$

□

1.3 Exercise. Show that Theorem 1.2 implies the classical isoperimetry inequality, i.e. for a measurable set A in $\mathbb{R}^n$ we have

$$|\partial A| \geq n|B_2^n|^{\frac{1}{n}}|A|^{\frac{n-1}{n}}.$$

1.2 Isoperimetry in the Gaussian space

By the n-dimensional Gaussian space we mean $\mathbb{R}^n$ equipped with the standard Gaussian measure γ. We begin with stating the analogue of Brunn-Minkowski inequality.

1.4 Theorem. *Let A and B be nonempty compact sets in $\mathbb{R}^n$ and let numbers α, β be such that $\alpha + \beta \geq 1$ and $|\alpha - \beta| \leq 1$. Then*

$$\Phi^{-1}\gamma(\alpha A + \beta B) \geq \alpha\Phi^{-1}\gamma(A) + \beta\Phi^{-1}\gamma(B). \tag{1.4}$$

Historically, it was A. Ehrhard who proved this inequality in 1983 (see [Ehr]) yet for convex sets. Then it was refined by R. Latała in 1996 to the case when A or B is convex (see [Lat]). Finally, in 2003 C. Borell (see [Bor2]) gave the proof of the above general statement. However, we will occupy ourselves with functional versions.

1.5 Theorem. *Fix numbers α, β such that $\alpha+\beta \geq 1$ and $|\alpha-\beta| \leq 1$. Let Borel functions $f, g, h \colon \mathbb{R}^n \longrightarrow [0,1]$ satisfy for any $x, y \in \mathbb{R}^n$*

$$\Phi^{-1}\big(h(\alpha x+\beta y\big) \geq \alpha\Phi^{-1}\big(f(x)\big)+\beta\Phi^{-1}\big(g(x)\big). \tag{1.5}$$

Then

$$\Phi^{-1}\left(\int h\,\mathrm{d}\gamma\right) \geq \alpha\Phi^{-1}\left(\int f\,\mathrm{d}\gamma\right)+\beta\Phi^{-1}\left(\int g\,\mathrm{d}\gamma\right) \tag{1.6}$$

1.6 Exercise. Prove that Theorem 1.5 implies Theorem 1.4.

1.7 Theorem. *Let $f, g, h \colon \mathbb{R}^n \longrightarrow]0,1[$ be Borel functions. If for any $x, y \in \mathbb{R}^n$ inequality (1.5) is satisfied then for any $t \geq 0$ and any $x, y \in \mathbb{R}^n$ we have*

$$\Phi^{-1}\Big(Q_t h(\alpha x+\beta y)\Big) \geq \alpha\Phi^{-1}\Big(Q_t f(x)\Big)+\beta\Phi^{-1}\Big(Q_t g(y)\Big). \tag{1.7}$$

Here, the notion of the *heat semi-group* has been used. For an integrable function $f \colon \mathbb{R}^n \longrightarrow \mathbb{R}$ and a nonnegative number t we define the *heat operator*

$$Q_t f(x) = \int_{\mathbb{R}^n} f(x+\sqrt{t}z)\,\mathrm{d}\gamma(z). \tag{1.8}$$

Note the following basic properties.

1.8 Proposition. *The heat semi-group operator Q_t satisfies*

1) $Q_t(f+g) = Q_t(f)+Q_t(g)$,

2) $Q_{s+t} = Q_s \circ Q_t$,

3) $Q_t f \geq 0$ *if* $f \geq 0$,

4) $Q_t 1 = 1$,

5) $Q_0 f = f$,

6) $\partial_t Q_t f = \frac{1}{2}\Delta(Q_t f)$,

7) $\int g(x)(Q_t f)(x)\,\mathrm{d}x = \int f(x)(Q_t g)(x)\,\mathrm{d}x$,

8) $(Q_t f)(x) \approx \frac{1}{(2\pi t)^{n/2}}\int f(x)\,\mathrm{d}x$ *when* $t \to \infty$,

9) $Q_1 f(0) = \int f\,\mathrm{d}\gamma$.

Our goal is to prove Theorem 1.7. We begin with a lemma concerning matrices.

1.9 Lemma. *Let $A, B \in M_{n\times n}(\mathbb{R})$ be two matrices of size $n \times n$ with real entries. Assume they are symmetric and positive. Then the matrix*

$$A * B = [a_{ij}b_{ij}]_{i,j}$$

is also symmetric and positive.

Proof. Take a vector $v \in \mathbb{R}^n$. We may write

$$v^T(A * B)v = \sum_{i,j} v_i a_{ij} b_{ij} v_j.$$

Let $C = [c_{ij}]$ be a symmetric and positive matrix such that $B = C^2$ (it exists as B is symmetric and positive). Then by symmetry of C

$$v^T(A * B)v = \sum_{i,j} v_i a_{ij} \sum_k c_{jk} c_{ki} v_j = \sum_{i,j,k} c_{ki}(v_i a_{ij} v_j) c_{jk} = \mathrm{tr}(CDADC),$$

where $D = \mathrm{diag}(v_1, \ldots, v_n)$. Since A is positive, DAD is also positive, and, consequently, $CDADC$ is positive, so the trace of this matrix is positive. $\square$

Proof of Theorem 1.7. The proof is quite involved and will be divided into several steps. Let us discuss the strategy first. We start with proving the assertion in the class of smooth functions satisfying certain regularity conditions. Out of this we will be able to deduce Theorem 1.4. Having this at hand, we will show how to derive Theorem 1.7 in its full generality.

Step I (Theorem 1.7 for *nice* functions). Given parameters $a > 0$ and $0 < \epsilon < \rho < 1$ we define

$$\delta_{\epsilon,\rho} := \max\left(\alpha\Phi^{-1}(2\epsilon) + \beta\Phi^{-1}(\rho), \alpha\Phi^{-1}(\rho) + \beta\Phi^{-1}(2\epsilon)\right) \tag{1.9}$$

and the set of triples of *nice* functions

$$\begin{aligned} \mathcal{N}^n_{a,\epsilon,\rho} = \{(f,g,h),\ & f,g,h\colon \mathbb{R}^n \longrightarrow]0,1[\text{ are of } C^\infty \text{ class}, \\ & f,g \le \epsilon \text{ outside } [-a,a]^n, \\ & f,g \le \rho \text{ everywhere}, \\ & h \ge \delta_{\epsilon,\rho} \text{ everywhere}\} \end{aligned} \tag{1.10}$$

We are to show that for any n if functions (f,g,h) belong to the class $\mathcal{N}^n_{a,\epsilon,\rho}$ for some parameters a, ϵ, ρ and satisfy (1.5), then (1.7) holds. We proceed by induction on n. Let $n=1$. Let us define the functions

$$\begin{aligned} F_t(x) &= \Phi^{-1}\big(Q_t f(x)\big), \\ G_t(x) &= \Phi^{-1}\big(Q_t g(x)\big), \\ H_t(x) &= \Phi^{-1}\big(Q_t h(x)\big), \\ C(t,x,y) &= H_t(\alpha x+\beta y) - \alpha F_t(x) - \beta G_t(y). \end{aligned}$$

We want to establish that $C(t,x,y) \geq 0$ for $t \geq 0$ and $x, y \in \mathbb{R}$. The idea is to use the maximum principle on the set $[0,T] \times \mathbb{R}^2$. Suppose that, on the contrary, there is a point (t,x,y) in $[0,T]\times\mathbb{R}^2$ such that $C(t,x,y) < 0$. Recall that by the hypothesis $C(0,x,y) \geq 0$. Let r be such that $\gamma([-r,r]) = 1-\epsilon$. Take $b = a + r\sqrt{T}$. Then for any x such that $|x| > b$ and $t \in [0,T]$ the fact that f is *nice* implies

$$\begin{aligned} Q_t f(x) &= \int_{|z|\leq r} f(x+\sqrt{t}z)\,\mathrm{d}\gamma(z) + \int_{|z|>r} f(x+\sqrt{t}z)\,\mathrm{d}\gamma(z) \\ &\leq \epsilon\gamma([-r,r]) + \gamma([-r,r]^c) \leq 2\epsilon, \end{aligned}$$

and the same for g. Since $h \geq \delta$, we have also $Q_t h \geq \delta$, so

$$H_t(\alpha x+\beta y) \geq \Phi^{-1}(\delta) \geq \alpha F_t(x) + \beta G_t(y),$$

if $|x|$ or $|y| > b$. Consequently, $C(t,x,y) \geq 0$ on $[0,T] \times ([-b,b]^2)^c$. Thus

$$\inf_{[0,T]\times\mathbb{R}^2} C(t,x,y) = \inf_{[0,T]\times[-b,b]^2} C(t,x,y).$$

Consider the function $C_\epsilon(t,x,y) = C(t,x,y) + \eta t$ and choose η so small that also $C_\eta \geq 0$ does not hold everywhere and

$$\inf_{[0,T]\times\mathbb{R}^2} C_\eta(t,x,y) = \inf_{[0,T]\times[-b,b]^2} C_\eta(t,x,y).$$

Moreover, this infimum is attained, say at (t_0,x_0,y_0). Then

$$\begin{aligned} 0 &= \partial_x C_\eta(t_0,x_0,y_0) = \partial_x C(t_0,x_0,y_0), \\ 0 &= \partial_y C_\eta(t_0,x_0,y_0) = \partial_y C(t_0,x_0,y_0), \\ 0 &\geq \partial_t C_\eta(t_0,x_0,y_0) = \partial_t C(t_0,x_0,y_0) + \eta, \\ 0 &\leq \mathrm{Hess}_{(x,y)} C(t_0,x_0,y_0). \end{aligned} \tag{1.11}$$

Before we calculate the derivatives of C one piece of notation

$$\begin{aligned} \Phi(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\,\mathrm{d}t, \\ \Phi'(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} =: \varphi(x), \\ \Phi''(x) &= \varphi'(x) = -x\varphi(x). \end{aligned} \tag{1.12}$$

The crucial observation is the following (actually here we use smoothness, cf. Proposition 1.8, 6))

$$u_t \text{ satisfies } \partial_t u = \frac{1}{2}\Delta u \text{ means that } u_t = Q_t u_0.$$

For $U(t,\cdot) = \Phi^{-1} \circ u(t,\cdot)$ we get

$$\begin{aligned}
\partial_t u &= \partial_t \Phi(U) = \partial_t U \cdot \varphi(U),\\
\partial_x u &= \partial_x U \cdot \varphi(U),\\
\partial_{xx} u &= \partial_{xx} U \cdot \varphi(U) + (\partial_x U)^2 \varphi'(U) = \varphi(U)(\partial_{xx} U - U(\partial_x U)^2).
\end{aligned}$$

Thus $2\partial_t u = \Delta u = \partial_{xx} u$ yields

$$2\partial_t U = \partial_{xx} U - U(\partial_x U)^2.$$

Since F_t, G_t, H_t are the images with respect to Φ^{-1} of the heat operator, they satisfy above equation. Therefore (to shorten the notation we will write $H_t = H_t(\alpha x + \beta y)$, $F_t = F_t(x)$, $G_t = G_t(y)$)

$$\begin{aligned}
2\partial_t C &= 2\partial_t H_t - \alpha \cdot 2\partial_t F_t - \beta \cdot 2\partial_t G_t\\
&= H_t'' - H_t H_t'^2 - \alpha(F_t'' - F_t F_t'^2) - \beta(G_t'' - G_t G_t'^2)\\
&= H_t'' - \alpha F_t'' - \beta G_t'' - H_t H_t'^2 + \alpha F_t F_t'^2 + \beta G_t G_t'^2,\\
\partial_x C &= \alpha(H_t' - F_t'),\\
\partial_y C &= \beta(H_t' - G_t'),\\
\partial_{xx} C &= \alpha^2 H_t'' - \alpha F_t'',\\
\partial_{yy} C &= \beta^2 H_t'' - \beta G_t'',\\
\partial_{xy} C &= \alpha\beta H_t''.
\end{aligned}$$

We have done all the necessary calculations. Now we show how to obtain the desired contradiction. Combining the first two conditions of (1.11) with the above computation of $\partial_x C, \partial_y C$ we obtain $H_{t_0}' = F_{t_0}' = G_{t_0}'$. Whence $C(t_0, x_0, y_0) < 0$ yields

$$\begin{aligned}
-2\eta \geq 2\partial_t C(t_0, x_0, y_0) &= H_{t_0}'' - \alpha F_{t_0}'' - \beta G_{t_0}'' - F_{t_0}'^2 C(t_0, x_0, y_0)\\
&\geq H_{t_0}'' - \alpha F_{t_0}'' - \beta G_{t_0}''.
\end{aligned}$$

Yet

$$\begin{aligned}
H_t'' - \alpha F_t'' - \beta G_t'' &= H_t'' + \partial_{xx} C_t + \partial_{yy} C_t - \alpha^2 H_t'' - \beta^2 H_t''\\
&= \partial_{xx} C_t + \partial_{yy} C_t + \underbrace{\frac{1 - \alpha^2 - \beta^2}{\alpha\beta}}_{2c} \partial_{xy} C_t
\end{aligned}$$

is an elliptic operator, for the matrix $A = \left[\begin{smallmatrix} 1 & c \\ c & 1 \end{smallmatrix}\right]$ is positive because $|\alpha - \beta| \leq 1, \alpha + \beta \geq 1$. This, the fact that the matrix $\mathrm{Hess}_{x,y}C$ is positive at the point (t_0, x_0, y_0), and Lemma 1.9 imply that the matrix $E = A * \mathrm{Hess}_{x,y}C(t_0, x_0, y_0) = \left[\begin{smallmatrix} \partial_{xx}C(t_0,x_0,y_0) & c\partial_{xy}C(t_0,x_0,y_0) \\ c\partial_{xy}C(t_0,x_0,y_0) & \partial_{yy}C(t_0,x_0,y_0) \end{smallmatrix}\right]$ is also positive and as a consequence

$$-2\eta \geq H''_{t_0} - \alpha F''_{t_0} - \beta G''_{t_0} = [1\ 1]E \begin{bmatrix} 1 \\ 1 \end{bmatrix} \geq 0,$$

a contradiction.

Thanks to certain structure of inequality (1.7), the induction step is relatively easy (this is the same as for Prekopa-Leindler inequality). For ease of notation, we prove the theorem for $n = 2$. Let functions $f, g, h \colon \mathbb{R}^2 \longrightarrow]0,1[$ be nice (with parameters a, ϵ, ρ) and satisfy for $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$

$$\Phi^{-1}\big(h(\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2)\big) \geq \alpha\Phi^{-1}\big(f(x_1, x_2)\big) + \beta\Phi^{-1}\big(g(y_1, y_2)\big).$$

We fix x_2, y_2, note that functions $f(\cdot, x_2), g(\cdot, y_2), h(\cdot, \alpha x_2 + \beta y_2)$ are also nice and apply the one dimensional result at the points x_1, y_1

$$\begin{aligned} \Phi^{-1}\Big(Q_t^{(1)} h(\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2)\Big) \geq & \alpha\Phi^{-1}\Big(Q_t^{(1)} f(x_1, x_2)\Big) \\ & + \beta\Phi^{-1}\Big(Q_t^{(1)} g(y_1, y_2)\Big), \end{aligned}$$

where $Q_t^{(1)} f(x_1, x_2) = \int_{\mathbb{R}} f(x_1 + \sqrt{t} z_1, x_2) \,\mathrm{d}\gamma(z_1)$. Now we fix x_1, y_1 and use the induction hypothesis for the nice functions $Q_t^{(1)} f(x_1, \cdot)$, $Q_t^{(1)} g(y_1, \cdot)$, $Q_t^{(1)} h(\alpha x_1 + \beta y_1, \cdot)$ and the points x_2, y_2, getting

$$\Phi^{-1}\Big(Q_t^{(2)} Q_t^{(1)} h(\alpha x + \beta y)\Big) \geq \alpha\Phi^{-1}\Big(Q_t^{(2)} Q_t^{(1)} f(x)\Big) + \beta\Phi^{-1}\Big(Q_t^{(2)} Q_t^{(1)} g(y)\Big).$$

But $Q_t = Q_t^{(2)} Q_t^{(1)}$, so the proof is complete.

Step II (Step I $\Longrightarrow$ Theorem 1.4). Let $A, B \subset \mathbb{R}^n$ be compact. Fix $0 < 2\epsilon < \rho < 1$ and $\eta > 0$. There is a smooth function $f \colon \mathbb{R}^n \longrightarrow]0,1[$ such that $f = \rho$ on A and $f = \epsilon$ outside A_η. Similarly for B, there is certain smooth function g. Let $h \colon \mathbb{R}^n \longrightarrow]0,1[$ be a smooth function such that $h = \Phi\big((\alpha + \beta)\Phi^{-1}(\rho)\big)$ on $\alpha A_\eta + \beta B_\eta$ and $h = \delta_{\epsilon,\rho}$ outside $\big(\alpha A_\eta + \beta B_\eta\big)_\eta$. Note that $(f, g, h) \in \mathcal{N}^n_{a,\epsilon,\rho}$ for some $a > 0$ as A and B are compact and $h \geq \delta_{\epsilon,\rho}$ because $2\epsilon < \rho$. One verifies that (1.4) holds for any $x, y \in \mathbb{R}^n$. Indeed, if $x \in A_\eta$ and $y \in B_\eta$ we have the equality $\alpha\Phi^{-1} f(x) + \beta\Phi^{-1} g(y) =$

$(\alpha+\beta)\Phi^{-1}(\rho) = \Phi^{-1}h(\alpha x+\beta y)$. Otherwise, we use the estimation $h \geq \delta_{\epsilon,\rho}$. Therefore we might use the result of Step I and get by virtue of (1.7) for $t = 1$, $x = y = 0$

$$\begin{aligned}\Phi^{-1}\left(\int h\,\mathrm{d}\gamma\right) &\geq \alpha\Phi^{-1}\left(\int f\,\mathrm{d}\gamma\right) + \beta\Phi^{-1}\left(\int g\,\mathrm{d}\gamma\right)\\ &\geq \alpha\Phi^{-1}(\rho\gamma(A)) + \beta\Phi^{-1}(\rho\gamma(B)).\end{aligned}$$

Yet $\delta_{\epsilon,\rho} \xrightarrow[\epsilon\to 0]{} 0$, whence

$$\Phi^{-1}\left(\int h\,\mathrm{d}\gamma\right) \xrightarrow[\eta\to 0,\epsilon\to 0]{} \Phi^{-1}\Big(\Phi\big((\alpha+\beta)\Phi^{-1}(\rho)\big)\gamma(\alpha A+\beta B)\Big).$$

Taking the limit when $\rho \to 1$ we derive (1.4).

Step III (Theorem 1.4 $\Longrightarrow$ Theorem 1.7). To see that Theorem 1.4 in dimension $n+1$ implies Theorem 1.7 in dimension n, for a function $f\colon \mathbb{R}^n \longrightarrow]0,1[$ and a point $x \in \mathbb{R}^n$ consider the set $B_f^x = \{(s,z) \in \mathbb{R}\times\mathbb{R}^n \mid s \leq \Phi^{-1}\big(f(x+\sqrt{t}z)\big)\}$. Since $\gamma(B_f^x) = Q_t f(x)$ it is enough to check that $\alpha B_f^x + \beta B_g^y \subset B_h^{\alpha x+\beta y}$ and conclude Theorem 1.7 from inequality 1.4.

$\square$

Now we would like to infer the Gaussian isoperimetry. We need two simple corollaries and an analytic lemma.

1.10 Corollary (The Sudakov-Tsierlson inequality)**.** *Let $K \subset \mathbb{R}^n$ be a convex set and $t \geq 1$. Then*

$$\Phi^{-1}\gamma(tK) \geq t\Phi^{-1}\gamma(K). \tag{1.13}$$

Proof. Apply Theorem 1.7 for $A = B = K$, $\alpha = \beta = t/2$ (convexity of K gives $K = \frac{K+K}{2}$). $\square$

1.11 Corollary (The Ehrhard inequality for convex/nonconvex)**.** *Let $A \subset \mathbb{R}^n$ be a Borell set, $K \subset \mathbb{R}^n$ be a convex set and $\alpha, \beta > 0$ such that $\alpha+\beta \geq 1$, $\alpha - \beta \leq 1$. Then*

$$\Phi^{-1}\gamma(\alpha A + \beta K) \geq \alpha\Phi^{-1}\gamma(A) + \beta\Phi^{-1}\gamma(K). \tag{1.14}$$

Proof. If $\beta - \alpha \leq 1$ the assumptions of Theorem 1.7 are satisfied. Otherwise we use this theorem for the weights $\alpha, 1+\alpha$ and then Corollary 1.10 for $\frac{\beta}{1+\alpha} > 1$

$$\begin{aligned}\Phi^{-1}\gamma(\alpha A + \beta K) &= \Phi^{-1}\gamma\left(\alpha A + (1+\alpha)\frac{\beta}{1+\alpha}K\right)\\ &\geq \alpha\Phi^{-1}\gamma(A) + (1+\alpha)\Phi^{-1}\gamma\left(\frac{\beta}{1+\alpha}K\right)\\ &\geq \alpha\Phi^{-1}\gamma(A) + \beta\Phi^{-1}\gamma(K).\end{aligned}$$

□

1.12 Lemma. *We have* $\lim_{r\to\infty}\frac{1}{r}\Phi^{-1}\gamma(rB_2^n) = 1$ *and as a consequence* $\sup_r \frac{1}{r}\Phi^{-1}\gamma(rB_2^n) = 1$.

Proof. Integrating by parts we obtain

$$\left(\frac{1}{r} - \frac{1}{r^3}\right)e^{-r^2/2} \leq \int_r^\infty e^{-t^2/2}\,\mathrm{d}t \leq \frac{1}{r}e^{-r^2/2}. \tag{1.15}$$

We will adopt throughout these notes the standard notation that $f(x) \sim_{x\to\infty} g(x)$ means $\lim_{x\to\infty}\frac{f}{g} = 1$. Thus $\ln\big(1-\Phi(r)\big) \sim_{r\to\infty} -\frac{r^2}{2}$. Taking $x = \Phi(r) \xrightarrow{r\to\infty} 1$ yields

$$\Phi^{-1}(x) \sim_{x\to 1} \sqrt{-2\ln(1-x)}. \tag{1.16}$$

Yet

$$\begin{aligned}1 - \gamma(rB_2^n) = \int_{|x|>r} e^{-|x|^2/2}\frac{\mathrm{d}x}{\sqrt{2\pi}^n} &= \int_r^\infty t^{n-1}e^{-t^2/2}\frac{n|B_2^n|}{\sqrt{2\pi}^n}\\ &\sim_{r\to\infty} r^{n-2}e^{-r^2/2}\frac{n|B_2^n|}{\sqrt{2\pi}^n}.\end{aligned}$$

Therefore

$$\Phi^{-1}\gamma(rB_2^n) \sim_{r\to\infty} \sqrt{-2\ln(1-\Phi^{-1}\gamma(rB_2^n))} \sim_{r\to\infty} r.$$

Since $rB_2^n \subset]-\infty, r] \times \mathbb{R}^{n-1}$, we have $\Phi^{-1}\gamma(rB_2^n) \leq \Phi^{-1}\Phi(r) = r$, hence

$$\sup_r \frac{1}{r}\Phi^{-1}\gamma(rB_2^n) = 1.$$

□

Now we are able to give an elegant proof of the Gaussian ispoterimetry inequality due to Ch. Borell [Bor1] and, independently, V. Sudakov and B. Tsirelson [ST].

1.13 Theorem (Gaussian isoperimetry)**.** *Let A be a Borel set in $\mathbb{R}^n$ and H a half-space such that $\gamma(A) = \gamma(H)$. Then*

$$\gamma(A_\epsilon) \geq \gamma(H_\epsilon). \tag{1.17}$$

Consequently,

$$\gamma^+(\partial A) \geq \gamma^+(\partial H), \tag{1.18}$$

or, in other words,

$$\gamma^+(\partial A) \geq I(\gamma(A)), \tag{1.19}$$

where the Gaussian isoperimetric profile $I\colon [0,1] \longrightarrow \mathbb{R}_+$ *is given by (see* (1.12)*)*

$$I = \varphi \circ \Phi^{-1}. \tag{1.20}$$

Proof of the theorem. Set $\epsilon > 0$. Applying Corollary 1.11 for A and the ball rB_2^n we have

$$\Phi^{-1}\gamma\left(A + \frac{\epsilon}{r} rB_2^n\right) \geq \Phi^{-1}\gamma(A) + \frac{\epsilon}{r}\Phi^{-1}\gamma(rB_2^n).$$

With the aid of Lemma 1.12 we optimize the right hand side with respect to r and get

$$\Phi^{-1}\gamma(A_\epsilon) \geq \Phi^{-1}\gamma(A) + \epsilon,$$

which is (1.17). $\square$

2 Lecture III & IV — Gaussian functional inequalities and the hypercontractivity

We continue studying functional inequalities in the Gaussian space. Namely, first we show the functional version of the Gaussian isoperimetric inequality, i.e. the L_1-Sobolev inequality (cf. (1.3)). Then we derive the so-called *Log*-Sobolev inequality. At the end we also touch the topic of hypercontractivity.

2.1 L_1 and Log-Sobolev inequalities for the Gaussian measure

Recall the classical isoperimetric inequality altogether with the co-area inequality have allowed us to obtain the L_1-Sobolev inequality (see Theorem 1.2). The same happens for the Gaussian measure.

2.1 Theorem (Bobkov's inequality)**.** *Let $f\colon \mathbb{R}^n \longrightarrow [0,1]$ be a locally Lipschitz function. Then*

$$I\left(\int_{\mathbb{R}^n} f \,\mathrm{d}\gamma\right) \le \int_{\mathbb{R}^n} \sqrt{I\big(f(x)\big)^2 + |\nabla f(x)|^2}\,\mathrm{d}\gamma(x). \tag{2.1}$$

Proof. Given a function function f consider the subgraph of the function $\Phi^{-1}\circ f$

$$E(f) = \{(x,t)\in\mathbb{R}^n\times\mathbb{R},\ t\le \Phi^{-1}\big(f(x)\big)\}.$$

Note that

$$\gamma\big(E(f)\big) = \int^{\mathbb{R}^n}\int_{-\infty}^{\Phi^{-1}\left(f(x)\right)} \mathrm{d}\gamma(t)\,\mathrm{d}\gamma(x) = \int_{\mathbb{R}^n} f(x)\,\mathrm{d}\gamma(x).$$

The idea is to apply the isoperimetric inequality (1.19) for the set $A = E(f)$. First we would like to see what is A_ϵ. Denote $g = \Phi^{-1}\circ f$ and define the function

$$D_\epsilon g(x) = \sup\left\{\frac{|g(x)-g(y)|}{|x-y|},\ 0<|x-y|<\epsilon\right\}.$$

Take a point $(x,t)\in A_\epsilon$. It means there is a point $(y,u)\in A$ such that $|x-y|^2+|t-u|^2\le\epsilon^2$. Then

$$\begin{aligned} t &\le u + \sqrt{\epsilon^2-|x-y|^2} \le g(y) + \sqrt{\epsilon^2-|x-y|^2} \\ &\le g(x) + D_\epsilon(x)|x-y| + \sqrt{\epsilon^2-|x-y|^2} \le g(x) + \epsilon\sqrt{\big(D_\epsilon(x)\big)^2+1}, \end{aligned}$$

where in the second estimate we use definition of $D_\epsilon g$ and the last one follows from Cauchy-Schwarz inequality $ab+c\le\sqrt{a^2+1}\sqrt{b^2+c^2}$. In fact, we have checked that $(x,t)\in E(\Phi\circ h)$, where $h(x) = g(x)+\epsilon\sqrt{\big(D_\epsilon(x)\big)^2+1}$, which means that $A_\epsilon\subset E(\Phi\circ h)$. Thus

$$\begin{aligned} \frac{\gamma(A_\epsilon)-\gamma(A)}{\epsilon} &\le \frac{1}{\epsilon}\int_{\mathbb{R}^n}\big(\Phi\circ h - f\big)\,\mathrm{d}\gamma \\ &= \int_{\mathbb{R}^n}\frac{\Phi\big(g+\epsilon\sqrt{1+(D_\epsilon g)^2}\big)-\Phi\circ g}{\epsilon}\,\mathrm{d}\gamma. \end{aligned}$$

Letting $\epsilon \to 0$ we get

$$\gamma^+(\partial A) \leq \int_{\mathbb{R}^n} (\varphi \circ g)\sqrt{1+|\nabla g|^2}\,\mathrm{d}\gamma = \int_{\mathbb{R}^n} \sqrt{I(f)^2 + |\nabla f|^2}\,\mathrm{d}\gamma,$$

as $\nabla g = \nabla(\Phi^{-1} \circ f) = \frac{\nabla f}{\varphi(\Phi^{-1} \circ f)} = \frac{\nabla f}{I(f)}$. □

2.2 Exercise. Prove that the Gaussian isoperimetric inequality follows from the Bobkov's inequality (2.1).

To state the *Log*-Sobolev inequality we need the notion of entropy. For a probability measure μ and a nonnegative function f we define its *entropy* with respect to the measure μ as

$$\mathrm{Ent}_\mu f = \int f \ln f \,\mathrm{d}\mu - \Big(\int f\,\mathrm{d}\mu\Big)\ln\Big(\int f\,\mathrm{d}\mu\Big). \tag{2.2}$$

2.3 Theorem (The *Log*-Sobolev inequality for the Gaussian measure)**.** *Let $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function of class C^1. Then*

$$\mathrm{Ent}_\gamma(g^2) \leq 2\int_{\mathbb{R}^n} |\nabla g|^2\,\mathrm{d}\gamma. \tag{2.3}$$

It is said that this can be deduced from the Bobkov's inequality. The argument is due to W. Beckner (see [Led, p. 331]) and hinges on putting $f = \epsilon g^2$ in (2.1) with $\epsilon \to 0$. We will give another proof in the next subsection which uses semi-group tools. However now let us follow Beckner. Before we proceed we need to know how the functions Φ^{-1} and I behave near zero.

2.4 Lemma. *We have*

$$\lim_{x\to 0} \frac{|\Phi^{-1}(x)|^2}{2\ln\frac{1}{x}} = 1. \tag{2.4}$$

More precisely,

$$|\Phi^{-1}(x)|^2 - 2\ln\frac{1}{x} = \ln\ln\frac{1}{x} + s(x), \qquad s(x) \xrightarrow[x\to 0]{} -\ln\pi. \tag{2.5}$$

Proof. The first formula easily follows from (1.16).

Now we deal with the second one. Putting $-r = \Phi^{-1}(x)$ $(x \approx 0)$ into (1.15) and taking logarithm we obtain

$$\begin{aligned} -\frac{|\Phi^{-1}(x)|^2}{2} - \ln|\Phi^{-1}(x)| + \ln\left(1 - \frac{1}{|\Phi^{-1}(x)|^2}\right) \\ \leq \ln\sqrt{2\pi} + \ln x \\ \leq -\frac{|\Phi^{-1}(x)|^2}{2} - \ln|\Phi^{-1}(x)|. \end{aligned}$$

Rearranging yields

$$
\begin{aligned}
-\ln(2\pi) + \ln|\Phi^{-1}(x)|^2 + \ln\left(1 - \frac{1}{|\Phi^{-1}(x)|^2}\right) \\
\le |\Phi^{-1}(x)|^2 - 2\ln\frac{1}{x} \\
\le -\ln(2\pi) + \ln|\Phi^{-1}(x)|^2.
\end{aligned}
$$

Let us rewrite for instance the estimate from above using (2.4)

$$
-\ln(2\pi) + \ln\frac{|\Phi^{-1}(x)|^2}{2\ln\frac{1}{x}} + \ln\left(2\ln\frac{1}{x}\right) = \ln\ln\frac{1}{x} - \ln\pi + \ln\frac{|\Phi^{-1}(x)|}{2\ln\frac{1}{x}}.
$$

We do the same for the estimate from below and as a result we obtain (2.5). $\square$

2.5 Corollary. *We have*

$$
I(x)^2 = x^2\left(2\ln\frac{1}{x} + \ln\ln\frac{1}{x} + r(x)\right), \qquad r(x) \xrightarrow[x\to 0]{} -\ln\pi. \tag{2.6}
$$

Proof. Recall (1.15) and put it slightly different

$$
\frac{1}{r}\varphi(r)\left(1 - \frac{1}{r^2}\right) \le \Phi(-r) \le \frac{1}{r}\varphi(r).
$$

Hence, putting $-r = \Phi^{-1}(x)$, we find that

$$
|\Phi^{-1}(x)| \le \frac{I(x)}{x} \le |\Phi^{-1}(x)|\frac{1}{1 - \frac{1}{|\Phi^{-1}(x)|^2}}.
$$

Combining this with (2.5) we conclude the assertion of the Corollary. $\square$

Now we go back and prove the *Log*-Sobolev inequality exploiting the Bobkov's inequality. Fix a C^1 function g. Without loss of generality we may assume that $1/N \le g \le N$ for some natural number N (one needs to properly approximate the function $g\mathbf{1}_{\{1/N \le g \le N\}}$). By homogeneity we might also consider only the case when $\int g^2 \,\mathrm{d}\gamma = 1$. Taking $f = \epsilon g^2$ in (2.1) yields

$$
I(\epsilon) \le \int \sqrt{I(\epsilon g^2)^2 + 4\epsilon^2 g^2 |\nabla g|^2}\,\mathrm{d}\gamma.
$$

Let us multiply both sides by $\frac{1}{\epsilon}\sqrt{2\ln\frac{1}{\epsilon}}$ and see what happens. Define the function $2R(x) = \ln\ln\frac{1}{x} + r(x)$, where $r(x)$ comes from Corollary 2.5. Then $I(x)^2 = x^2 \cdot 2\left(\ln\frac{1}{x} + R(x)\right)$ and the left hand side gives

$$\sqrt{2\ln\frac{1}{\epsilon} + 2R(\epsilon)}\sqrt{2\ln\frac{1}{\epsilon}} = 2\ln\frac{1}{\epsilon}\left(1 + \frac{R(\epsilon)}{2\ln\frac{1}{\epsilon}} + O\left(\frac{R(\epsilon)^2}{\ln^2\frac{1}{\epsilon}}\right)\right),$$

while the right hand side now reads (we use $\sqrt{1+t} \leq 1 + t/2$)

$$\begin{aligned}
&\frac{1}{\epsilon}\sqrt{2\ln\frac{1}{\epsilon}}\int\sqrt{I(\epsilon g^2)^2 + 4\epsilon^2 g^2|\nabla g|^2}\,\mathrm{d}\gamma \\
&= 2\sqrt{\ln\frac{1}{\epsilon}}\int g^2\sqrt{\ln\frac{1}{\epsilon} + \ln\frac{1}{g^2} + R(\epsilon g^2) + 2\frac{|\nabla g|^2}{g^2}}\,\mathrm{d}\gamma \\
&\leq 2\ln\frac{1}{\epsilon}\int g^2\left(1 + \frac{\ln\frac{1}{g^2} + 2\frac{|\nabla g|^2}{g^2} + R(\epsilon g^2)}{2\ln\frac{1}{\epsilon}}\right)\mathrm{d}\gamma \\
&= 2\ln\frac{1}{\epsilon} + \int g^2 R(\epsilon g^2)\,\mathrm{d}\gamma + \int g^2\ln\frac{1}{g^2}\,\mathrm{d}\gamma + 2\int|\nabla g|^2\,\mathrm{d}\gamma.
\end{aligned}$$

Thus we obtain

$$O\left(\frac{R(\epsilon)^2}{\ln\frac{1}{\epsilon}}\right) + \left(R(\epsilon) - \int g^2 R(\epsilon g^2)\,\mathrm{d}\gamma\right) \leq 2\int|\nabla g|^2\,\mathrm{d}\gamma - \int g^2\ln g^2\,\mathrm{d}\gamma.$$

We take $\epsilon \to 0$ and observe that the left hand side vanishes in the limit, which completes the proof.

Indeed, by the definition of the function R it is clear that the first term tends to 0. The expression in the brackets up to $\frac{1}{2}$ factor equals

$$\begin{aligned}
&\ln\ln\frac{1}{\epsilon} + r(\epsilon) - \int g^2\ln\left(\ln\frac{1}{\epsilon} + \ln\frac{1}{g^2}\right)\mathrm{d}\gamma - \int g^2 r(\epsilon g^2)\,\mathrm{d}\gamma \\
&= \left(r(\epsilon) - \int g^2 r(\epsilon g^2)\,\mathrm{d}\gamma\right) - \int g^2\ln\left(1 + \frac{\ln\frac{1}{g^2}}{\ln\frac{1}{\epsilon}}\right)\mathrm{d}\gamma.
\end{aligned}$$

By Lebesgue's dominated convergence theorem (recall that $1/N \leq g \leq N$) we get that the second term vanishes. The same theorem altogether with the fact that $r(\epsilon) \to -\ln\pi$ yields the terms in the brackets cancel out in the limit.

2.2 Hypercontractivity

We have already seen in Proposition 1.8, 6) that the $\Delta/2$ is the generator of the heat semi-group, which has proved to be very useful in the isoperimetric problems for the Gaussian measure. Here we introduce another powerful tool, the Ornstein-Uhlenbeck semi-group, the generator of which is the operator $L = \Delta - x \cdot \nabla$ responsible for the integration by parts formula for the Gaussian measure

$$\int fLg \, \mathrm{d}\gamma = -\int \nabla f \cdot \nabla g \, \mathrm{d}\gamma. \tag{2.7}$$

The aim is to explore the hypercontractivity properties of this semi-group.

2.6 Definition. For a continuous function $f \colon \mathbb{R}^n \longrightarrow \mathbb{R}$ of *a moderate growth* (we would not precise this), a number $t \geq 0$ and a point $x \in \mathbb{R}^n$ we define the *operator of the Ornstein-Uhlenbeck semi-group*

$$P_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1 - e^{-2t}}y) \, \mathrm{d}\gamma(y). \tag{2.8}$$

Let us begin with stating a few rudimentary yet useful properties, which are rather easy in proof.

2.7 Proposition. *The Ornstein-Uhlenbeck operator P_t satisfies*

1) $P_t(\lambda f + \mu g) = \lambda P_t f + \mu P_t g$

2) $P_s \circ P_t = P_{s+t}$

3) $P_t f \geq 0$, *if* $f \geq 0$

4) $P_s 1 = 1$

5) $P_0 f = f$, $P_t f(x) \xrightarrow[t \to \infty]{} \int f \, \mathrm{d}\gamma$

6) $\partial_{x_i} P_t f(x) = e^{-t} P_t(\partial_{x_i} f)(x)$ *and* $\nabla P_t f = e^{-t} P_t(\nabla f)$

7) $\partial_t P_t f(x) = L P_t f(x)$

8) $L P_t = P_t L$

9) $\int f P_t g \, \mathrm{d}\gamma = \int g P_t f \, \mathrm{d}\gamma$, *in particular* $\int P_t f \, \mathrm{d}\gamma = \int f \, \mathrm{d}\gamma$

10) $|P_t f(x)|^p \leq P_t(|f|^p)(x)$ *for* $p \geq 1$. *In particular*

$$\|P_t f\|_p \leq \|f\|_p$$

11) $P_t(fg) \le \sqrt{P_t(f^2)}\sqrt{P_t(g^2)}$ *(the Cauchy-Schwarz inequality).*

Proof. This is a matter of direct computations to check these properties, so this is left to the reader as an exercise. We only show how to proceed in 9), as a nice change of variables might be applied. Set θ such that $e^{-t} = \cos\theta$ and $\sqrt{1-e^{-2t}} = \sin\theta$. Then

$$\int f P_t g \,\mathrm{d}\gamma = \iint f(x) g(x\cos\theta + y\sin\theta)\,\mathrm{d}\gamma(y)\,\mathrm{d}\gamma(x),$$

so we put new variables $z = x\cos\theta + y\sin\theta$, $w = x\sin\theta - y\cos\theta$ and get

$$\int f P_t g \,\mathrm{d}\gamma = \iint f(z\cos\theta + w\sin\theta) g(z)\,\mathrm{d}\gamma(z)\,\mathrm{d}\gamma(w) = \int g P_t f \,\mathrm{d}\gamma.$$

□

As an example of application let us see how to derive the *Log*-Sobolev inequality for the Gaussian measure.

Proof of Theorem 2.3. Since entropy is homogeneous (this is left as an exercise to verify that $\mathrm{Ent}(tf) = t\,\mathrm{Ent}\, f$) we might and will assume without loss of generality that $\int f^2 \,\mathrm{d}\gamma = 1$. Notice that thanks to Property 5)

$$\mathrm{Ent}_\gamma f^2 = -\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{Ent}_\gamma P_t f^2 \,\mathrm{d}t.$$

Let us calculate the derivative. By virtue of Property 9) and our simplifying assumption $\int P_t f^2 \,\mathrm{d}\gamma = \int f^2 \,\mathrm{d}\gamma = 1$, so $\mathrm{Ent}_\gamma P_t f^2 = \int P_t f^2 \ln P_t f^2 \,\mathrm{d}\gamma$. Hence,

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{Ent}_\gamma P_t f^2 = \int \partial_t (P_t f^2) \ln P_t f^2 \,\mathrm{d}\gamma + \int P_t f^2 \frac{\partial_t (P_t f^2)}{P_t f^2} \,\mathrm{d}\gamma.$$

Because of Property 7) the second integral equals $\int L(P_t f^2)\,\mathrm{d}\gamma = 0$, as by integration by parts formula (2.7) with $f \equiv 1$ we get that in general $\int Lf \,\mathrm{d}\gamma = 0$. The first integral is tackled with the aid of integration by parts

$$\begin{aligned}
\int L(P_t f^2) \ln P_t f^2 \,\mathrm{d}\gamma &= -\int \nabla P_t f^2 \cdot \frac{1}{P_t f^2} \nabla P_t f^2 \,\mathrm{d}\gamma \\
&= -\int \frac{1}{P_t f^2} \left|\nabla P_t f^2\right|^2 \mathrm{d}\gamma \\
&= -\int 4e^{-2t} \frac{1}{P_t f^2} \left|P_t (f\nabla f)\right|^2 \mathrm{d}\gamma \\
&= -\sum_{i=1}^n \int 4e^{-2t} \frac{1}{P_t f^2} \big(P_t(f\partial_{x_i} f)\big)^2 \,\mathrm{d}\gamma.
\end{aligned}$$

The Cauchy-Schwarz inequality (see Property 11)) yields

$$\big(P_t(f\partial_{x_i}f)\big)^2 \le (P_tf^2)\big(P_t(\partial_{x_i}f)^2\big).$$

Therefore,

$$\int L(P_tf^2)\ln P_tf^2 \,\mathrm{d}\gamma \ge -4e^{-2t}\int P_t\big(|\nabla f|^2\big)\,\mathrm{d}\gamma = -4e^{-2t}\int |\nabla f|^2\,\mathrm{d}\gamma,$$

where in the last equality we use Property 9). We thus conclude as follows

$$\mathrm{Ent}_\gamma f^2 \le \int_0^\infty 4e^{-2t}\Big(\int |\nabla f|^2\,\mathrm{d}\gamma\Big)\,\mathrm{d}t = 2\int |\nabla f|^2\,\mathrm{d}\gamma.$$

□

2.8 Remark. Using the same method one can establish the *Poincaré inequality* for the Gaussian measure which states that for a differentiable function $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $|\nabla f| \in L_2(\gamma)$ there holds

$$\mathrm{Var}_\gamma f \le \int |\nabla f|^2\,\mathrm{d}\gamma. \tag{2.9}$$

2.9 Exercise. Prove that $\mathrm{Ent}_\gamma P_tf \le e^{-2t}\,\mathrm{Ent}_t f$.

Having the *Log*-Sobolev at our disposal we now prove in a neat way the main theorem of this section.

2.10 Theorem (Hypercontractivity for the Gaussian measure)**.** *Given numbers $1 < p < q < +\infty$ and $t \ge 0$ such that $e^{-2t} \le \frac{p-1}{q-1}$ there holds the hypercontractivity inequality*

$$\|P_tf\|_q \le \|f\|_p. \tag{2.10}$$

Moreover, if $q > 1+e^{2t}(p-1)$, then P_t is not continuous as a linear operator from $L_p(\gamma)$ to $L_q(\gamma)$.

Stating the main result it seems appropriate to briefly mention its history as it is quite involved. It was E. Nelson who gave an early version of the above theorem in 1966 [Nel1]. Then his result was improved several times and at last in 1973 Nelson himself gave the full proof in the Gaussian setting [Nel2]. L. Gross followed a different path through Theorem 2.10. In the paper [Gro] in 1975 he discovered the *Log*-Sobolev inequality on the discrete cube and using CLT he obtained Theorem 2.3. Out of this result inequality

(2.10) can be neatly deduced as we will see following Gross. Yet it was not he who first proved *Log*-Sobolev inequality in the boolean setting, for A. Bonami did it only in 1970, [Bon]. In computer science literature the name of W. Beckner is used to being attached to the theorem due to the paper [Bec].

Proof of Theorem 2.10. To see the second part take $f_\lambda(x) = e^{-\lambda x}$, check that
$$\frac{\|P_t f\|_q}{\|f\|_p} = \exp\left[\frac{\lambda^2}{2}\Big(1-p-e^{-2t}(1-q)\Big)\right]$$
and conclude letting $\lambda \to \infty$.

For the proof of the desired inequality let us define
$$\begin{aligned} q(t) &= 1 + e^{2t}(p-1), \\ F(t) &= \ln \|P_t f\|_{q(t)}. \end{aligned}$$
The assertion is equivalent to $F(t) \le F(0)$. Since $|P_t f| \le P_t|f|$ we may assume without loss of generality that f is nonnegative. We write explicitly
$$F(t) = \frac{1}{q(t)} \ln \int e^{q(t)\ln P_t f}\, \mathrm{d}\gamma$$
and differentiate
$$F'(t) = -\frac{q}{q^2} \ln \int (P_t f)^q + \frac{1}{q\int (P_t f)^q} \int (P_t f)^q \left(q' \ln P_t f + q\frac{LP_t f}{P_t f}\right).$$
To shorten the notation we put $g = P_t f$ and continue the computation
$$\begin{aligned} F'(t) &= -\frac{q'}{q^2\int g^q}\left(\int g^q \ln \int g^q - \int g^q \ln g^q - \frac{q^2}{q'}\int g^{q-1}Lg\right) \\ &= \frac{q'}{q^2\int g^q}\left(\mathrm{Ent}_\gamma(g^q) + \frac{q^2}{q'}\int g^{q-1}Lg\right). \end{aligned}$$
Observe that integration by parts yields
$$\begin{aligned} \frac{q^2}{q'}\int g^{q-1}Lg &= -\frac{q^2}{2(q-1)}\int \nabla(g^{q-1})\cdot\nabla g \\ &= -\frac{q^2}{2}\int g^{q-2}|\nabla g|^2 = -2\int |\nabla g^{q/2}|^2. \end{aligned}$$
Thus by virtue of the *Log*-Sobolev inequality (2.3)
$$F'(t) = \frac{q'}{q^2\int g^q}\left(\mathrm{Ent}_\gamma(g^q) - 2\int |\nabla g^{q/2}|^2\right) \le 0.$$
This completes the proof. $\square$

3 Lecture V — Gaussian chaoses via the Hermite polynomials

The Hermite polynomials are yet another tool which addresses inequalities involving the Gaussian measure. In this section we are going to describe their basic properties first in dimension one and then in higher dimensions. We stress the connections of these polynomials with the operators L and P_t. Making use of the hypercontractivity we end up with the estimation concerning moments of the so-called Gaussian chaoses, which is an analogue of the Khintchine inequality for polynomials of higher degree than one.

3.1 Dimension one

We begin with a definition. 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 $L_2(\mathbb{R},\gamma)$ is called the sequence of *Hermite polynomials*. That is,

$$\begin{aligned} H_0 &= 1, \\ H_m &= x^m - \sum_{i=0}^{m-1} \langle x^m, H_i\rangle \frac{H_i}{\|H_i\|^2}. \end{aligned} \tag{3.1}$$

There are other equivalent definitions. For instance,

$$H_m(x) = (-1)^n e^{x^2/2} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(e^{-x^2/2}), \tag{3.2}$$

or via the generating function

$$e^{xt-t^2/2} = \sum_{m\geq 0} \frac{H_m(x)}{m!} t^m. \tag{3.3}$$

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

$$H_m(x) = m!\sum_{k=0}^{m} \frac{\cos(k\pi/2)}{2^{k/2}(k/2)!}\frac{x^{m-k}}{(m-k)!} = m!\sum_{0\leq k\leq m/2} \frac{(-1)^k}{2^k\cdot k!}\frac{x^{m-2k}}{(m-2k)!}. \tag{3.4}$$

What makes the Hermite polynomials significant is the fact that they form a basis.

3.1 Theorem. *The Hermite polynomials* $(H_m)_{m\geq 0}$ *form an orthogonal basis of* $L_2(\mathbb{R},\gamma)$.

Proof. Let $f \in L_2(\gamma)$ be such that $\langle f, H_m\rangle = 0$ for every $m \geq 0$. Exploiting a little bit of Fourier analysis we show that $f = 0$ γ-a.e. Since also $\langle f, x^m\rangle = 0$ for all $m \geq 0$, we get that the entire function

$$F(z) = \int_{\mathbb{R}} f(x)e^{-x^2/2}e^{zx}\,\mathrm{d}x$$

is zero as it equals $\sum_{m\geq 0}\frac{z^m}{m!}\sqrt{2\pi}\langle f, x^m\rangle$. Therefore

$$F(-it) = 0 = \int_{\mathbb{R}} f(x)e^{-x^2/2}e^{-itx}\,\mathrm{d}x = \widehat{f(x)e^{-x^2/2}}(t)$$

for every $t \in \mathbb{R}$. Thus $f(x)e^{-x^2/2} = 0$ a.e. $\square$

Now let us link the Hermite polynomials with the integration by parts operator L (see 2.7) as well as with the Ornstein-Uhlenbeck operator P_t (see 2.8). To start with we make two simple observations.

3.2 Proposition. *(i)* $H_m' = mH_{m-1}$ *for* $m \geq 1$,

(ii) $LH_m = -mH_m$ *for* $m \geq 0$.

Proof. *(i)* It follows by formula (3.4) 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

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

as the polynomial $H_i'(x) - xH_i(x)$ is of degree $i+1 \leq m-1$.

(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 $i \leq m-1$

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

as LH_i is of degree i. $\square$

Clearly, $LH_m = -mH_m$ means that H_m is an eigenvalue of the operator L. Due to Theorem 3.1 we arrive at the desired links.

3.3 Theorem. *All the eigenvalues of the operator L acting on $L_2(\gamma)$ are nonpositive integers $-m$, $m \geq 0$ and the corresponding eigenvectors are H_m.*

3.4 Theorem. *All the eigenvalues of the operator P_t acting on $L_2(\gamma)$ are e^{-tm}, $m \geq 0$ and the corresponding eigenvectors are H_m.*

Proof. Note that (Proposition 2.7, 8))

$$LP_tH_m = P_tLH_m = -mP_tH_m,$$

so Theorem 3.3 yields that P_tH_m is an eigenvector of L. As a consequence there is a number $\lambda(t)$ such that $P_tH_m = \lambda(t)H_m$. By virtue of Property 7) of the Ornstein-Uhlenbeck operator we get that $\lambda(t)$ is differentiable as a function of $t \geq 0$ and

$$\lambda'(t)H_m = \partial_t P_tH_m = LP_tH_m = -m\lambda(t)H_m.$$

Moreover,

$$H_m = P_0H_m = \lambda(0)H_m.$$

Thus λ solves the Cauchy problem

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

so $\lambda(t) = e^{-tm}$. □

We set up the notation

$$h_m = H_m/\|H_m\|_{L_2(\gamma)}.$$

Then $(h_m)_{m\geq 0}$ is an orthogonal basis in $L_2(\gamma)$.

3.2 Dimension n

We consider the space $L_2(\mathbb{R}^n, \gamma)$ with the usual inner product. Multiindices will be useful. Let us recall the standard notation. Adopting the convention that $\mathbb{N} = \{0, 1, 2, \ldots\}$ for an index $m = (m_1, \ldots, m_n) \in \mathbb{N}^n$ we denote $|m| = \sum_{i=1}^n m_i$ and $m! = \prod_{i=1}^n m_i!$.

Let us introduce the multidimensional Hermite polynomials on $\mathbb{R}^n$. Given $m \in \mathbb{N}^n$ we put

$$H_m(x) = \prod_{i=1}^n H_{m_i}(x_i), \tag{3.5}$$

where H_{m_i} is a Hermite polynomial of one variable defined in (3.1). This is a polynomial of degree $|m|$. We also define

$$h_m(x) = H_m / \|H_m\|_{L_2(\mathbb{R}^n,\gamma)}.$$

3.5 Theorem. *The Hermite polynomials $(h_m)_{m\in\mathbb{N}^n}$ forms an orthogonal basis of $L_2(\mathbb{R}^n,\gamma)$.*

For instance, all the Hermite polynomials of degree up to 2 read as follows

$$\begin{array}{lll} |m| = 0, & & H_0 \equiv 1, \\ |m| = 1, & m = e_i, & H_{e_i}(x) = x_i, \\ |m| = 2, & m = e_i + e_j, & H_{e_i e_j}(x) = \begin{cases} x_i x_j & \text{if } i \neq j \\ x_i^2 - 1 & \text{if } i = j \end{cases}. \end{array}$$

Given $k \geq 0$ let us define the linear subspace $\mathcal{H}_k$ in $L_2(\gamma)$ spanned by the Hermite polynomials of degree k

$$\mathcal{H}_k = \operatorname{span}\{H_m,\ |m| = k\}. \tag{3.6}$$

Let Π_k be the orthogonal projection onto $\mathcal{H}_k$. Then for any function $f \in L_2(\gamma)$

$$f = \sum_{m\in\mathbb{N}^n} \langle f, h_m\rangle h_m = \sum_{k=0}^{\infty} \sum_{|m|=k} \langle f, h_m\rangle h_m = \sum_{k=0}^{\infty} \Pi_k f,$$

where $\Pi_k f$ is called the *Gaussian chaos* of order k of the function f. For example, the Gaussian chaoses of order 1 are of the form

$$\sum_i a_i x_i,$$

while these of order 2 take form

$$\sum_{i,j} a_{ij} x_i x_j + \sum_i a_i (x_i^2 - 1).$$

Our task is now to devise estimations of the norms of Gaussian chaoses, e.g. we ask whether the quantity

$$\left(\mathbb{E} \left| \sum_{i,j} a_{ij} g_i g_j + \sum_i a_i (g_i^2 - 1) \right|^p \right)^{1/p}$$

is comparable with the same one for, say $p = 1$? This would be the generalization of Khintchine type inequalities but for nonlinear expressions (chaoses) involving i.i.d. standard normal random variables.

Like in dimension one the importance of the Hermite polynomials is revealed in their spectral properties with respect to the operators L and P_t.

3.6 Proposition. *For any* $m \in \mathbb{N}^n$

$$LH_m = -|m|H_m, \tag{3.7}$$

$$P_t H_m = e^{-t|m|}H_m. \tag{3.8}$$

Proof. 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,

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

$$= \sum_{i=1}^m -m_iH_m(x) = -|m|H_m.$$

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. $\square$

3.7 Corollary. $P_t\Pi_k f = e^{-tk}\Pi_k f$.

Now we are ready to formulate and prove the advertised main result.

3.8 Theorem. *Let* $q \geq 2$. *For a Gaussian chaos of degree* d *of a function* f *we have*

$$\|\Pi_d f\|_{L_q(\gamma)} \leq \sqrt{q-1}^{\,d}\|\Pi_d f\|_{L_2(\gamma)}. \tag{3.9}$$

More generally, if Q *is a polynomial of degree* d *of* n *variables, then*

$$\|Q\|_{L_q(\gamma,\mathbb{R}^n)} \leq \sqrt{d+1}\sqrt{q-1}^{\,d}\|Q\|_{L_2(\gamma,\mathbb{R}^n)}. \tag{3.10}$$

Proof. The comparison (3.9) of moments of chaoses is a simple consequence of the hypercontraction thanks to Corollary 3.7. Namely, take t so that $e^t = \sqrt{q-1}$. Then by Theorem 2.10 with $p = 2$ we obtain

$$\|\Pi_d f\|_q = e^{td}\|P_t \Pi_d f\|_q \leq \sqrt{q-1}^d \|\Pi_d f\|_2.$$

For the proof of the second formula write $Q = \sum_{k=0}^{d} Q_k$, where $Q_k = \Pi_k Q$. Then

$$\begin{aligned}
\|Q\|_q &\leq \sum_{k=0}^{d} \|Q_k\|_q \leq \sqrt{q-1}^d \sum_{k=0}^{d} \|Q_k\|_2 \\
&\leq \sqrt{q-1}^d \sqrt{d+1} \left(\sum_{k=0}^{d} \|Q_k\|_2^2 \right)^{1/2} = \sqrt{q-1}^d \sqrt{d+1} \|Q\|_2,
\end{aligned}$$

where in the last equality Pythagorean theorem is used. □

3.9 Remark. If $Q = \sum_{k=l}^{d} Q_k$, then

$$\|Q\|_q \leq \sqrt{q-1}^d \sqrt{d-l+1}\|Q\|_2.$$

In particular, applying this for linear forms $Q(x) = \sum_i a_i x_i$ we recover the Khintchine inequality for the Gaussian measure

$$\left\| \sum a_i g_i \right\|_q \leq \sqrt{q-1} \left\| \sum a_i g_i \right\|_2, \qquad a_i \in \mathbb{R}.$$

3.10 Remark (Research Problem)**.** We have seen that for the Gaussian measure q-th moments of polynomials are of order $\sqrt{q}$. It is known that for arbitrary log-concave measures there is no chance to get it better than q (see e.g. [Bob2]). However, log-concave measures μ with densities $\mathrm{d}\mu(x) = e^{-V(x)}\,\mathrm{d}x$, where $\mathrm{Hess} V \geq c\,\mathrm{id}$, for a constant $c > 0$ are expected to exhibit behaviour closer to the Gaussian case. For example, using the localization technique the optimal isoperimetry has been derived for such measures (see [Bob1]). Our question is whether

$$\|Q\|_{L_q(\mu)} \leq C\sqrt{d}\sqrt{q}^d \|Q\|_{L_2(\mu)}?$$

To face the problem one might try to adopt the program of the paper [Fr] by M. Fradelizi exploiting the localization approach (see also the aforementioned paper by S. Bobkov [Bob1] in order to get familiar with techniques applicable to measures μ).

3.11 Exercise. Using the Hermite polynomials prove the Poincaré inequality (2.9).

Part II
Olivier's lectures

4 Lecture III — Distances between convex bodies

We try to set up basic facts concerning measurement of a distance between convex sets. The goal is to present a metric space of convex bodies which is compact. Let us start with two definitions.

4.1 Definition. Given two symmetric convex bodies K and L in $\mathbb{R}^n$ we define
$$d(K,L) = \inf\{s > 0, \exists T \in \mathrm{GL}_n\ K \subset TL \subset sK\}. \tag{4.1}$$

4.2 Definition. Given two isomorphic Banach spaces X and Y we define
$$\rho(X,Y) = \inf\{\|T\| \cdot \|T^{-1}\|,\ T\colon X \longrightarrow Y \text{ isomorphism}\}. \tag{4.2}$$

The functions d and ρ are sometimes referred to as the Banach-Mazur distances.

4.3 Proposition. *Let K and L be two symmetric convex bodies in $\mathbb{R}^n$ and let $X = (\mathbb{R}^n, K)$, $Y = (\mathbb{R}^n, L)$ be two Banach spaces with the norms such that K, L are the unit balls. Then*
$$\rho(X,Y) = d(K,L).$$

Proof. First assume s is a number such that there exists a matrix $T \in \mathrm{GL}_n$ and $K \subset TL \subset sK$. Let us think of T as an isomorphism from Y to X. Since $TL \subset sK$ we have $\|T\|_{Y\to X} = \sup_{y\in L} \|Ty\|_X \le s$. Similarly, $T^{-1}K \subset L$ yields $\|T^{-1}\|_{X\to Y} = \sup_{x\in K} \|Tx\|_Y \le 1$. Thus $\|T\|\cdot\|T^{-1}\| \le s$ which implies $\rho(X,Y) \le s$. Taking infimum over all possible s we get $\rho(X,Y) \le d(K,L)$.

Now consider any isomorphism $T\colon X \longrightarrow Y$. By the definition of the operator norm for any $x \in K$ we have $\|Tx\|_Y \le \|T\|$, which means that $TK \subset \|T\|L$. Similarly, $T^{-1}L \subset \|T^{-1}\|K$. Therefore $L \subset (\|T^{-1}\|T)K \subset \|T^{-1}\| \cdot \|T\|L$, so the inequality $d(K,L) \le \|T^{-1}\| \cdot \|T\|$ follows from (4.1). Taking infimum over all possible T we obtain $d(K,L) \le \rho(X,Y)$. $\square$

4.4 Example. The balls B_1^n and B_∞^n are extremal in some sense. It is interesting to ask on the distance between them. Equivalently, what is $\rho(\ell_1^n, \ell_\infty^n)$?

For n being a power of 2, say $n = 2^m$, we are to construct explicitly a matrix T such that $B_\infty^n \subset TB_1^n \subset \sqrt{n}B_\infty^n$ which proves that

$$\rho(\ell_1^n, \ell_\infty^n) \leq \sqrt{n}, \qquad \text{when } n = 2^m.$$

Let $W_0 = [1]$, $W_1 = \left[\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right]$, and define by recurrence the *k-th Walsh matrix* W_k of size $2^k \times 2^k$ by $W_k = \left[\begin{smallmatrix} W_{k-1} & -W_{k-1} \\ -W_{k-1} & W_{k-1} \end{smallmatrix}\right]$. One checks that $W_k^T W_k = 2^k I$, i.e. the matrix $\frac{1}{\sqrt{2^m}} W_m$ is orthogonal. Since the columns of W_m have entries ± 1, we have $W_m e_i \in B_\infty^n$. By linearity $W_m B_1^n = \operatorname{conv}\{\pm W_m(e_i), 1 \leq i \leq n\} \subset B_\infty^n$. On the other hand, $B_1^n \supset \frac{1}{\sqrt{n}} B_2^n$, whence $W_m B_1^n \supset \frac{1}{\sqrt{n}} W_m B_2^n = B_2^m$, as $\frac{1}{\sqrt{2^m}} W_m$ is orthogonal. Therefore

$$B_\infty^n \supset W_m B_1^n \supset B_2^n \supset \frac{1}{\sqrt{n}} B_\infty^n.$$

We may take $T = \sqrt{n} W_m$.

Now we show that

$$\rho(\ell_1^n, \ell_\infty^n) \geq \frac{\sqrt{n}}{e}.$$

For this purpose take T and s satisfying $B_\infty^n \subset TB_1 \subset sB_\infty^n$. It is enough to prove that $s \geq \sqrt{n}/e$. Observe that $|Te_i|_\infty \leq s$, for $Te_i \in sB_\infty^n$. Volume argument yields $|B_\infty^n| \leq |TB_1^n| = |\det T| \cdot |B_1^n|$. By Hadamard's inequality

$$|\det T| \leq \prod_{i=1}^n |Te_i| \leq \prod_{i=1}^n \sqrt{n} |Te_i|_\infty \leq (\sqrt{n} s)^n.$$

Thus

$$2 = |B_\infty^n|^{1/n} \leq |\det T|^{1/n} \cdot |B_1^n|^{1/n} \leq s\sqrt{n} \frac{2}{(n!)^{1/n}},$$

hence

$$s \geq \frac{(n!)^{1/n}}{\sqrt{n}} \geq \frac{\sqrt{n}}{e}.$$

It is time to establish a very useful lemma due to H. Auerbach, which helps investigating distances between spaces.

4.5 Lemma (Auerbach)**.** *Let X be a Banach space of dimension n. Then there exists a biorthogonal system (x_j, x_j^*), $1 \leq j \leq n$, in $X \times X^*$ (i.e. these vectors satisfy $\langle x_k^*, x_l \rangle = \delta_{kl}$) such that*

$$\|x_j\| = 1 \text{ and } \|x_j^*\| = 1, \qquad \text{for every } j.$$

Proof. Take any biorthogonal system (y_j, y_j^*) in $X \times X^*$ (it is enough to consider a base and its dual). We define the function

$$V(z_1, \dots, z_n) = \det \left[y_j^*(z_i) \right]_{i,j=1,\dots,n} \qquad \text{on } X \times \ldots \times X.$$

It is continuous, hence it attains its supremum on the compact set $B_X \times \ldots \times B_X$ at, say $(x_1, \dots, x_n) \in B_X \times \ldots \times B_X$. Necessarily $\|x_i\| = 1$ for all i, for otherwise $(x_1, \dots, x_n)$ would not give the maximal value as one would take $\lambda x_i \in B_X$, for some $\lambda > 1$ obtaining a greater value. For a fixed index j let us define the functional

$$x_j^*(x) = \frac{V(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_n)}{V(x_1, \dots, x_j)}, \qquad x \in X.$$

Then $x_j^*(x_i) = 0$, if $i \neq j$, as the determinant of a matrix with two identical columns equals 0. Clearly $x_j^*(x_j) = 1$. Moreover, since V on the set $B_X \times \ldots \times B_X$ attains its maximum at $(x_1, \dots, x_n)$, we have $\sup_{x \in B_X} x_j(x) = 1$, so $\|x_j^*\| = 1$. The proof is now complete. $\square$

4.6 Exercise. Let $(E, \|\cdot\|)$ be a n dimensional Banach space. Prove that there exist a basis $(x_1, \dots, x_n)$ such that for any scalars $a_1, \dots, a_n \in \mathbb{R}$

$$\sup_{1 \leq i \leq n} |a_i| \leq \left\| \sum_{i=1}^n a_i x_i \right\| \leq \sum_{i=1}^n |a_i|.$$

Now we are able to define a metric space of classes of convex bodies in given dimension, or equivalently, a metric space of classes of Banach spaces with fixed dimension, and then prove it is compact.

4.7 Definition. Given n let $\mathcal{F}_n$ be the set of all Banach spaces of dimension n. Take the equivalence relation $\sim$ defined as $E \sim F$ iff $\rho(E, F) = 1$ and its quotient set $\widetilde{\mathcal{F}}_n = \mathcal{F}_n / \sim$. We define the metric p on $\widetilde{\mathcal{F}}_n$ by the formula $p(E, F) = \ln \rho(E, F)$.

4.8 Proposition. *The metric space $(\widetilde{\mathcal{F}}_n, p)$ is compact.*

Proof. We start with showing that for every space $E \in \mathcal{F}_n$

$$\rho(E, \ell_1^n) \leq n.$$

Let us consider the isomorphism $T \colon \ell_1^n \longrightarrow E$ defined so that $e_i \mapsto x_i$, for every $i = 1, \dots, n$, where (x_j, x_j^*), $j = 1, \dots, n$ is a biorthogonal system which comes from Auerbach's lemma. Consider a point $\lambda = (\lambda_1, \dots, \lambda_j) \in \ell_1^n$. On the one hand

$$\|T\lambda\| = \left\|\sum \lambda_j x_j\right\| \leq \sum |\lambda_j| \cdot \|x_j\| = \sum |\lambda_j| = |\lambda|_1,$$

so $\|T\| \leq 1$. On the other hand for any i

$$\|T\lambda\| \geq |x_i^*(T\lambda)| = |\lambda_i|,$$

thus

$$\|T\lambda\| \geq \max_i |\lambda_i| \geq \frac{1}{n}|\lambda|_1,$$

which gives that $\|T^{-1}\| \leq n$. Therefore $\|T\| \cdot \|T^{-1}\| \leq n$ and $\rho(E, \ell_1^n) \leq n$.

Let Φ_n be the set of all norms $\|\cdot\|$ on $\mathbb{R}^n$ which satisfy the inequality

$$\frac{1}{n}|x|_1 \leq \|x\| \leq |x|_1, \qquad x \in \mathbb{R}^n. \tag{4.3}$$

Having this define the subset of $C(B_1^n)$

$$\widetilde{\Phi}_n = \{f \colon B_1^n \longrightarrow \mathbb{R},\ f(\cdot) = \|\cdot\| \text{ for some } \|\cdot\| \in \Phi_n\}.$$

By what has been shown at the beginning, i.e. $\rho(E, \ell_1^n) \leq n$, the natural mapping $(\widetilde{\Phi}_n, \|\cdot\|_{C(B_1^n)}) \xrightarrow{S} (\widetilde{\mathcal{F}}_n, p)$ is surjective. Moreover, it is continuous. Indeed, let us take two close points $f, g \in \widetilde{\Phi}_n$, say $\epsilon = \sup_{x \in B_1^n} |f(x) - g(x)|$. By this and (4.3) we get for every x

$$|\|x\|_f - \|x\|_g| \leq \epsilon |x|_1 \leq n\epsilon \|x\|_f, n\epsilon\|x\|_g.$$

Thus

$$(1 - \epsilon n)\|x\|_g \leq \|x\|_f \leq (1 + \epsilon n)\|x\|_g$$

and, in other words,

$$p(S(f), S(g)) \leq \ln \frac{1 + \epsilon n}{1 - \epsilon n} \xrightarrow[\epsilon \to 0]{} 0.$$

If we know that the set $\widetilde{\Phi}_n$ is compact in the space $C(B_1^n)$, the proof will be finished because then its image $\widetilde{\mathcal{F}}_n$ under the continuous mapping S also has to be compact. This is the case due to the Arzelà-Ascoli theorem (functions from the set $\widetilde{\Phi}_n$ are uniformly bounded by 1 thanks to (4.3) and obviously equicontinuous as norms are 1-Lipschitz mappings). $\square$

5 Lecture IV — Ideals of operators and John's theorem

We start with collecting a few abstract facts about various norms of linear operators. Making use of them we then derive a famous John's theorem on an ellipsoid of maximal volume contained in a convex body (see e.g. [Ball1] for a short different proof).

5.1 Operator norms

Given two Banach spaces X and Y we consider the set $\mathcal{L}(X,Y)$ of all linear and continuous mappings from X to Y. This is a Banach algebra with the norm given by

$$\|u\| = \sup_{\|x\|_X \leq 1} \|u(x)\|_Y,$$

which is called an *operator norm*. There is a particular nice ideal $\mathcal{F}(X,Y) \subset \mathcal{L}(X,Y)$ of all mappings of finite rank. It can be seen as a tensor product $X^* \otimes Y$ which is a linear space spanned by rank one operators

$$\begin{aligned} \xi \otimes y \colon \ X &\longrightarrow Y, \\ x &\longmapsto \xi(x)y. \end{aligned}$$

Each operator $u \in \mathcal{F}(X,Y)$ can be written as

$$u = \sum_{i=1}^{m} \xi_i \otimes y_i, \qquad \text{for some } \xi_i \in X^*,\ y_i \in Y,$$

although such a representation might not be unique. We come to the two important definitions of *injective norm* of an finite-rank operator $u \in \mathcal{F}(X,Y)$

$$\|u\|_\wedge = \inf \left\{ \sum_{i=1}^{m} \|\xi_i\|_{X^*} \|y_i\|_Y,\ u = \sum_{i=1}^{m} \xi_i \otimes y_j \right\} \tag{5.1}$$

and its *trace*

$$\operatorname{tr} u = \sum_{i=1}^{m} \xi_i(x_i). \tag{5.2}$$

5.1 Exercise. Prove that the trace is well defined, i.e. $\sum_{i=1}^m \xi_i \otimes x_i = 0$ implies $\sum_{i=1}^m \xi_i(x_i) = 0$.

5.2 Exercise. Let $u \in \mathcal{F}(X,Y)$ and $v \in \mathcal{F}(Y,X)$. Then $\operatorname{tr} uv = \operatorname{tr} vu$.

From now on we will be concerned only with finite dimensional spaces. Then the situation is quite simpler. Observe that $\mathcal{L}(X,Y)=\mathcal{F}(X,Y)$. Then one should ask what is a natural representation of a mapping $u\in\mathcal{L}(X,Y)$ given by a matrix A. Denote its columns by $a_1,\dots,a_m$, i.e. $u(e_i)=a_i$ and let $\xi_i=e_i^*$ be the dual basis to the basis $(e_i)_{i=1}^m$ of X in which the matrix A is written. We have $u=\sum_{i=1}^m e_i^*\otimes a_i$.

5.3 Exercise. Let ϕ be a linear functional on $\mathcal{L}(X,Y)$. Then there exists an operator $w\in\mathcal{L}(X,Y)$ such that $\phi(u)=\operatorname{tr} uw$ for any $u\in\mathcal{L}(X,Y)$.

The last but not least definition addresses duality concept in the space of linear transformations. For a norm α on the space $\mathcal{L}(X,Y)$ we set its *dual norm* α^* to be

$$\alpha^*(u)=\sup\{|\operatorname{tr} uw|,\ \alpha(w)\le 1, w\in\mathcal{L}(X,Y)\}\,. \tag{5.3}$$

To get familiar with these concepts let us look at some basic facts, which later turn out to be indispensable in our pursuit of the John's theorem. In the first preposition the injective norm of the simplest operator is calculated while the second preposition reveals how one can actually devise the concept of this norm.

5.4 Proposition. *If* $\dim X=n$ *then* $\|\operatorname{id}\colon X\longrightarrow X\|_\wedge=n$.

Proof. Assume $\operatorname{id}=\sum_{i=1}^m \xi_i\otimes x_i$ for some $\xi_i\in X^*$ and $x_i\in X$. Then

$$n=\operatorname{tr}\operatorname{id}=\sum\operatorname{tr}\xi_i\otimes x_i=\sum\xi_i(x_i)\le\sum\|\xi_i\|\cdot\|x_i\|,$$

so $n\le\|\operatorname{id}\|_\wedge$. On the other hand, by the Auerbach lemma there is a biorthogonal system $(x_i^*,x_i)_{i=1}^n$ in $X^*\times X$ such that $\|x_i^*\|=1=\|x_i\|$. In particular $\sum\|x_i^*\|\cdot\|x_i\|=n$, but $\operatorname{id}=\sum x_i^*\otimes x_i$. Thus $\|\operatorname{id}\|_\wedge=n$. $\square$

5.5 Proposition. *The projective and operator norms on* $\mathcal{L}(X,Y)$ *are dual in the sense of* (5.3), *i.e.*

$$|\!|\!|\cdot|\!|\!|^*=\|\cdot\|_\wedge\,.$$

Proof. Define the sets

$$D=\{f^*\otimes e,\ f\in X^*, e\in Y, \|f\|\le 1, \|e\|\le 1\},$$
$$C=B_{|\!|\!|\cdot|\!|\!|^*}=\{u\in\mathcal{L}(X,Y),\ |\!|\!|u|\!|\!|^*\le 1\}.$$

Since $\operatorname{conv} D = B_{\|\cdot\|_\wedge}$, the target is to establish that $\operatorname{conv} D = C$. To see the first inclusion, take an element $f^* \otimes e$ from D and observe that for any $w \in \mathcal{L}(X,Y)$ with $|||w||| \le 1$ we have

$$\begin{aligned} |\operatorname{tr} w(f^* \otimes e)| = |\operatorname{tr} f^* \otimes w(e)| = |f^*(w(e))| \le \|f^*\| \cdot \|w(e)\| \\ \le \|f^*\| \cdot |||w||| \cdot \|e\| \le 1, \end{aligned}$$

which means that $f \in C$. Therefore $D \subset C$ and consequently $\operatorname{conv} D \subset C$.

To see the second inclusion $\operatorname{conv} D \supset C$, by standard duality arguments it is enough to show that $D^\circ \subset C^\circ$, as then $\operatorname{conv} D = (D^\circ)^\circ \supset (C^\circ)^\circ = C$. Let us stress that this duality is understood with respect to tr in order to be coherent with our definition of the dual norm (5.3). Namely for a set $K \subset \mathcal{L}(X,Y)$ we put here its dual to be

$$K^\circ = \{u \in \mathcal{L}(X,Y),\ \forall v \in K\ |\operatorname{tr} uv| \le 1\}.$$

The place where the trace come into play is that we would like to know $B_{|||\cdot|||} \subset C^\circ$ since then it will be enough to prove that $D^\circ \subset B_{|||\cdot|||}$. To see the former, take $u \in B_{|||\cdot|||}$. By (5.3) we have $|\operatorname{tr} uv| \le 1$ for any v with $|||v|||^* \le 1$, which proves that indeed $u \in C^\circ$.

After this cumbersome clarification let us check that $D^\circ \subset B_{|||\cdot|||}$. Take $u \in D^\circ$. Then we know that for any $v \in D$ we have $|\operatorname{tr} uv| \le 1$. Let us put here $v = f^* \otimes e$ where $e \in B_{\|\cdot\|}$ is chosen so that $|||u||| = \|u(e)\|$ and f^* so that $\|u(e)\| = f^*(u(e))$. We get $\operatorname{tr} uv = f^*(u(e)) = |||u|||$ which combined with the fact that this trace is at most 1 finishes the proof. $\square$

The alluded abstract tool addresses existence of certain operator.

5.6 Theorem (Lewis, 78')**.** *Let X and Y be normed spaces of dimension n. Suppose α is a norm on $\mathcal{L}(X,Y)$. Then there exists an operator $v \in \mathcal{L}(X,Y)$ such that $\alpha(v) = 1$ and $\alpha^*(v^{-1}) = n$.*

5.7 Remark. If $\alpha(u) = 1$, then $\alpha^*(u^{-1}) \ge \operatorname{tr} uu^{-1} = n$.

5.8 Remark. The examples of norms which play an important role in the theory read as follows

a) $\alpha(u) = \left(\int_{\mathbb{R}^n} \|u(x)\|^2 \,\mathrm{d}\gamma(x)\right)^{1/2}$

b) $\alpha(u) = |||u|||$, $\alpha^*(u) = \|u\|_\wedge$ (see Proposition 5.5)

c) See Chapter 2 of the monograph [Pi] for the notion of K-convexity.

Proof of Theorem 5.6. Let $K = \{u \in \mathcal{L}(X,Y),\ \alpha(u) \le 1\}$. We consider the function $u \longmapsto \det u$ on K, which being continuous attains its maximum on compacts. Therefore there is $v \in \mathcal{L}(X,Y)$ such that

$$\det v = \max_{u \in K} \det u.$$

Clearly, $\alpha(v) = 1$ (the relevant argument supporting this has been given in the proof of Lemma 4.5). In order to prove $\alpha^*(v) = n$ let us notice that

$$\begin{aligned}\det v \ge \det\left(\frac{v+\epsilon w}{\alpha(v+\epsilon w)}\right) &\ge \frac{1}{(1+\epsilon\alpha(w))^n}\det(v+\epsilon w)\\ &= \frac{1}{(1+\epsilon\alpha(w))^n}(\det v)\det(\mathrm{id} + \epsilon v^{-1}w),\end{aligned}$$

hence

$$1 + \epsilon\operatorname{tr}(v^{-1}w) + o(\epsilon) = \det(\mathrm{id} + \epsilon v^{-1}w) \le (1+\epsilon\alpha(w))^n = 1 + n\epsilon\alpha(w) + o(\epsilon).$$

This yields that for any nonzero $w \in \mathcal{L}(X,Y)$ we have $\operatorname{tr} v^{-1}\frac{w}{\alpha(w)} \le n$, whence $\alpha^*(v^{-1}) \le n$. The observation that necessarily $\alpha^*(v^{-1}) \ge n$ made in Remark 5.7 completes the proof. $\square$

5.2 John's theorem

This is a matter of making use of compactness that given a convex body $K \subset \mathbb{R}^n$ there exists an ellipsoid $\mathcal{E} \subset K$ which has the maximal volume among all ellipsoids contained in K. Such ellipsoid $\mathcal{E}$ is called the *John ellipsoid* (of maximal volume) for K. Obviously we can always find a position of K in which $\mathcal{E} = B_2^n$. It turns out that such a position of K is not arbitrary due to the existence of special *contact points* provided by the famous John theorem (see [John]).

5.9 Theorem (John)**.** *Let K be a symmetric convex body in $\mathbb{R}^n$ such that B_2^n is its John's ellipsoid. Then there exist numbers $c_1, \ldots, c_m > 0$ and contact points $u_1, \ldots, u_m$, i.e. $|u_i| = \|u_i\|_K = 1$, such that*

$$\mathrm{id} = \sum_{i=1}^{m} c_i u_i \cdot u_i^T. \tag{5.4}$$

Moreover, $m \le n^2 + 1$.

Proof. Let $E = (\mathbb{R}^n, \|\cdot\|_K)$. Consider on $\mathcal{L}(\ell_2^n, E)$ the operator norm $\alpha(u) = \|u\colon \ell_2^n \longrightarrow E\|$. In Proposition 5.5 we have seen that $\alpha^* = \|\cdot\|_\wedge$. Note that for any $u \in \mathcal{L}(\ell_2^n, E)$ the condition $\alpha(u) \le 1$ is equivalent to $u(B_2^n) \subset K$. Since $u(B_2^n)$ is an ellipsoid and B_2^n is the ellipsoid of maximal volume contained in K for such u we may write that $|\det u| \cdot |B_2^n| = |u(B_2^n)| \le |B_2^n|$ which gives $|\det u| \le 1$. Therefore,

$$\max_{\alpha(u)\le 1} \det u \le 1$$

with equality for $u = \mathrm{id}$. Theorem 5.6 provides an operator $v \in \mathcal{L}(\ell_2^n, E)$ for which $\alpha(v) = 1$, $\alpha^*(v^{-1}) = n$. By its proof and the above observation we might take $v = \mathrm{id}_{\ell_2^n \to E}$. The condition $\|v^{-1}\|_\wedge = \alpha^*(v^{-1}) = n$ means that there are m points $y_j^* \in E^*$, $x_j \in \ell_2^n$, with $m \le n^2 + 1$ by Caratheodory's theorem, such that

$$v^{-1} = \sum_{i=1}^m y_i^* \otimes x_i$$

and

$$\sum_{i=1}^m \|y_i^*\|_{E^*} \cdot |x_i| = n.$$

We normalize $y_i^* = \|y_i^*\|_{E^*} v_i^*$, $x_i = |x_i| u_i$, put $c_i = \|y_i^*\|_{E^*} \cdot |x_i|$ and get

$$v^{-1} = \sum_{i=1}^m c_i v_i^* \otimes u_i, \qquad \text{with} \sum_{i=1}^m c_i = n.$$

Let us now show how vectors v_i^* relate to u_i. Observe that $B_2^n \supset K^\circ$ implies $|\cdot| \le \|\cdot\|_{E^*}$. So by the Cauchy-Schwarz inequality $\mathrm{tr}\, v_i^* \otimes u_i = \langle v_i^*, u_i \rangle \le |v_i^*| \cdot |u_i| \le \|v_i^*\|_{E^*} \cdot |u_i| = 1$. Thus,

$$n = \mathrm{tr}\, v v^{-1} = \sum c_i \,\mathrm{tr}\, v_i^* \otimes u_i \le \sum c_i = n.$$

In particular, there must be equality in the Cauchy-Schwarz inequality. This yields $v_i^* = \lambda_i u_i$. By $1 = \langle v_i^*, u_i \rangle$ we infer that necessarily $\lambda_i = 1$. Thus $v_i^* = u_i$, so $\|u_i\|_{E^*} = |u_i| = 1$ which implies that $\|u_i\|_K = 1$.

Taking into account that actually $v = \mathrm{id}$ we have obtained

$$\mathrm{id} = \sum_{i=1}^m c_i u_i \cdot u_i^T,$$

where $|u_i| = 1 = \|u_i\|_K$ and $\sum c_i = n$. ☐

One immediate consequence concerns distances between linear spaces of prescribed dimension.

5.10 Corollary. *For every n dimensional real vector space E we have $\rho(E, \ell_2^n) \le \sqrt{n}$.*

Proof. Let $E = (\mathbb{R}^n, \|\cdot\|)$ and $K = \{x \in \mathbb{R}^n,\ \|x\| \le 1\}$. We pick a linear isomorphism T of $\mathbb{R}^n$ such that B_2^n is the John's ellipsoid of the body $K' = TK$. Therefore it suffices to show that

$$K' \subset \sqrt{n} B_2^n.$$

By the John theorem there are numbers $c_j > 0$ and contact points $u_j \in \mathbb{R}^n$ such that (5.4) holds. Take $x \in K'$ and observe that

$$|x|_2^2 = \langle x, x\rangle = \left\langle \sum c_j \langle u_j, x\rangle u_j, x\right\rangle = \sum c_j \langle u_j, x\rangle^2.$$

Yet, $\|u_j\|_{K'^\circ} = 1$, so $|\langle u_j, x\rangle| \le 1$ which yields

$$|x|_2^2 \le \sum c_j = n.$$

□

By the triangle inequality we can conclude an estimate of the distance between two arbitrary spaces.

5.11 Corollary. *For any two n dimensional real vector spaces E and F we have $\rho(E, F) \le n$.*

We have seen that ℓ_1^n and ℓ_∞^n are in the distance of order $\sqrt{n}$. Intuitively, these two spaces are the extreme cases. Therefore it is reasonable to ask whether $\rho(E, F) \le n$ is a rough estimate? Unexpectedly, it is not at all as it was shown by E. Gluskin [Glu].

5.12 Theorem (Gluskin '81)**.** *There exist n dimensional real vector spaces E and F such that $\rho(E, F) \ge cn$, where c is a universal constant.*

6 Lecture V — Volumes of sections of convex bodies

Studies on volumes of sections of the cube find applications in number theory. The very first link between these two fields is Minkowski's theorem.

6.1 Theorem (Minkowski). *If $K \subset \mathbb{R}^n$ is a symmetric convex body with volume $|K| \leq 2^n$, then it contains a nonzero lattice point (i.e. a point with all integer coordinates).*

6.2 Corollary. *Let $A = [a_{ij}]_{i,j=1,\dots,n}$ be a matrix such that $|\det A| \leq 1$. Then there is a nonzero lattice point $z = (z_1, \dots, z_n)$ such that $-1 \leq \sum_j a_{ij} z_j \leq 1$ for every i.*

Proof. The assertion $|\sum_j a_{ij} z_j| \leq 1$ means that $Az \in B_\infty^n$, or $z \in A^{-1} B_\infty^n$. We finish by Minkowski's theorem as

$$|A^{-1} B_\infty^n| = \frac{1}{|\det A|} |B_\infty^n| \geq 2^n.$$

□

It was A. Good who posed more sophisticated question about sufficient conditions which would guarantee an existence of integer solutions in $\mathbb{R}^n$ to a system of m inequalities ($m \geq n$).

6.3 Conjecture (Good). Let $A = [a_{ij}]_{i \leq m, j \leq n}$ be a matrix such that $\det A^T A \leq 1$. Then there is a nonzero lattice point $z = (z_1, \dots, z_n)$ such that $-1 \leq \sum_{j=1}^n a_{ij} z_j \leq 1$ for every i.

Again, by Minkowski's theorem it is enough to prove that under the hypothesis the convex body

$$K = \left\{ x \in \mathbb{R}^n, \ \left| \sum_{j=1}^n a_{ij} x_j \right| \leq 1, \ i = 1, \dots, m \right\}$$

has volume greater or equal to 2^n. Notice that $K = B_\infty^m \cap \mathrm{im} A$. This gives rise to the question of estimations of volumes of sections of the cube. The answer was given by J. Vaaler [Vaa].

6.4 Theorem (Vaaler). *For every k dimensional vector subspace E of $\mathbb{R}^n$ we have*

$$|E \cap B_\infty^n| \geq 2^k = |B_\infty^k|. \tag{6.1}$$

The case of $k = n-1$ was solved by D. Hensley (see [Hen]) slightly before the result of J. Vaaler appeared. Natural question of maximal sections arises as well. D. Hensley established that for $k = n-1$ we have $|E \cap B_\infty^n| \leq 5 \cdot |B_\infty^{n-1}|$, which is not optimal. Due to K. Ball [Ball2] we know the sharp inequality with the constant $\sqrt{2}$ instead of 5. Later on M. Mayer and A. Pajor (see [MP]) dealt with other p-balls generalizing Vaaler's result.

In the following subsections we will collect a few tools and then provide proofs of two theorems which altogether give sharp bounds on volumes of $n-1$ dimensional sections of the cube.

6.5 Theorem (Meyer-Pajor)**.** *Given k dimensional linear subspace E of $\mathbb{R}^n$ the function*

$$p \longmapsto \frac{|E \cap B_p^n|}{|B_p^k|}, \qquad p \geq 1$$

is increasing.

6.6 Corollary. *Given k dimensional linear subspace E of $\mathbb{R}^n$ the following inequalities hold*

$$\begin{aligned} |E \cap B_p^n| &\geq |B_p^k|, \qquad p \geq 2, \\ |E \cap B_p^n| &\leq |B_p^k|, \qquad p \in [1, 2). \end{aligned} \tag{6.2}$$

Proof. Notice that any section of a Euclidean ball is again a ball and as a result $|E \cap B_2^n| = |B_2^k|$. Thus the assertion easily follows by the theorem. $\square$

6.7 Theorem (Ball)**.** *For every $n-1$ dimensional linear subspace H in $\mathbb{R}^n$ we have*

$$|H \cap B_\infty^n| \leq \sqrt{2}|B_\infty^{n-1}|. \tag{6.3}$$

The equality is attained for $H = (1/\sqrt{2}, 1/\sqrt{2}, 0, \ldots, 0)^\perp$.

Almost all the material presented here comes from the excellent notes [Ball3] by K. Ball.

6.1 The comparison method

We define an order $\succ$ on the set of all measures on $\mathbb{R}^n$ saying that $\mu \succ \nu$, μ is *more peaked* than ν, if for every convex and symmetric (with respect to the origin) subset A of $\mathbb{R}^n$ we have $\mu(A) \geq \nu(A)$. This definition lies in the heart of the so-called comparison method, which has been thoroughly studied in [Kan] by M. Kanter. The significance of this notion is subject to its tensorization property.

6.8 Theorem (Kanter)**.** *Let μ_1, ν_1 be two log-concave measures on $\mathbb{R}^n$ which are symmetric, i.e. measures of sets A and $-A$ are the same. Let μ_2, ν_2 be two symmetric and log-concave measures on $\mathbb{R}^m$. If $\mu_i \succ \nu_i$, $i = 1, 2$, then*

$$\mu_1 \otimes \mu_2 \succ \nu_1 \otimes \nu_2.$$

6.9 Lemma. *Given two measures μ, ν on $\mathbb{R}^n$ such that μ is more peaked that ν for every even log-concave function f on $\mathbb{R}^n$ we have*

$$\int f \,\mathrm{d}\mu \geq \int f \,\mathrm{d}\nu.$$

Proof. It follows by integration by parts (f is nonnegative) and the definition as $\int f \,\mathrm{d}\mu = \int_0^\infty \mu(f \geq t) \,\mathrm{d}t$ and the set $\{f \geq t\}$ is convex (f is log-concave) and symmetric (f is even). $\square$

Proof of Theorem 6.8. Let $C \subset \mathbb{R}^n \times \mathbb{R}^m$ be convex and symmetric. By Fubini's theorem we get

$$\mu_1 \otimes \mu_2(C) = \int \mu_2(C_{x_1}) \,\mathrm{d}\mu_1(x_1),$$

where $C_{x_1} = \{x_2 \in \mathbb{R}^m,\ (x_1, x_2) \in C\}$ is a section of the set C. Consider the function $f(x_1) = \mu_2(C_{x_1})$. By symmetry of C we have $C_{-x_1} = -C_{x_1}$ so by symmetry of μ_2 we infer that f is even. Moreover, f is log-concave as convexity of C implies that $C_{\lambda x + (1-\lambda)y} \supset \lambda C_x + (1-\lambda)C_y$ and recall that μ_2 is log-concave. Thus by Lemma 6.9 applied to $\mu_1 \succ \nu_1$ and f

$$\mu_1 \otimes \mu_2(C) \geq \int f(x_1) \,\mathrm{d}\nu_1(x_1) = \nu_1 \otimes \mu_2(C).$$

Following the same line of reasoning yet for measures $\mu_2 \succ \nu_2$ we find that $\nu_1 \otimes \mu_2(C) \geq \nu_1 \otimes \nu_2(C)$. $\square$

Therefore to establish that one product measure is more peaked than the other one it suffices to consider their one dimensional factors. The point is that the one dimensional situation can e easily understood.

6.10 Remark. Suppose μ and ν are symmetric measures on $\mathbb{R}$ with densities f and g respectively. Then μ is more peaked that ν iff for every $a > 0$ we have $\int_0^a f(t) \,\mathrm{d}t > \int_0^a g(t) \,\mathrm{d}t$.

6.2 Expressions for volumes of convex bodies

Assume K is a symmetric convex body in $\mathbb{R}^n$ with the associated norm $\|\cdot\|_K$. Then for $p > 0$

$$\begin{aligned}\int_{\mathbb{R}^n} \exp(-\|x\|_K^p)\,\mathrm{d}x &= \int_{\mathbb{R}^n}\int_0^\infty e^{-t}\mathbf{1}_{\{t>\|x\|_K^p\}}\,\mathrm{d}t\,\mathrm{d}x \\ &= \int_0^\infty \int_{\|x\|_K<t^{1/p}} e^{-t}\,\mathrm{d}x\,\mathrm{d}t = \int_0^\infty e^{-t}|t^{1/p}K|\,\mathrm{d}t \\ &= |K|\int_0^\infty t^{n/p}e^{-t}\,\mathrm{d}t = |K|\Gamma\left(1+\frac{n}{p}\right).\end{aligned}$$

Actually we have proved

6.11 Proposition. *For a symmetric convex body in $\mathbb{R}^n$ and $p > 0$*

$$|K| = \frac{1}{\Gamma\left(1+\frac{n}{p}\right)}\int_{\mathbb{R}^n} \exp(-\|x\|_K^p)\,\mathrm{d}x. \tag{6.4}$$

6.12 Corollary. $|B_p^n| = \frac{\left[2\Gamma(1+1/p)\right]^n}{\Gamma(1+n/p)}$.

Now we use our formula to derive volume of sections of a convex body.

6.13 Proposition. *Suppose K is a symmetric convex body in $\mathbb{R}^n$ and E is a k dimensional linear subspace of $\mathbb{R}^n$. Then for $p > 0$*

$$\Gamma\left(1+\frac{k}{p}\right)|K\cap E| = \lim_{\epsilon\to 0}\frac{1}{\epsilon^{n-k}}\int_{d_\infty(x,E)\le\epsilon} \exp\big(-\|x\|_K^p\big)\,\mathrm{d}x. \tag{6.5}$$

Here $d_\infty(x,E)\le\epsilon$ means that for all $j\ge k+1$ there holds $|\langle \mathrm{proj}_{E^\perp}(x), e_j\rangle| \le \epsilon/2$, where $(e_1,\ldots,e_k)$ is a fixed basis of E and $(e_{k+1},\ldots,e_n)$ is a fixed basis of $E^\perp$.

Proof. Let $x=(u,v)$ with $u\in E$, $v\in E^\perp$. We change variables $v=\epsilon w$ and write

$$\begin{aligned}&\frac{1}{\epsilon^{n-k}}\int_{d_\infty(x,E)\le\epsilon}\exp\big(-\|x\|_K^p\big)\,\mathrm{d}x \\ &= \frac{1}{\epsilon^{n-k}}\int_{\substack{u\in E, w\in E^\perp \\ |\langle w,e_j\rangle|\le 1/2, j>k}} \exp\big(-\|(u,0)+\epsilon(0,w)\|_K^p\big)\,\mathrm{d}u\,\mathrm{d}(\epsilon w) \\ &= \int \exp\big(-\|(u,0)+\epsilon(0,w)\|_K^p\big)\,\mathrm{d}u\,\mathrm{d}w.\end{aligned}$$

Taking the limit $\epsilon \to 0$, by virtue of the fact that $|\{w \in E^\perp,\ |\langle w, e_j\rangle| \le 1/2, j > k\}| = 1$ we obtain

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{1}{\epsilon^{n-k}}&\int_{d_\infty(x,E)\le\epsilon}\exp\big(-\|x\|_K^p\big)\,\mathrm{d}x\\ =&\int_{\substack{u\in E, w\in E^\perp\\ |\langle w,e_j\rangle|\le 1/2, j>k}}\exp\big(-\|(u,0)\|_K^p\big)\,\mathrm{d}u\,\mathrm{d}w\\ =&\int_{u\in E}\exp\big(-\|u\|_{K\cap E}^p\big)\,\mathrm{d}u\\ =&\Gamma\left(1+\frac{k}{p}\right)|K\cap E|.\end{aligned}$$

□

6.3 Pólya's formula for volumes of sections of the cube

Consider a unit vector $v \in S^{n-1}$. It defines the hyperplane $v^\perp$ which can be moved along v producing an affine subspace $tv + v^\perp$ of codimension one in $\mathbb{R}^n$, where $t \in \mathbb{R}$. We are interested in volumes of sections of the cube B_∞^n with such subspaces. We define the relevant function

$$A_v(t) = |B_\infty^n \cap (tv + v^\perp)|. \tag{6.6}$$

In the proof of Ball's theorem 6.7 we will need the following formula, which is due to G. Pólya

$$A_v(t) = \frac{2^n}{\pi}\int_0^\infty \cos(tr)\prod_{k=1}^n \frac{\sin(v_k r)}{v_k r}\,\mathrm{d}r, \qquad t \in \mathbb{R}, v \in S^{n-1}. \tag{6.7}$$

For the proof let us consider two cases depending on integrability of the integrated function.

Case I (not in L_1). If all but one coordinates of v are zero, say $v_1 = 1$ and $v_k = 0$, $k > 1$, then obviously

$$A_v(t) = \begin{cases} 0, & \text{if } |t| > 1 \\ 2^{n-1}, & \text{if } |t| \le 1 \end{cases}.$$

On the other hand, suppose $t > 0$ (we do not loose generality as cos is an even function) and observe that

$$\int_0^\infty \cos(tr)\frac{\sin(vr)}{vr}\,\mathrm{d}r = \int_0^\infty \frac{\sin((t+1)r) - \sin((t-1)r)}{2r}\,\mathrm{d}r.$$

Yet for any $a > 0$ simple change of variables yields $\int_0^\infty \frac{sin(ar)}{r}\,\mathrm{d}r = \int_0^\infty \frac{\sin r}{r}\,\mathrm{d}r = \pi/2$, thus

$$\int_0^\infty \cos(tr)\frac{\sin(vr)}{vr}\,\mathrm{d}r = \begin{cases} 0, & \text{if } t > 1 \\ \pi/2, & \text{if } t < 1 \end{cases}$$

which agrees with (6.7) (what's up when $t = 1$?).

Case II (in L_1). Assume at least 2 coordinates of v are non-zero. Our function A_v is compactly supported hence we may calculate its Fourier transform (we use $tv + v^\perp = \{x \in \mathbb{R}^n,\ \langle v, x\rangle = t\}$)

$$\begin{aligned}
\widehat{A_v}(r) &= \int_{\mathbb{R}} A_v(t)e^{-itr}\,\mathrm{d}t = \int_{\mathbb{R}}\left(\int_{\langle v,x\rangle = t} \mathbf{1}_{B_\infty^n}(x)\,\mathrm{d}x\right)e^{-itr}\,\mathrm{d}r \\
&= \int_{\mathbb{R}^n}\prod_{k=1}^n \mathbf{1}_{[-1,1]}(x_k)e^{-i\langle v,x\rangle r}\,\mathrm{d}x = \prod_{k=1}^n \int_{-1}^1 e^{-irx_kv_k}\,\mathrm{d}x_k \\
&= \prod_{k=1}^n \frac{2\sin(rv_k)}{rv_k}.
\end{aligned}$$

Now we can see by the assumption that $\widehat{A_v} \in L_1$, so using the inverse Fourier transform we get the desired formula

$$\begin{aligned}
2\pi A_v(t) = \widehat{\widehat{A_v}}(t) &= 2^n \int_{\mathbb{R}} e^{-itr}\prod_{k=1}^n \frac{\sin(rv_k)}{rv_k}\,\mathrm{d}r \\
&= 2^n \int_{\mathbb{R}} \cos(tr)\prod_{k=1}^n \frac{\sin(rv_k)}{rv_k}\,\mathrm{d}r.
\end{aligned}$$

6.4 Proof of Meyer and Pajor's theorem

Take $p \geq 1$ and introduce the normalization constant $\alpha_p = \big(2\Gamma(1 + 1/p)\big)^p$ so that

$$\int_{\mathbb{R}} e^{-\alpha_p|t|^p}\,\mathrm{d}t = 1.$$

For $K = B_p^n$ we have $\|x\|_K^p = \sum_{i=1}^n |x_i|^p$, hence applying Proposition 6.13 we obtain

$$\Gamma\left(1 + \frac{k}{p}\right)|B_p^n \cap E| = \lim_{\epsilon\to 0}\frac{1}{\epsilon^{n-k}}\int_{d_\infty(x,E)\leq\epsilon} \exp\big(-\sum|x_i|^p\big)\,\mathrm{d}x.$$

We change variables putting $x_i = \alpha_p^{1/p} y_i$, $\eta = \epsilon/\alpha_p^{1/p}$ and get

$$\Gamma\left(1+\frac{k}{p}\right)|B_p^n \cap E| = \alpha_p^{k/p} \lim_{\eta\to 0}\frac{1}{\eta^{n-k}}\int_{d_\infty(y,E)\le\eta} \exp\big(-\alpha_p \sum |y_i|^p\big)\,\mathrm{d}y.$$

In view of Corollary 6.12 $\alpha_p^{k/p}/\Gamma(1+k/p) = |B_p^k|$, thus we find a nice formula for the function we are interested in

$$\begin{aligned}\frac{|B_p^n \cap E|}{|B_p^k|} &= \lim_{\eta\to 0}\frac{1}{\eta^{n-k}}\int_{d_\infty(y,E)\le\eta} \exp\big(-\alpha_p \sum |y_i|^p\big)\,\mathrm{d}y \\ &= \lim_{\eta\to 0}\frac{1}{\eta^{n-k}}\mu_p^{\otimes n}\big(\{y\in\mathbb{R}^n,\ d_\infty(y,E)\le\eta\}\big),\end{aligned}$$

where $\mu_p^{\otimes n}$ is the product measure of n copies of the probability measure μ_p on $\mathbb{R}$ with the density $t\mapsto e^{-\alpha_p|t|^p}$. It is left to the Reader to check that μ_p is more peaked than μ_q for $p>q$ (evoke Remark 6.10). The measure μ_p is symmetric and for $p\ge 1$ it is log-concave, so Theorem 6.8 yields that $\mu_p^{\otimes n}$ is more peaked than $\mu_q^{\otimes n}$ for $p>q$. Then, by definition,

$$\mu_p^{\otimes n}\big(\{y\in\mathbb{R}^n,\ d_\infty(y,E)\le\eta\}\big) \ge \mu_p^{\otimes n}\big(\{y\in\mathbb{R}^n,\ d_\infty(y,E)\le\eta\}\big),$$

as the set $\{y\in\mathbb{R}^n,\ d_\infty(y,E)\le\eta\}$ is convex and symmetric. This finishes the proof.

6.5 Proof of Ball's theorem

We fix a unit vector $v\in S^{n-1}$ and consider the section of the cube B_∞^n with the hyperspace $v^\perp$. By definition (6.6) and Póya's formula (6.7) we may write

$$\frac{|B_\infty^n\cap v^\perp|}{|B_\infty^{n-1}|} = \frac{A_v(0)}{2^{n-1}} = \frac{1}{\pi}\int_{\mathbb{R}}\prod_{k=1}^n \frac{\sin(rv_k)}{rv_k}\,\mathrm{d}r,$$

whence our goal is to prove that the last quantity does not exceed $\sqrt{2}$. Taking into account the symmetry of the problem we might assume without loss of generality that $v_i\ge 0$ for all $i\le n$.

First we exclude the easy case when v possesses a big coordinate, i.e. suppose at least one coordinate of v is grater than $1/\sqrt{2}$, say $v_1 > 1/\sqrt{2}$. We roughly estimate $|B_\infty^n\cap v^\perp| \le |C\cap v^\perp|$ where C is a cylinder $\{x\in\mathbb{R}^n,\ |x_i|\le 1,\ i>1\}$. Let T denote the orthogonal projection onto the hyperspace $\{x_1=0\}$. Then clearly $T(C\cap v^\perp) = B_\infty^{n-1}$, so

$$|C\cap v^\perp| = \frac{1}{\cos\theta}|T(C\cap v^\perp)| = \frac{1}{\cos\theta}|B_\infty^{n-1}|.$$

The angle θ between $v^\perp$ and $\{x_1 = 0\}$ can be computed as follows. Denote by S the orthogonal projection on $v^\perp$. Then $S(e_1) = (1 - v_1^2, -v_1 v_2, \dots, -v_1 v_n)$ and

$$\frac{1}{\cos\theta} = \frac{|S(e_1)|}{|TS(e_1)|} = \frac{\left((1-v_1^2)^2 + v_1^2 \sum_{i>1} v_i^2\right)^{1/2}}{\left(v_1^2 \sum_{i>1} v_1^2\right)^{1/2}} = \frac{1}{v_1} < \sqrt{2},$$

for $\sum_i v_i^2 = 1$. Finally we get

$$\frac{|B_\infty^n \cap v^\perp|}{|B_\infty^{n-1}|} \le \frac{1}{\cos\theta} < \sqrt{2}.$$

Now we assume that all coordinates are small, i.e. $v_i \le 1/\sqrt{2}$, $i \le n$. Take the weights $p_k = 1/v_k^2$ so that $\sum \frac{1}{p_k} = 1$ and $p_k \ge 2$. Note that by virtue of Hölder's inequality

$$\begin{aligned}\frac{1}{\pi}\int_{\mathbb{R}} \prod_{k=1}^n \frac{\sin(rv_k)}{rv_k}\,\mathrm{d}r &\le \frac{1}{\pi}\prod_{k=1}^n \left(\int_{\mathbb{R}} \left|\frac{\sin(rv_k)}{rv_k}\right|^{p_k} \mathrm{d}r\right)^{1/p_k} \\ &= \prod_{k=1}^n \left(\frac{1}{\pi v_k}\int_{\mathbb{R}} \left|\frac{\sin(r)}{r}\right|^{p_k} \mathrm{d}r\right)^{1/p_k}.\end{aligned}$$

A miracle happens thanks to the following lemma.

6.14 Lemma. *For* $p \ge 2$

$$\frac{1}{\pi}\int_{\mathbb{R}} \left|\frac{\sin r}{r}\right|^p \mathrm{d}r \le \frac{\sqrt{2}}{\sqrt{p}}, \tag{6.8}$$

with equality iff $p = 2$.

(We skip the proof because it is involved and technical. Consult [Ball2].) Indeed,

$$\prod_{k=1}^n \left(\frac{1}{\pi v_k}\int_{\mathbb{R}} \left|\frac{\sin(r)}{r}\right|^{p_k} \mathrm{d}r\right)^{1/p_k} \le \prod_{k=1}^n \left(\frac{\sqrt{2}}{v_k\sqrt{p_k}}\right)^{1/p_k} = \sqrt{2},$$

since by definition $v_k\sqrt{p_k} = 1$. Observe that for the equality we need that all nontrivial weights amount to 2, i.e. $v_i = v_j = 1/\sqrt{2}$ for some $i \ne j$ and $v_k = 0$ for $k \ne i, j$. This completes the proof.

6.15 Remark. For sections with subspaces of lower dimension than $n-1$ one repeat all the arguments, yet instead of using Hölder's inequality, the Brascamp-Lieb inequality could be applied.

6.6 Particular case of The Gaussian correlation conjecture

At the end of this section we discuss *ad hoc* a particular case of the Gaussian correlation conjecture which can be easily established exploiting the comparison method.

Is it true that for any symmetric convex bodies K and L in $\mathbb{R}^n$

$$\gamma_n(K \cap L) \geq \gamma_n(K)\gamma_n(L)? \tag{6.9}$$

This is the well-known *Gaussian correlation conjecture.* Apart being quite old as formulated in 1972 in [GEO] (see therein for a detailed history), this question still remains open. However, there are lots of partial results. C. Khatri and independently Z. Sidak proved in 1967 that the question has an affirmative answer if one of the sets is a strip, i.e. a set of the form $\{x \in \mathbb{R}^n, |\langle x, v\rangle| \leq 1\}$ for some $v \in \mathbb{R}^n$ (see [Kha] and [Sid]). L. Pitt showed in 1977 that it is true for $n = 2$ (see [Pit]). G. Schechtman, T. Schlumprecht and J. Zinn in 1998 (see [SSZ]) gave a few results, e.g. the inequality holds for two ellipsoids, or there is a constant c such that it holds whenever K and L are contained in the ball with radius $c\sqrt{n}$. Finally, in 1999, G. Hargé [Har] refined this and showed that the conjecture is true in the case when one of the bodies is an ellipsoid.

Here we recover in a striking way the result of C. Khatri and Z. Sidak. The proof is due to A. Giannopoulos (his private communication with K. Ball).

6.16 Theorem. *Let K be a symmetric convex set in $\mathbb{R}^n$ and let $S = \{x \in \mathbb{R}^n,\ |\langle x, u\rangle| \leq t\}$ be a strip, where $u \in S^{n-1}$ and $t > 0$. Then*

$$\gamma_n(K \cap S) \geq \gamma_n(K)\gamma_n(S).$$

Proof. Since the Gaussian measure is rotationally invariant we may assume that $S = \{x \in \mathbb{R}^n,\ |x_1| \leq t\}$. We would like to prove that

$$\frac{\gamma_n(K \cap S)}{\gamma_n(S)} \geq \gamma_n(K).$$

Thus we define the probability measure $\mu(A) = \gamma_n(A \cap S)/\gamma_n(S)$ and now our goal is to show that μ is more peaked that γ_n. By virtue of Lemma 6.9 it is reduced to one dimensional problem as $\gamma_n = \gamma_1 \otimes \gamma_{n-1}$ and $\mu = \nu \otimes \gamma_{n-1}$, where ν has the density proportional to $e^{-s^2/2}\mathbf{1}_{[-t,t]}(s)$. Indeed, we now only need to check that ν is more peaked than γ_1. For this purpose, due to Remark 6.10, we would like to verify that for any $a > 0$

$$\frac{1}{\int_0^t e^{-s^2/2}\,\mathrm{d}s}\int_0^a e^{-s^2/2}\mathbf{1}_{[0,t]}(s)\,\mathrm{d}s \geq \frac{1}{\int_0^\infty e^{-s^2/2}\,\mathrm{d}s}\int_0^a e^{-s^2/2}\,\mathrm{d}s.$$

Yet this is obvious. $\square$

7 Lecture VI — Volume estimates of entropy numbers and embeddings into ℓ_∞^N

The main question which is addressed here concerns the embeddings of finite dimensional normed vector spaces into ℓ_∞^N. This is a sort of introduction to celebrated Dvoretzky's theorem. In the first subsection we present the so-called volume estimates which we will exploit in the second subsection where we provide a relevant theorem regarding embeddings.

7.1 Volume estimates of entropy numbers

Let (X, d) be a metric space. Given a subset T of X we say that a collection of points $\{x_i\}_{i=1,\dots,N} \subset T$ is an *ϵ-net* of T if $T \subset \bigcup_{i=1}^N B(x_i, \epsilon)$, where $B(a, r) = \{x \in X,\ d(x, a) < r\}$ is an open ball in X. Such a collection is called an *ϵ-separated set* in T if $d(x_i, x_j) \geq \epsilon$ for any $i \neq j$. The cardinality of a minimal ϵ-net of T is called the *covering number* and denoted by $N(T, d, \epsilon)$ while the cardinality of a maximal ϵ-separated set in T is called the *packing number* and denoted by $K(T, d, \epsilon)$. The logarithm of the packing and covering numbers are sometimes referred to as *entropy numbers.* There is a link between them.

7.1 Proposition. $N(T, d, \epsilon) \leq K(T, d, \epsilon) \leq N(T, d, \epsilon/2)$.

Proof. For the proof of the first inequality suppose $S = \{x_1, \dots, x_N\} \subset T$ is a maximal ϵ-separated set in T. Take any but different from $x_1, \dots, x_N$ point $x \in T$. Then there is i such that $d(x, x_i) < \epsilon$ as otherwise the set $S \cup \{x\}$ would be also ϵ-separated which contradicts maximality. Therefore $T \subset \bigcup_{i=1}^N B(x_i, \epsilon)$, i.e. S is an ϵ-net. Thus $N(T, d, \epsilon) \leq N = K(T, d, \epsilon)$.

For the proof of the second inequality let us consider an $\epsilon/2$-net $\{x_1, \dots, x_N\}$ and an ϵ-separated set $\{y_1, \dots, y_K\}$. We prove that $K \leq N$. For each $i \in \{1, \dots, K\}$ there is $f(i) \in \{1, \dots, N\}$ such that $d(x_i, y_{f(i)}) < \epsilon/2$ since $\{x_i\}_{i=1,\dots,N}$ is an $\epsilon/2$-net. This defines a function $\{1, \dots, K\} \underset{f}{\to} \{1, \dots, N\}$. It is one-to-one and consequently $K \leq N$. Indeed, if there were $i \neq j$ such that $f(i) = f(j)$ then we would get

$$\epsilon \leq d(x_i, x_j) \leq d(x_i, y_{f(i)}) + d(y_{f(j)}, x_j) < \epsilon/2 + \epsilon/2 = \epsilon$$

which is a contradiction. $\square$

In particular, for two convex bodies A, B in $\mathbb{R}^n$ we define

$$N(A, \epsilon B) = N(A, \|\cdot\|_B, \epsilon) = \inf\{N,\ \exists\, x_1, \ldots, x_N\ A \subset \bigcup_{i=1}^{N}(x_i + \epsilon B)\}. \tag{7.1}$$

The question is how to estimate this covering number. The notion of volume will help.

7.2 Lemma. *For any convex bodies A, B in $\mathbb{R}^n$ and $\epsilon > 0$*

$$\frac{1}{\epsilon^n}\frac{|A|}{|B|} \le N(A, \epsilon B). \tag{7.2}$$

Proof. This is obvious as $|A| \le \left|\bigcup_{i=1}^{N}(x_i + \epsilon B)\right| \le N\epsilon^n|B|$. □

7.3 Lemma (Volume estimate of the entropy number)**.** *For any convex bodies A, B in $\mathbb{R}^n$ such that $B \subset A$ and $\epsilon > 0$*

$$N(A, \epsilon B) \le \left(1 + \frac{2}{\epsilon}\right)^n \frac{|A|}{|B|}. \tag{7.3}$$

Proof. Let $\{x_1, \ldots, x_N\}$ be an ϵ-separated set in A with respect to the metric given by the norm $\|\cdot\|_B$. By Proposition 7.1 we know that it is enough to estimate N. Since $B \subset A$, $A + \frac{\epsilon}{2}B \supset \bigcup_{i=1}^{N}(x_i + \frac{\epsilon}{2}B)$ and these balls are mutually disjoint, we obtain

$$\left(1 + \frac{\epsilon}{2}\right)^n |A| \ge |A + \epsilon/2B| \ge N|(\epsilon/2)B| = N(\epsilon/2)^n|B|.$$

□

7.4 Corollary. *For a convex body A in $\mathbb{R}^n$ and $\epsilon > 0$*

$$\frac{1}{\epsilon^n} \le N(A, \epsilon A) \le \left(1 + \frac{2}{\epsilon}\right)^n. \tag{7.4}$$

7.2 Embeddings into ℓ_∞^N

Now we will try to figure out the geometrical meaning of the volume estimate (7.4). It says that for a convex body A in $\mathbb{R}^n$ and a positive number ϵ there is an ϵ-net Λ of A with $\sharp\Lambda \le (1 + 2/\epsilon)^n$. It means that $A \subset \Lambda + \epsilon A$, or, in different words, using support functions, $h_A \le h_{\Lambda+\epsilon A} = h_\Lambda + h_{\epsilon A} = h_\Lambda + \epsilon h_A$. Moreover, $\Lambda \subset A$, so $h_\Lambda \le h_A$. Altogether,

$$h_\Lambda \le h_A \le \frac{1}{1-\epsilon}h_\Lambda.$$

The first meaning of this inequality is that A can be approximated with a polytope $P = \operatorname{conv}\Lambda$ which has a controlled number of vertices $\sharp\Lambda$

$$P \subset A \subset \frac{1}{1-\epsilon} P.$$

The second meaning is revealed when we put $A = K^\circ$ for a symmetric convex body K, for then $h_A = \|\cdot\|_K$. Consequently, for every $x \in \mathbb{R}^n$

$$\max_{\lambda\in\Lambda} \langle x, \lambda\rangle \leq \|x\|_K \leq \frac{1}{1-\epsilon} \max_{\lambda\in\Lambda} \langle x, \lambda\rangle.$$

Therefore, defining the embedding

$$\begin{aligned} T\colon (\mathbb{R}^n, \|\cdot\|_K) &\longrightarrow (\mathbb{R}^N, |\cdot|_\infty) \\ x &\longmapsto (\langle x, \lambda\rangle)_{\lambda\in\Lambda}, \end{aligned}$$

where $N = \sharp\Lambda$, we find that

$$|Tx|_\infty \leq \|x\|_K \leq \frac{1}{1-\epsilon} |Tx|_\infty.$$

That is, the Banach-Mazur distance between the bodies K and $\mathrm{im}T \cap B_\infty^N$ is lower or equal to $\frac{1}{1-\epsilon} \approx 1+\epsilon$, with $N \leq (1+2/\epsilon)^n$. In other words, we have derived the following proposition.

7.5 Proposition. *For any n dimensional real Banach space X there is a subspace Y of ℓ_∞^N such that $\rho(X,Y) \leq 1+\epsilon$, with $N \leq (1+2/\epsilon)^n$.*

In particular, if we take for X the Euclidean space ℓ_2^n we discover that ℓ_∞^N admits subspaces of dimension $n = \frac{\ln N}{\ln(1+2/\epsilon)}$ which are almost Euclidean (ϵ-close to ℓ_2^n in the sense of the Banach-Mazur distance). Much more general phenomenon takes place, as, stated informally, Dvoretzky's theorem says that such Euclidean subspaces exist not only for ℓ_∞ spaces but for any Banach spaces.

References

[Ball1] K. Ball, Ellipsoids of maximal volume in convex bodies, *Geom. Dedicata* **41** (2) (1992), 241–250.

[Ball2] K. Ball, Cube slicing in $\mathbb{R}^n$, *Proc. Amer. Math. Soc.* **97** (1986), 465–473.

[Ball3] K. Ball, Convex geometry and functional analysis. In: Handbook of the geometry of Banach spaces. Volume 1. *Johnson, W. B. (ed.) et al., Amsterdam: Elsevier*, (2001), 161-194.

[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.

[Fr] M. Fradelizi, Concentration inequalities for s-concave measures of dilations of Borel sets and applications, *Electron. J. Probab.*, Vol. **14** (2009), 2068–2090.

[Glu] E. Gluskin, The diameter of Minkowski compactum roughly equals to n, *Funct. Anal. Appl.* **15** (1981), 57–58.

[Gro] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.* **97** (1975), 1061–1083.

[GEO] S. Das Gupta, M. Eaton, I. Olkin, M. Perlman, L. Savage, M. Sobel, Inequalities and the probability content of convex regions for elleptically contourned distributions, *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2**, Univ. California Press, Berkeley, 241–264, 1972.

[Har] G. Hargé, A particular case of the correlation inequality for the Gaussian measure, *Ann. Probab.* **27** (1999), No. 4, 1939–1951.

[Hen] D. Hensley, Slicing the cube in $\mathbb{R}^n$ and probability, *Proc. Amer. Math. Soc.* **73** (1) (1979) 95–100.

[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.

[Kan] M. Kanter, Unimodality and dominance for symmetric random vectors, *Trans. Amer. Math. Soc.* **229** (1977), 65–85.

[Kha] C. Khatri, On certain inequalities for normal distributions and their applications to simultaneous confidence bounds, *Ann. Math. Statist.* **38** (1967), 1853–1867.

[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.

[MP] M. Meyer, A. Pajor, Section of the unit ball of ℓ_p^n, *J. Funct. Anal.* **80** (1988) 109–123.

[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.

[Pit] L. Pitt, A Gaussian correlation inequality for symetric convex sets, *Ann. Probab.* **5** (1977), 470–474.

[Sid] Z. Sidak, Rectangular confidence regions for the means of multivariate normal distributions, *J. Amer. Statist. Assoc.* **62** (1967), 626–633.

[SSZ] G. Schechtman, T. Schlumprecht, J. Zinn, On the Gaussian measure of the intersection of symmetric convex sets, *Ann. Probab.* **26** (1998), 346–357.

[ST] V. Sudakov, B. Tsirelson, Extremal propertiesof half-spaces for spherically invariant measures (in Russian), *Zap. Nauchn. Sem. L.O.M.I.* **41** (1974), 14–24.

[Vaa] J. Vaaler, A geometric inequality with applications to linear forms, *Pacif. J. Math* **83**, No 2, (1979), 543–553.