CHAPTER 1

Background from asymptotic convex geometry

In this introductory chapter we survey the prerequisites from the theory of convex bodies and the asymptotic theory of finite dimensional normed spaces. Short proofs are provided for the most important results that are used in the sequel. Basic references on convex geometry are the monographs by Schneider [**463**] and Gruber [**237**]. The asymptotic theory of finite dimensional normed spaces is presented in the books by V. Milman and Schechtman [**387**], Pisier [**430**] and Tomczak-Jaegermann [**493**].

The books of Rockafellar [**442**], Bogachev [**92**], Brezis [**121**] and Feller [**169**], [**170**] are very useful sources of information on facts from convex analysis, functional analysis and probability theory that are being used throughout this book.

The first four sections of this chapter contain background material from classical convexity: the Brunn-Minkowski inequality and its functional forms, mixed volumes and classical geometric inequalities.

Section 1.5 introduces three classical positions of convex bodies: John's position, the minimal mean width position and the minimal surface area position. All of them arise as solutions of extremal problems and can be characterized as satisfying an isotropic condition with respect to an appropriate measure. This relates them to the Brascamp-Lieb inequality and its reverse. In Section 1.6 we discuss Barthe's proof of these inequalities and their applications to geometric problems; an example is K. Ball's sharp reverse isoperimetric inequality.

Section 1.7 introduces the concept of measure concentration and the main examples of metric probability spaces that will be used in this book: the sphere, the Gauss space and the discrete cube. The next two sections survey basic probabilistic tools that we will use: covering numbers and basic inequalities for them, Gaussian and sub-Gaussian processes and bounds for the expectation of their supremum.

The last sections of the chapter give a brief synopsis of the major results of asymptotic convex geometry: Dvoretzky type theorems, the notion of volume ratio and Kashin's theorem, the ℓ-position and Pisier's inequality on the Rademacher projection, the MM^*-estimate, Milman's low M^*-estimate and the quotient of subspace theorem. Finally, we present the reverse Santaló inequality and the reverse Brunn-Minkowski inequality; during this discussion M-ellipsoids and their basic properties are also introduced.

1.1. Convex bodies

We work in $\mathbb{R}^n$, which is equipped with a Euclidean structure $\langle \cdot, \cdot \rangle$. We denote by $\| \cdot \|_2$ the corresponding Euclidean norm, and write B_2^n for the Euclidean unit ball and S^{n-1} for the unit sphere. Volume is denoted by $| \cdot |$. We write ω_n for the volume of B_2^n and σ for the rotationally invariant probability measure on S^{n-1}.

We also denote the Haar measure on $O(n)$ by ν. The Grassmann manifold $G_{n,i}$ of i-dimensional subspaces of $\mathbb{R}^n$ is equipped with the Haar probability measure $\nu_{n,i}$. Let $i \leqslant n$ and $F \in G_{n,i}$. We will denote the orthogonal projection from $\mathbb{R}^n$ onto F by P_F. We also define $B_F = B_2^n \cap F$ and $S_F = S^{n-1} \cap F$.

We say that a subset A of $\mathbb{R}^n$ is convex if $(1-\lambda)x + \lambda y \in A$ for any $x, y \in A$ and any $\lambda \in [0,1]$. The Minkowski sum of two sets $A, B \subset \mathbb{R}^n$ is defined by

$$A + B = \{a + b : a \in A, b \in B\},$$

and for every $\lambda \in \mathbb{R}$ we set

$$\lambda A = \{\lambda a : a \in A\}.$$

Note that both operations preserve convexity; also, A is convex if and only if $\lambda A + (1-\lambda)A = A$ for every $\lambda \in (0,1)$. We denote by $\tilde{\mathcal{K}}_n$ the convex cone (under Minkowski addition and multiplication by nonnegative real numbers) of all non-empty, compact convex subsets of $\mathbb{R}^n$. In this book we are mainly interested in convex bodies.

1.1.1. Convex bodies

Definition 1.1.1. A *convex body* is a convex subset K of $\mathbb{R}^n$ which is compact and has non-empty interior. The class of convex bodies in $\mathbb{R}^n$ is denoted by $\mathcal{K}_n$.

We say that $K \in \mathcal{K}_n$ is a *symmetric convex body* if $x \in K$ implies that $-x \in K$. We also say that K is *centered* if the *barycenter*

$$\mathrm{bar}(K) = \frac{1}{|K|}\int_K x\,dx$$

of K is at the origin.

Definition 1.1.2. The *support function* of a convex body K (and more generally of a compact convex set) in $\mathbb{R}^n$ is defined by

$$h_K(x) = \sup\{\langle x, y\rangle : y \in K\}$$

for all $x \in \mathbb{R}^n$.

One may check that h_K is positively homogeneous and convex. Note that if K and T are two convex bodies in $\mathbb{R}^n$ then $K \subseteq T$ if and only if $h_K \leqslant h_T$. Given $u \in S^{n-1}$, the quantity $h_K(u) + h_K(-u)$ is the *width* of K in the direction of u.

A compact set K in $\mathbb{R}^n$ will be called *star-shaped* at 0 if it contains the origin in its interior and every line through 0 meets K in a line segment. For such a set, the *radial function* ρ_K is defined for all $x \neq 0$ by

$$\rho_K(x) = \max\{\lambda > 0 : \lambda x \in K\}.$$

If ρ_K is continuous then we say that K is a *star body*. The volume of a star body K can be expressed in polar coordinates as

$$|K| = \omega_n \int_{S^{n-1}} \rho_K^n(\theta)\,d\sigma(\theta).$$

Definition 1.1.3. Let K be a convex body in $\mathbb{R}^n$ with 0 in the interior of K. The *polar body* of K is the set

$$K^\circ = \Big\{y \in \mathbb{R}^n : \sup_{x \in K}\langle x, y\rangle \leqslant 1\Big\}.$$

It is easily checked that K° is a convex body, the mapping $K \mapsto K^\circ$ reverses order, and $K^{\circ\circ} = K$ for every K in the class $\mathcal{K}_n^{(0)}$ of convex bodies which contain 0 in their interior. Note that $(K \cap T)^\circ = \mathrm{conv}(K^\circ \cup T^\circ)$ for all $K, T \in \mathcal{K}_n^{(0)}$.

The natural topology on the space of convex bodies in $\mathbb{R}^n$ is induced by the Hausdorff metric δ^H: More generally, if $K, T \in \tilde{\mathcal{K}}_n$ then we define

$$\delta^H(K,T) = \max\Big\{ \max_{x\in K}\min_{y\in T} \|x-y\|_2, \max_{x\in T}\min_{y\in K} \|x-y\|_2 \Big\}.$$

Equivalently,

$$\begin{aligned}\delta^H(K,T) &= \inf\{\delta \geqslant 0 : K \subseteq T + \delta B_2^n \text{ and } T \subseteq K + \delta B_2^n\}\\ &= \max\{|h_K(u) - h_T(u)| : u \in S^{n-1}\}.\end{aligned}$$

This shows that the embedding $K \mapsto h_K$ from $\mathcal{K}_n$ to the space $C(S^{n-1})$ of continuous functions on the sphere is an isometry between $(\mathcal{K}_n, \delta^H)$ and a subset of $C(S^{n-1})$ endowed with the supremum norm. Note that this mapping is positively linear (mapping Minkowski addition to sum of functions) and order-preserving (between inclusion and point-wise inequality). The *Blaschke selection theorem* provides a very useful compactness principle.

THEOREM 1.1.4 (Blaschke). *Let $\{K_j\}$ be a sequence of compact convex sets in $\mathbb{R}^n$. Assume that there exists $R > 0$ such that $K_j \subseteq RB_2^n$ for all j. Then, $\{K_j\}$ has a subsequence which converges to some $K \in \tilde{\mathcal{K}}_n$ with respect to δ^H.*

1.1.2. Symmetric convex bodies

Let K be a symmetric convex body in $\mathbb{R}^n$. The function

$$\|x\|_K = \min\{\lambda \geqslant 0 : x \in \lambda K\}$$

is a norm on $\mathbb{R}^n$. We denote the normed space $(\mathbb{R}^n, \|\cdot\|_K)$ by X_K. Conversely, if $X = (\mathbb{R}^n, \|\cdot\|)$ is a normed space, then its unit ball $K_X = \{x \in \mathbb{R}^n : \|x\| \leqslant 1\}$ is a symmetric convex body in $\mathbb{R}^n$.

The dual norm $\|\cdot\|_*$ of $\|\cdot\|$ is defined by

$$\|y\|_* = \max\{|\langle x, y\rangle| : \|x\| \leqslant 1\}.$$

From the definition it is clear that $|\langle x, y\rangle| \leqslant \|y\|_*\|x\|$ for all $x, y \in \mathbb{R}^n$. If $X^* = (\mathbb{R}^n, \|\cdot\|_*)$ is the dual space of X, then $K_{X^*} = K_X^\circ$. We will use the notation $\|\cdot\|_{K^\circ}$ or $\|\cdot\|_*$, and $\|\cdot\|_K$ or $\|\cdot\|$.

If K and T are two convex bodies in $\mathbb{R}^n$ that contain the origin in their interior, their *geometric distance* $d_{\mathrm{G}}(K,T)$ is defined by

$$d_{\mathrm{G}}(K,T) = \inf\{ab : a, b > 0, K \subseteq bT \text{ and } T \subseteq aK\}.$$

The natural distance between two n-dimensional normed spaces X_K and X_T is the Banach-Mazur distance

$$d_{\mathrm{BM}}(X_K, X_T) = \inf_{A\in GL(n)} \|A : X_K \to X_T\|\,\|A^{-1} : X_T \to X_K\|.$$

From the definition of the geometric distance we see that

$$d_{\mathrm{BM}}(X_K, X_T) = \inf\{d_{\mathrm{G}}(K, A(T)) : A \in GL(n)\}.$$

In other words, the Banach-Mazur distance $d_{\mathrm{BM}}(X_K, X_T)$ is the smallest positive real λ for which we may find $A \in GL(n)$ such that $K \subseteq A(T) \subseteq \lambda K$. It is clear that

$d_{\mathrm{BM}}(X_K, X_T) \geqslant 1$ with equality if and only if X_K are X_T isometrically isomorphic. Note that $d_{\mathrm{BM}}(X,Z) \leqslant d_{\mathrm{BM}}(X,Y)d_{\mathrm{BM}}(Y,Z)$ for any triple of n-dimensional normed spaces.

If K and T are symmetric convex bodies in $\mathbb{R}^n$ we set $d_{\mathrm{BM}}(K,T) = d_{\mathrm{BM}}(X_K.X_T)$. The definition of the Banach-Mazur distance can be extended to the class of not necessarily symmetric convex bodies as follows: if $K, T \in \mathcal{K}_n$ then we set

$$d_{\mathrm{BM}}(K,T) = \inf\{\lambda > 0 : K - z \subseteq A(T-w) \subseteq \lambda(K-z)\},$$

where the infimum is over all $z, w \in \mathbb{R}^n$ and all $A \in GL(n)$. In the sequel, we usually denote d_{BM} simply by d. Also, the distance from an n-dimensional normed space X to ℓ_2^n will be denoted by $d_X\,(= d(X, \ell_2^n))$, and similarly we set $d_K = d(K, B_2^n)\,(= d_{\mathrm{BM}}(K, B_2^n))$.

1.2. Brunn–Minkowski inequality

The Brunn-Minkowski inequality relates Minkowski addition and volume in $\mathbb{R}^n$.

THEOREM 1.2.1 (Brunn-Minkowski). *Let K and T be two non-empty compact subsets of $\mathbb{R}^n$. Then,*

$$|K+T|^{1/n} \geqslant |K|^{1/n} + |T|^{1/n}. \tag{1.2.1}$$

Theorem 1.2.1 expresses the fact that volume is a concave function with respect to Minkowski addition. For this reason we also write it in the following form: If K and T are non-empty compact subsets of $\mathbb{R}^n$ then for every $\lambda \in (0,1)$ we have

$$|\lambda K + (1-\lambda)T|^{1/n} \geqslant \lambda |K|^{1/n} + (1-\lambda)|T|^{1/n}. \tag{1.2.2}$$

From (1.2.2) and the arithmetic-geometric means inequality we get

$$|\lambda K + (1-\lambda)T| \geqslant |K|^{\lambda} |T|^{1-\lambda}. \tag{1.2.3}$$

This form of the Brunn-Minkowski inequality has the advantage of being dimension free. In fact, one can show that it is equivalent to (1.2.1) in the sense that knowing (1.2.3) for all K, T and λ we can then show that the stronger inequality (1.2.1) holds true.

1.2.1. Brunn's principle

There are many interesting proofs of the Brunn-Minkowski inequality. The first one, in chronological order, was restricted to the class of convex bodies and it was based on Brunn's concavity principle.

THEOREM 1.2.2 (Brunn). *Let K be a convex body in $\mathbb{R}^n$ and let F be a k-dimensional subspace. Then, the function $f : F^{\perp} \to \mathbb{R}$ defined by $f(x) = |K \cap (F + x)|^{1/k}$ is concave on its support.*

For the proof of Brunn's principle we introduce *Steiner symmetrization*. For any convex body K in $\mathbb{R}^n$ and any $\theta \in S^{n-1}$ we consider the set $S_\theta(K)$ consisting of all points of the form $x + \lambda\theta$, where x is in the projection $P_{\theta^\perp}(K)$ of K onto $\theta^\perp$ and $|\lambda| \leqslant \frac{1}{2} \times \text{length}[(x + \mathbb{R}\theta) \cap K]$. In other words, we obtain $S_\theta(K)$ by sliding its chords so that their midpoint will be on $\theta^\perp$ and take the union of all resulting chords. The set $S_\theta(K)$ is the Steiner symmetrization of K in the direction of θ.

From the definition one can check a number of basic properties of Steiner symmetrization which are summarized below:

(i) Steiner symmetrization preserves convexity: if K is a convex body then $S_\theta(K)$ is also a convex body.

(ii) $S_\theta(K)$ can be described as follows:

$$S_\theta(K) = \left\{ x + \frac{t_1 - t_2}{2}\theta : x \in P_{\theta^\perp}K, x + t_1\theta \in K, x + t_2\theta \in K \right\}.$$

(iii) Steiner symmetrization preserves volume: $|S_\theta(K)| = |K|$.

(iv) If K_1 and K_2 are two convex bodies then, for every $\lambda \in (0,1)$,

$$S_\theta(\lambda K_1 + (1-\lambda)K_2) \supseteq \lambda S_\theta(K_1) + (1-\lambda)S_\theta(K_2).$$

A very useful fact is that, given a convex body K in $\mathbb{R}^n$, we can find a sequence $\{\theta_j\}$ of directions so that applying successive Steiner symmetrizations with respect to θ_j we obtain a sequence of convex bodies which converges to a Euclidean ball in the Hausdorff metric. More generally, if F is a k-dimensional subspace of $\mathbb{R}^n$, $1 \leqslant k \leqslant n$, it is a well known fact, which goes back to Steiner and Schwarz, that for every convex body K one can find a sequence of successive Steiner symmetrizations in directions $\theta \in F$ so that the limiting convex body $\tilde{K}$ has the following property: For every $x \in F^\perp$, $\tilde{K} \cap (F + x)$ is a ball with center at x and radius $r(x)$ such that $|\tilde{K} \cap (F + x)| = |K \cap (F + x)|$.

Then, the proof of Theorem 1.2.2 is easily completed: using the convexity of $\tilde{K}$ we see that the function r is concave on its support, and hence f is also concave. □

Proof of Theorem 1.2.1. Brunn's concavity principle implies the Brunn-Minkowski inequality for convex bodies as follows. If K and T are convex bodies in $\mathbb{R}^n$, we define

$$K_1 = K \times \{0\} \quad \text{and} \quad T_1 = T \times \{1\}$$

in $\mathbb{R}^{n+1}$ and consider their convex hull L. If

$$L(t) = \{x \in \mathbb{R}^n : (x,t) \in L\} \qquad (t \in [0,1])$$

we easily check that $L(0) = K$, $L(1) = T$ and

$$L(1/2) = \frac{K+T}{2}.$$

Then, Brunn's concavity principle for $F = \mathbb{R}^n$ shows that

$$\left|\frac{K+T}{2}\right|^{1/n} \geqslant \frac{1}{2}|K|^{1/n} + \frac{1}{2}|T|^{1/n},$$

and (1.2.1) is proved. □

1.2.2. Prékopa-Leindler inequality

We describe one more proof of the Brunn-Minkowski inequality, using an inequality of Prékopa and Leindler.

Theorem 1.2.3 (Prékopa-Leindler). *Let $f, g, h : \mathbb{R}^n \to \mathbb{R}^+$ be measurable functions, and let $\lambda \in (0,1)$. We assume that f and g are integrable, and for every $x, y \in \mathbb{R}^n$*

$$h(\lambda x + (1-\lambda)y) \geqslant f(x)^\lambda g(y)^{1-\lambda}. \tag{1.2.4}$$

Then,

$$\int_{\mathbb{R}^n} h \geqslant \left(\int_{\mathbb{R}^n} f\right)^{\lambda} \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda}.$$

Proof. We present a proof which uses induction on the dimension n. We first work in dimension one: we may assume that f and g are continuous and strictly positive, and we define $x, y : (0,1) \to \mathbb{R}$ by the equations

$$\int_{-\infty}^{x(t)} f = t \int_{\mathbb{R}} f \quad \text{and} \quad \int_{-\infty}^{y(t)} g = t \int_{\mathbb{R}} g.$$

In view of our assumptions, x and y are differentiable, and for every $t \in (0,1)$ we have

$$x'(t) f(x(t)) = \int_{\mathbb{R}} f \quad \text{and} \quad y'(t) g(y(t)) = \int_{\mathbb{R}} g.$$

We now define $z : (0,1) \to \mathbb{R}$ by

$$z(t) = \lambda x(t) + (1-\lambda) y(t).$$

Since x and y are strictly increasing, z is also strictly increasing, and the arithmetic-geometric means inequality shows that

$$z'(t) = \lambda x'(t) + (1-\lambda) y'(t) \geqslant (x'(t))^{\lambda} (y'(t))^{1-\lambda}.$$

Hence, we can estimate the integral of h making the change of variables $s = z(t)$, as follows:

$$\begin{aligned}
\int_{\mathbb{R}} h &= \int_0^1 h(z(t)) z'(t) dt \\
&\geqslant \int_0^1 h(\lambda x(t) + (1-\lambda) y(t)) (x'(t))^{\lambda} (y'(t))^{1-\lambda} dt \\
&\geqslant \int_0^1 f^{\lambda}(x(t)) g^{1-\lambda}(y(t)) \left(\frac{\int f}{f(x(t))}\right)^{\lambda} \left(\frac{\int g}{g(y(t))}\right)^{1-\lambda} dt \\
&= \left(\int_{\mathbb{R}} f\right)^{\lambda} \left(\int_{\mathbb{R}} g\right)^{1-\lambda}.
\end{aligned}$$

Next, we assume that $n \geqslant 2$ and the theorem has been proved in all dimensions $k \in \{1, \dots, n-1\}$. Let f, g and h be as in the theorem. For every $s \in \mathbb{R}$ we define $h_s : \mathbb{R}^{n-1} \to \mathbb{R}^+$ setting $h_s(w) = h(w, s)$, and $f_s, g_s : \mathbb{R}^{n-1} \to \mathbb{R}^+$ in an analogous way. From (1.2.4) it follows that if $x, y \in \mathbb{R}^{n-1}$ and $s_0, s_1 \in \mathbb{R}$ then

$$h_{\lambda s_1 + (1-\lambda) s_0}(\lambda x + (1-\lambda) y) \geqslant f_{s_1}(x)^{\lambda} g_{s_0}(y)^{1-\lambda},$$

and our inductive hypothesis gives

$$\begin{aligned}
H(\lambda s_1 + (1-\lambda) s_0) := \int_{\mathbb{R}^{n-1}} h_{\lambda s_1 + (1-\lambda) s_0} &\geqslant \left(\int_{\mathbb{R}^{n-1}} f_{s_1}\right)^{\lambda} \left(\int_{\mathbb{R}^{n-1}} g_{s_0}\right)^{1-\lambda} \\
&=: F^{\lambda}(s_1) G^{1-\lambda}(s_0).
\end{aligned}$$

Applying the inductive hypothesis once again, this time with $n = 1$, to the functions F, G and H, we get

$$\int_{\mathbb{R}^n} h = \int_{\mathbb{R}} H \geqslant \left(\int_{\mathbb{R}} F\right)^{\lambda} \left(\int_{\mathbb{R}} G\right)^{1-\lambda} = \left(\int_{\mathbb{R}^n} f\right)^{\lambda} \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda}.$$

This completes the proof. $\square$

The dimension free version (1.2.3) of the Brunn-Minkowski inequality is a simple consequence of the Prékopa-Leindler inequality. We consider two non-empty compact subsets K and T of $\mathbb{R}^n$ and, given $\lambda \in (0,1)$, we define $f = \mathbf{1}_K$, $g = \mathbf{1}_T$ and $h = \mathbf{1}_{\lambda K + (1-\lambda)T}$, where $\mathbf{1}_A$ denotes the indicator function of a set A. We check that the assumptions of Theorem 1.2.3 are satisfied, therefore

$$|\lambda K + (1-\lambda)T| = \int_{\mathbb{R}^n} h \geqslant \left(\int_{\mathbb{R}^n} f\right)^{\lambda} \left(\int_{\mathbb{R}^n} g\right)^{1-\lambda} = |K|^{\lambda}|T|^{1-\lambda}.$$

1.2.3. Knothe map

We fix an orthonormal basis $\{e_1, \ldots, e_n\}$ in $\mathbb{R}^n$, and consider two open convex bodies K and T. The properties of the *Knothe map* from K to T with respect to the given coordinate system are described in the following theorem.

Theorem 1.2.4 (Knothe). *There exists a map $\phi : K \to T$ with the following properties:*

(i) *ϕ is triangular: the i-th coordinate function of ϕ depends only on $x_1, \ldots, x_i$. That is,*

$$\phi(x_1, \ldots, x_n) = (\phi_1(x_1), \phi_2(x_1, x_2), \ldots, \phi_n(x_1, \ldots, x_n)).$$

(ii) *The partial derivatives $\frac{\partial \phi_i}{\partial x_i}$ are positive on K, and the Jacobian determinant J_ϕ of ϕ is constant. More precisely, for every $x \in K$,*

$$J_\phi(x) = \prod_{i=1}^{n} \frac{\partial \phi_i}{\partial x_i}(x) = \frac{|T|}{|K|}.$$

Proof. For each $i = 1, \ldots, n$ and $s = (s_1, \ldots, s_i) \in \mathbb{R}^i$ we consider the section

$$K_s = \{y \in \mathbb{R}^{n-i} : (s, y) \in K\}$$

of K (similarly for T). We shall define a one to one and onto map $\phi : K \to T$ as follows.

Let $x = (x_1, \ldots, x_n) \in K$. Then, $K_{x_1} \neq \emptyset$ and we can define $\phi_1(x) = \phi_1(x_1)$ by

$$\frac{1}{|K|} \int_{-\infty}^{x_1} |K_{s_1}|_{n-1} ds_1 = \frac{1}{|T|} \int_{-\infty}^{\phi_1(x_1)} |T_{t_1}|_{n-1} dt_1.$$

In other words, we move in the direction of e_1 until we "catch" a percentage of T which is equal to the percentage of K occupied by $K \cap \{s = (s_1, \ldots, s_n) : s_1 \leqslant x_1\}$. Note that ϕ_1 is defined on K but $\phi_1(x)$ depends only on the first coordinate of $x \in K$. Also,

$$\frac{\partial \phi_1}{\partial x_1}(x) = \frac{|T|}{|K|} \frac{|K_{x_1}|_{n-1}}{|T_{\phi_1(x_1)}|_{n-1}}.$$

We continue by induction. Assume that we have defined $\phi_1(x) = \phi_1(x_1)$, $\phi_2(x) = \phi_2(x_1, x_2)$ and $\phi_{j-1}(x) = \phi_{j-1}(x_1, \ldots, x_{j-1})$ for some $j \geqslant 2$. If $x = (x_1, \ldots, x_n) \in K$ then $K_{(x_1, \ldots, x_{j-1})} \neq \emptyset$, and we define $\phi_j(x) = \phi_j(x_1, \ldots, x_j)$ by

$$\begin{aligned} \frac{|T_{(\phi_1(x_1), \ldots, \phi_{j-1}(x_1, \ldots, x_{j-1}))}|_{n-j+1}}{|K_{(x_1, \ldots, x_{j-1})}|_{n-j+1}} & \int_{-\infty}^{x_j} |K_{(x_1, \ldots, x_{j-1}, s_j)}|_{n-j} ds_j \\ &= \int_{-\infty}^{\phi_j(x_1, \ldots, x_j)} |T_{(\phi_1(x_1), \ldots, \phi_{j-1}(x_1, \ldots, x_{j-1}), t_j)}|_{n-j} dt_j. \end{aligned}$$

It is clear that

$$\frac{\partial \phi_j}{\partial x_j}(x) = \frac{|T_{(\phi_1(x),\dots,\phi_{j-1}(x))}|_{n-j+1}}{|K_{(x_1,\dots,x_{j-1})}|_{n-j+1}} \frac{|K_{(x_1,\dots,x_j)}|_{n-j}}{|T_{(\phi_1(x),\dots,\phi_j(x))}|_{n-j}}.$$

Continuing in this way, we obtain a map $\phi = (\phi_1, \dots, \phi_n) : K \to T$. It is easy to check that ϕ is one to one and onto. Note that

$$\frac{\partial \phi_n}{\partial x_n}(x) = \frac{|T_{(\phi_1(x),\dots,\phi_{n-1}(x))}|_1}{|K_{(x_1,\dots,x_{n-1})}|_1}.$$

By construction, ϕ has properties (i) and (ii). □

Remark 1.2.5. Observe that each choice of coordinate system in $\mathbb{R}^n$ produces a different Knothe map from K onto T. Using the Knothe map one can give one more proof of the Brunn-Minkowski inequality for convex bodies. We may clearly assume that K and T are open. Consider the Knothe map $\phi : K \to T$. It is clear that

$$(I + \phi)(K) \subseteq K + \phi(K) = K + T,$$

and hence, employing property (ii) of ϕ and the arithmetic-geometric means inequality, we write

$$\begin{aligned} |K + T| &\geqslant \int_{(I+\phi)(K)} dx = \int_K |J_{I+\phi}(x)|\, dx \\ &= \int_K \prod_{j=1}^n \Big(1 + \frac{\partial \phi_j}{\partial x_j}(x)\Big) dx \geqslant \int_K \Big(1 + \Big(\prod_{j=1}^n \frac{\partial \phi_j}{\partial x_j}(x)\Big)^{1/n}\Big)^n dx \\ &= |K| \Big(1 + \Big(\frac{|T|}{|K|}\Big)^{1/n}\Big)^n \\ &= \big(|K|^{1/n} + |T|^{1/n}\big)^n. \end{aligned}$$

1.3. Applications of the Brunn-Minkowski inequality

In this section we collect a few very important geometric inequalities that will be frequently used in this book. Most of them are consequences of the Brunn-Minkowski inequality (and more will appear in subsequent chapters).

1.3.1. An inequality of Rogers and Shephard

The *difference body* of a convex body K is the symmetric convex body

$$K - K = \{x - y \mid x, y \in K\}.$$

From the Brunn-Minkowski inequality it is clear that $|K - K| \geqslant 2^n |K|$ with equality if and only if K has a center of symmetry. Rogers and Shephard gave a sharp upper bound for the volume of the difference body.

Theorem 1.3.1 (Rogers-Shephard). *Let K be a convex body in $\mathbb{R}^n$. Then,*

$$|K - K| \leqslant \binom{2n}{n} |K|.$$

Proof. We consider the function $f(x) = |K \cap (x+K)|^{1/n}$; by the Brunn-Minkowski inequality this is a concave function supported on $K-K$. Note that every $x \in K-K$ can be written in the form $x = r\theta$, where $\theta \in S^{n-1}$ and $0 \leqslant r \leqslant \rho_{K-K}(\theta)$. We define a second function $g : K-K \to [0,\infty)$ by $g(r\theta) = f(0)(1-r/\rho_{K-K}(\theta))$. Then, g is linear on the interval $[0, \rho_{K-K}(\theta)\theta]$, it vanishes on the boundary of $K-K$, and $g(0) = f(0)$. Since f is concave, we see that $f \geqslant g$ on $K-K$. Therefore,

$$\begin{aligned}\int_{K-K} |K \cap (x+K)|dx &= \int_{K-K} f^n(x)dx \geqslant \int_{K-K} g^n(x)dx \\ &= [f(0)]^n n\omega_n \int_{S^{n-1}} \int_0^{\rho_{K-K}(\theta)} r^{n-1}(1-r/\rho_{K-K}(\theta))^n dr d\sigma(\theta) \\ &= n\omega_n |K| \int_{S^{n-1}} \rho_{K-K}^n(\theta)\, d\sigma(\theta) \int_0^1 t^{n-1}(1-t)^n dt \\ &= |K||K-K| \frac{n\Gamma(n)\Gamma(n+1)}{\Gamma(2n+1)} \\ &= \binom{2n}{n}^{-1} |K||K-K|.\end{aligned}$$

On the other hand, Fubini's theorem gives

$$\begin{aligned}\int_{K-K} |K \cap (x+K)|dx = \int_{\mathbb{R}^n} |K \cap (x+K)|dx \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathbf{1}_K(y)\mathbf{1}_{x+K}(y)dydx \\ &= \int_{\mathbb{R}^n} \mathbf{1}_K(y) \left(\int_{\mathbb{R}^n} \mathbf{1}_{y-K}(x)dx \right) dy \\ &= \int_K |y-K|dy = |K|^2.\end{aligned}$$

Combining the above, we conclude the proof. □

Theorem 1.3.1 will be used in the following way: we always have

$$|K-K|^{1/n} \leqslant 4|K|^{1/n},$$

and hence every convex body which contains the origin is contained in a symmetric convex body with the same more or less *volume radius* (for any convex body A in $\mathbb{R}^n$, its volume radius is defined as the radius $\text{v.rad}(A) := (|A|/|B_2^n|)^{1/n}$ of a Euclidean ball that has the same volume as A).

Rogers and Shephard also proved that, when the barycenter of K is at 0, then

$$|K \cap (-K)| \geqslant 2^{-n}|K|.$$

This result implies that every convex body contains a convex body which has a center of symmetry and the same more or less volume radius. Let us also note that Milman and Pajor obtained the following generalization:

THEOREM 1.3.2 (Milman-Pajor). *Let K and L be two convex bodies in $\mathbb{R}^n$ with barycenter at the origin. Then,*

$$|K|\,|L| \leqslant |K+L|\,|K \cap (-L)|.$$

1.3.2. Borell's lemma

Borell's lemma states that if the intersection of $A \cap K$ of a convex body K with a symmetric convex set A captures more than half of the volume of K, then the percentage of K which stays outside tA, $t > 1$ decreases exponentially with respect to t as $t \to \infty$, at a rate which is independent from the body K and the dimension n.

THEOREM 1.3.3 (Borell). *Let K be a convex body of volume 1 in $\mathbb{R}^n$, and let A be a closed, convex and symmetric set such that $|K \cap A| = \delta > \frac{1}{2}$. Then, for every $t > 1$ we have*

$$|K \cap (tA)^c| \leqslant \delta \left(\frac{1-\delta}{\delta}\right)^{\frac{t+1}{2}}.$$

Proof. Observe that

$$A^c \supseteq \frac{2}{t+1}(tA)^c + \frac{t-1}{t+1}A.$$

Otherwise, there exists $a \in A$ which can be written in the form $a = \frac{2}{t+1}y + \frac{t-1}{t+1}a_1$ for some $a_1 \in A$ and $y \notin tA$, and then, using the convexity and symmetry of A, we can write

$$\frac{1}{t}y = \frac{t+1}{2t}a + \frac{t-1}{2t}(-a_1) \in A,$$

which implies that $y \in tA$, a contradiction. Since K is convex, we then have

$$A^c \cap K \supseteq \frac{2}{t+1}\big((tA)^c \cap K\big) + \frac{t-1}{t+1}\big(A \cap K\big).$$

From the Brunn-Minkowski inequality we get

$$1 - \delta = |A^c \cap K| \geqslant |(tA)^c \cap K|^{\frac{2}{t+1}} |A \cap K|^{\frac{t-1}{t+1}} = |(tA)^c \cap K|^{\frac{2}{t+1}} \delta^{\frac{t-1}{t+1}},$$

and the result follows. □

1.3.3. The isoperimetric inequality for convex bodies

The *surface area* (or Minkowski content) $\partial(K)$ of a convex body K is defined by

$$\partial(K) = \lim_{t \to 0^+} \frac{|K + tB_2^n| - |K|}{t}.$$

It is a well-known fact that among all convex bodies of a given volume the ball has minimal surface area. This is an immediate consequence of the Brunn-Minkowski inequality: If K is a convex body in $\mathbb{R}^n$ and if we write $|K| = |rB_2^n|$ for some $r > 0$, then for every $t > 0$

$$|K + tB_2^n|^{1/n} \geqslant |K|^{1/n} + t|B_2^n|^{1/n} = (r+t)|B_2^n|^{1/n}.$$

It follows that the surface area $\partial(K)$ of K satisfies

$$\begin{aligned}\partial(K) &= \lim_{t \to 0^+} \frac{|K + tB_2^n| - |K|}{t} \geqslant \lim_{t \to 0^+} \frac{(r+t)^n - r^n}{t}|B_2^n| \\ &= nr^{n-1}|B_2^n|,\end{aligned}$$

which shows that

$$\partial(K) \geqslant n|B_2^n|^{\frac{1}{n}}|K|^{\frac{n-1}{n}} \tag{1.3.1}$$

with equality if $K = rB_2^n$. The question of uniqueness in the equality case is more delicate.

Actually, the argument above gives a stronger statement: if $|K| = |rB_2^n|$ then

$$|K + tB_2^n| \geqslant |rB_2^n + tB_2^n|$$

for every $t > 0$. If we fix the volume of K then, *for every* $t > 0$, the t-extension

$$K_t = \{y \mid d(y, K) \leqslant t\}$$

of K has minimal volume if K is a ball.

1.3.4. Blaschke-Santaló inequality

Let K be a symmetric convex body in $\mathbb{R}^n$. Recall that the polar body of K is the symmetric convex body

$$K^\circ = \{y \in \mathbb{R}^n \mid \forall x \in K \;\; |\langle x, y\rangle| \leqslant 1\}.$$

The *volume product* $s(K)$ of K is defined by

$$s(K) := |K|\,|K^\circ|.$$

Since $(TK)^\circ = (T^{-1})^*(K^\circ)$ for every $T \in GL(n)$, we readily see that $s(K) = s(TK)$. So, the volume product is an invariant of the linear class of K. The Blaschke-Santaló inequality states that $s(K)$ is maximized when K is an ellipsoid.

Theorem 1.3.4 (Blaschke-Santaló). *Let K be a symmetric convex body in $\mathbb{R}^n$. Then, $|K|\,|K^\circ| \leqslant \omega_n^2$.*

Meyer and Pajor gave a very simple proof of this fact which is based on Steiner symmetrization. The main step is to show that if K is a symmetric convex body in $\mathbb{R}^n$ and if $K_1 = S_\theta(K)$ is the Steiner symmetrization of K in the direction of any $\theta \in S^{n-1}$, then

$$|K^\circ| \leqslant |(K_1)^\circ|. \tag{1.3.2}$$

Since Steiner symmetrization preserves volume, this implies that

$$s(K) \leqslant s(K_1),$$

and then the theorem follows by applying a suitable sequence of Steiner symmetrizations to K.

For the proof of (1.3.2) we may clearly assume that $\theta^\perp = \mathbb{R}^{n-1}$ and then we write

$$S_\theta(K) = K_1 = \left\{ \left(x, \frac{t_1 - t_2}{2}\right) : x \in P_{\theta^\perp}K, (x, t_1) \in K, (x, t_2) \in K \right\}.$$

For every $A \subseteq \mathbb{R}^n$ we write

$$A(t) = \{x \in \mathbb{R}^{n-1} : (x, t) \in A\}.$$

With this notation we show that, for every $s \in \mathbb{R}$,

$$\frac{K^\circ(s) + K^\circ(-s)}{2} \subseteq (K_1)^\circ(s).$$

Then, we apply the Brunn-Minkowski inequality to get

$$|(K_1)^\circ(s)| \geqslant |K^\circ(s)|^{\frac{1}{2}}|K^\circ(-s)|^{\frac{1}{2}},$$

and since $|K^\circ(s)| = |K^\circ(-s)|$ by the symmetry of K°, we see that

$$|(K_1)^\circ(s)| \geqslant |K^\circ(s)|$$

for every $s \in \mathbb{R}$. Integrating with respect to s we have

$$|(K_1)^\circ| = \int_{-\infty}^{+\infty} |(K_1)^\circ(s)|ds \geqslant \int_{-\infty}^{+\infty} |K^\circ(s)|ds = |K^\circ|$$

and (1.3.2) is proved. □

1.3.5. Urysohn's inequality

The *mean width* $w(K)$ of a convex body K in $\mathbb{R}^n$ is defined by

$$w(K) = \int_{S^{n-1}} h_K(u)\, d\sigma(u),$$

where h_K is the support function of K. A classical inequality of Urysohn states that for fixed volume, Euclidean ball has minimal mean width.

THEOREM 1.3.5 (Urysohn). *Let K be a convex body in $\mathbb{R}^n$. Then,*

$$w(K) \geqslant \left(\frac{|K|}{|B_2^n|}\right)^{1/n}.$$

A simple proof of this fact may be given with the method of Steiner symmetrization. The main step is to show that

$$w(S_\theta(K)) \leqslant w(K)$$

for every $\theta \in S^{n-1}$. Urysohn's inequality then follows by applying a suitable sequence of Steiner symmetrizations to K.

We can give a second proof by averaging K using orthogonal transformations. One easily checks that $K_N = \frac{1}{N}\sum_{i=1}^N U_i(K)$ has the same mean width as K. Clearly, by taking U_i to be a net in $O(n)$ one may make sure that K_N converges to a multiple of the Euclidean ball, and thus it must converge to $w(K)B_2^n$. But, by the Brunn-Minkowski inequality, the volume of K_N is greater than the volume of K, and we get the inequality in the limit.

1.4. Mixed volumes

In this section we introduce mixed volumes and survey some of the fundamental results, formulas and inequalities, that will be used in this book.

1.4.1. Minkowski's theorem

Recall that $\tilde{\mathcal{K}}_n$ denotes the convex cone of all non-empty, compact convex subsets of $\mathbb{R}^n$. Minkowski's fundamental theorem on mixed volumes states that there exists a function $V : (\tilde{\mathcal{K}}_n)^n \to \mathbb{R}^+$ satisfying the following conditions:

(i) V has volume as its diagonal: if $K \in \tilde{\mathcal{K}}_n$ then $V(K,\dots,K) = |K|$.

(ii) V is positive linear in each of its arguments: if $K_1,\dots,K_i^{(1)},K_i^{(2)},\dots,K_n \in \tilde{\mathcal{K}}_n$ and $t_1,t_2 \geqslant 0$, then

$$V(K_1,\dots,t_1K_i^{(1)}+t_2K_i^{(2)},\dots,K_n) = \sum_{j=1}^2 t_jV(K_1,\dots,K_i^{(j)},\dots,K_n).$$

(iii) V is symmetric: if $K_1, \dots, K_n \in \tilde{\mathcal{K}}_n$ and σ is any permutation of the indices, then

$$V(K_{\sigma(1)}, \dots, K_{\sigma(n)}) = V(K_1, \dots, K_n).$$

The value $V(K_1, \dots, K_n)$ is called the *mixed volume* of $K_1, \dots, K_n$. It follows that if $K_1, \dots, K_m \in \tilde{\mathcal{K}}_n, m \in \mathbb{N}$, then the volume of $t_1K_1 + \cdots + t_mK_m$ is a homogeneous polynomial of degree n in $t_i > 0$. That is,

$$|t_1K_1 + \cdots + t_mK_m| = \sum_{1 \leqslant i_1, \dots, i_n \leqslant m} V(K_{i_1}, \dots, K_{i_n}) t_{i_1} \cdots t_{i_n}.$$

In particular, if $K, L \in \tilde{\mathcal{K}}_n$ then the function $|K + tL|$ is a polynomial in $t \in [0, \infty)$:

$$|K + tL| = \sum_{j=0}^{n} \binom{n}{j} V_j(K, L)\, t^j, \tag{1.4.1}$$

where $V_j(K, L) = V(K; n-j, L; j)$ is the j-th mixed volume of K and L. Here and elsewhere we use the notation $L; j$ for $L, \dots, L$ j-times.

From the last formula we obtain

$$V_1(K, L) = \frac{1}{n} \lim_{t \to 0^+} \frac{|K + tL| - |K|}{t},$$

which together with the classical Brunn-Minkowski inequality $|K+tL|^{1/n} \geqslant |K|^{1/n} + t|L|^{1/n}$ implies that

$$V_1(K, L) \geqslant |K|^{\frac{n-1}{n}} |L|^{1/n}$$

for all $K, L \in \tilde{\mathcal{K}}_n$. This is *Minkowski's first inequality.*

1.4.2. Steiner's formula and quermassintegrals

Steiner's formula may be viewed as a special case of (1.4.1). The volume of $K + tB_2^n$, $t > 0$, can be expanded as a polynomial in t:

$$|K + tB_2^n| = \sum_{j=0}^{n} \binom{n}{j} W_j(K) t^j,$$

where

$$W_j(K) = V_j(K, B_2^n) = V(K; n-j, B_2^n; j)$$

is the *j-th quermassintegral* of K. The quermassintegrals W_j inherit properties of mixed volumes: they are monotone, continuous with respect to the Hausdorff metric, and homogeneous of degree $n - j$.

It is easy to see that the surface area of K is given by

$$\partial(K) = nW_1(K).$$

Kubota's integral formula expresses the quermassintegral $W_j(K)$ as an average of the volumes of $(n-j)$-dimensional projections of K:

$$W_j(K) = \frac{\omega_n}{\omega_{n-j}} \int_{G_{n,n-j}} |P_F(K)| d\nu_{n,n-j}(F).$$

Applying this formula for $j = n - 1$ we see that

$$W_{n-1}(K) = \omega_n w(K).$$

It will be convenient for us to work with a normalized variant of $W_{n-j}(K)$: for every $1 \leqslant j \leqslant n$ we set

$$Q_j(K) = \left(\frac{1}{\omega_j}\int_{G_{n,j}} |P_F(K)|\, d\nu_{n,j}(F)\right)^{1/j}.$$

Note that $Q_1(K) = w(K)$. Kubota's formula shows that

$$Q_j(K) = \left(\frac{W_{n-j}(K)}{\omega_n}\right)^{1/j}.$$

1.4.3. Mixed area measures

We fix $K \in \tilde{\mathcal{K}}_n$, and for every $L \in \tilde{\mathcal{K}}_n$ we define $f(h_L) = V_1(K, L)$. We extend f linearly on the subspace $D(S^{n-1}) = \text{span}\{h_L|_{S^{n-1}}, L \in \tilde{\mathcal{K}}_n\}$ of $C(S^{n-1})$. From the additivity of V_1 with respect to L and the fact that $h_{L_1+L_2} = h_{L_1} + h_{L_2}$ whenever $L_1, L_2 \in \tilde{\mathcal{K}}_n$, f is a well-defined positive functional on $D(S^{n-1})$, and hence it extends to a positive functional on $C(S^{n-1})$. By the Riesz representation theorem, we can find a Borel measure σ_K on S^{n-1} for which

$$V_1(K, L) = \frac{1}{n}\int_{S^{n-1}} h_L(u) d\sigma_K(u), \qquad L \in \mathcal{K}_n.$$

σ_K is called the *surface area measure* of K. Equivalently, if B is a Borel subset of S^{n-1} then $\sigma_K(B)$ is the $(n-1)$-dimensional surface measure of the set of all boundary points of K at which there exists an exterior normal in B (for a polytope K with facets $F_1, \ldots, F_m$ having exterior normals $u_1, \ldots, u_m$ respectively, σ_K is the measure supported by $\{u_1, \ldots, u_m\}$ with $\sigma_K(\{u_j\}) = |F_j|, j = 1, \ldots, m$). As a consequence of the integral representation for $V_1(K, L)$ we also see that $L_1 \subseteq L_2$ implies $V_1(K, L_1) \leqslant V_1(K, L_2)$.

More generally, the *mixed area measures* were introduced by Alexandrov and may be viewed as a local generalization of the mixed volumes. For any $(n-1)$-tuple $\mathcal{C} = (K_1, \ldots, K_{n-1})$ of elements of $\tilde{\mathcal{K}}_n$, the Riesz representation theorem guarantees the existence of a Borel measure $S(\mathcal{C}, \cdot)$ on the unit sphere S^{n-1} such that

$$V(L, K_1, \ldots, K_{n-1}) = \frac{1}{n}\int_{S^{n-1}} h_L(u) dS(\mathcal{C}, u)$$

for every $L \in \tilde{\mathcal{K}}_n$. The local analogue of Minkowski's theorem is

$$\sigma_{\sum_{i=1}^m t_i K_i}(B) = \sum_{1\leqslant i_1,\ldots,i_n\leqslant m} S(K_{i_1}, \ldots, K_{i_{n-1}}, \omega) t_{i_1} \ldots t_{i_{n-1}}$$

for all Borel sets $B \subseteq S^{n-1}$ and all $t_i > 0$, $K_i \in \tilde{\mathcal{K}}_n$, $m \in \mathbb{N}$.

The *j-th area measure* of K is defined by $S_j(K, \cdot) = S(K; j, B_2^n; n-j-1, \cdot)$, $j = 0, 1, \ldots, n-1$. It follows that the quermassintegrals of K can be represented by

$$W_j(K) = \frac{1}{n}\int_{S^{n-1}} h_K(u) dS_{n-j-1}(K, u), \qquad j = 0, 1, \ldots, n-1$$

or alternatively,

$$W_j(K) = \frac{1}{n}\int_{S^{n-1}} dS_{n-j}(K, u), \qquad i = 1, \ldots, n.$$

1.4.4. The Alexandrov-Fenchel inequalities

The *Alexandrov-Fenchel inequality* generalizes the Brunn-Minkowski inequality and its consequences. It states that if $K, L, K_3, \ldots, K_n \in \tilde{\mathcal{K}}_n$, then

$$V(K, L, K_3, \ldots, K_n)^2 \geqslant V(K, K, K_3, \ldots, K_n)V(L, L, K_3, \ldots, K_n).$$

From this inequality one can recover the Brunn-Minkowski inequality as well as the following generalization for the quermassintegrals:

$$W_j(K+L)^{\frac{1}{n-j}} \geqslant W_j(K)^{\frac{1}{n-j}} + W_j(L)^{\frac{1}{n-j}}, \qquad j = 0, \ldots, n-1$$

for any pair of convex bodies in $\mathbb{R}^n$.

Steiner's formula and the Brunn-Minkowski inequality show that

$$\sum_{j=0}^{n} \binom{n}{j} \frac{W_j(K)}{|B_2^n|} t^j = \frac{|K + tB_2^n|}{|B_2^n|} \geqslant \Big(\Big(\frac{|K|}{|B_2^n|}\Big)^{1/n} + t\Big)^n$$

$$= \sum_{j=0}^{n} \binom{n}{j} \Big(\frac{|K|}{|B_2^n|}\Big)^{\frac{n-j}{n}} t^j$$

for every $t > 0$. Since the first and the last term are equal on both sides of this inequality, we must have

$$\frac{W_1(K)}{|B_2^n|} \geqslant \left(\frac{|K|}{|B_2^n|}\right)^{\frac{n-1}{n}}$$

which is the isoperimetric inequality for convex bodies, and

$$w(K) = \frac{W_{n-1}(K)}{|B_2^n|} \geqslant \left(\frac{|K|}{|B_2^n|}\right)^{\frac{1}{n}},$$

which is Urysohn's inequality. Both inequalities are special cases of the set of *Alexandrov inequalities*

$$\left(\frac{W_i(K)}{|B_2^n|}\right)^{\frac{1}{n-i}} \geqslant \left(\frac{W_j(K)}{|B_2^n|}\right)^{\frac{1}{n-j}} \qquad n > i > j \geqslant 0.$$

This implies that the sequence $(W_0(K), \ldots, W_n(K))$ is log-concave: we have

$$W_j^{k-i} \geqslant W_i^{k-j} W_k^{j-i}$$

if $0 \leqslant i < j < k \leqslant n$. From these inequalities one can check that $Q_j(K)$ is a decreasing function of j.

1.4.5. Projection bodies

Minkowski's existence theorem states that if $u_1, \ldots, u_m$ are distinct unit vectors with the origin in the interior of their convex hull, and if $\gamma_1, \ldots, \gamma_m$ are positive real numbers with $\sum_{j=1}^{m} \gamma_j u_j = 0$, then there exists a polytope K having $u_1, \ldots, u_m$ as its (only) normal vectors and satisfying $|F_j| = \gamma_j, j = 1, \ldots, m$, where F_j is the facet of K corresponding to u_j. By approximation one obtains that a finite Borel measure μ on S^{n-1} is the surface area measure of some $K \in \mathcal{K}_n$ if and only if μ is not concentrated on any great subsphere of S^{n-1}, and

$$\int_{S^{n-1}} u \, d\mu(u) = 0.$$

Note that the second condition is always satisfied if μ is an even measure.

If $K \in \mathcal{K}_n$, the *projection body* ΠK of K is the symmetric convex body whose support function is defined by

$$h_{\Pi K}(\theta) = |P_\theta(K)|, \qquad \theta \in S^{n-1}.$$

One has the integral representation

$$|P_\theta(K)| = \frac{1}{2}\int_{S^{n-1}} |\langle u, \theta\rangle| \, d\sigma_K(u),$$

which is easily verified in the case of a polytope, and extends to any $K \in \mathcal{K}_n$ by approximation (*Cauchy's formula*). It follows that the support function of a projection body is given by

$$h_{\Pi K}(\theta) = \int_{S^{n-1}} |\langle u, \theta\rangle| \, d\mu(u)$$

for some finite even Borel measure on S^{n-1}. This also shows that all projection bodies are indeed convex. The integral representation of $h_{\Pi K}$ also shows that ΠK is a zonoid: a Hausdorff limit of a sequence of finite Minkowski sums of line segments. Minkowski's existence theorem implies that, conversely, every zonoid is the projection body of some symmetric convex body in $\mathbb{R}^n$. Moreover, if we denote by $\mathcal{Z}$ the class of zonoids, *Alexandrov's uniqueness theorem* shows that the Minkowski map $\Pi : \mathcal{C}_n \to \mathcal{Z}$ with $K \to \Pi K$, is injective. Note also that $\mathcal{Z}$ is invariant under invertible linear transformations and closed in the Hausdorff metric.

1.5. Classical positions of convex bodies

The family of positions of a convex body K in $\mathbb{R}^n$ is the class $\{z + T(K) : z \in \mathbb{R}^n, T \in GL(n)\}$. The right choice of a position is often quite important for the study of affinely invariant quantities. For example, let K be a symmetric convex body in $\mathbb{R}^n$ and consider the volume product $s(K) = |K|\,|K^\circ|$. The Blaschke-Santaló inequality (Theorem 1.3.4) asserts that $s(K)$ is maximized if and only if K is an ellipsoid (note that $s(K)$ is invariant under $GL(n)$). On the other hand, a simple application of Hölder's inequality shows that, for every symmetric convex body A in $\mathbb{R}^n$,

$$\frac{|A|}{|B_2^n|} = \int_{S^{n-1}} \|\theta\|_A^{-n} d\sigma(\theta) \geqslant w(A^\circ)^{-n}.$$

This implies that

$$\left(\frac{s(B_2^n)}{s(K)}\right)^{1/n} \leqslant \min_{T \in GL(n)} w(TK) w((TK)^\circ).$$

Therefore, in order to obtain a "reverse Blaschke-Santaló inequality" it is useful to study the quantity

$$\max_K \min_{T \in GL(n)} w(TK) w((TK)^\circ),$$

or equivalently, to study the position $\tilde{K}$ of K which minimizes $w(TK)w((TK)^\circ)$ over all $T \in GL(n)$.

In this section we introduce three classical positions of convex bodies. All of them arise as solutions of extremal problems of the following type: we normalize the volume of K to be 1 and ask for the maximum or minimum of $f(TK)$ over all $T \in SL(n)$, where f is some functional on convex bodies (in the example above, f is the product of the mean widths of a body and its polar). An interesting feature of this procedure, which was put forward in [**204**], is that a simple variational method leads to a geometric description of the extremal position, and that in many cases this position satisfies an isotropic condition for an appropriate measure on S^{n-1}.

DEFINITION 1.5.1. A Borel measure μ on S^{n-1} is called *isotropic* if

$$\int_{S^{n-1}} \langle x, \theta\rangle^2 d\mu(x) = \frac{\mu(S^{n-1})}{n} \tag{1.5.1}$$

for every $\theta \in S^{n-1}$. We will make frequent use of the next standard lemma.

LEMMA 1.5.2. *Let μ be a Borel measure on S^{n-1}. The following are equivalent:*

(i) *μ is isotropic.*

(ii) *For every $i, j = 1, \dots, n$,*

$$\int_{S^{n-1}} \phi_i\phi_j d\mu(\phi) = \frac{\mu(S^{n-1})}{n}\delta_{i,j}. \tag{1.5.2}$$

(iii) *For every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ (we will write $T \in L(\mathbb{R}^n)$),*

$$\int_{S^{n-1}} \langle \phi, T\phi\rangle d\mu(\phi) = \frac{\mathrm{tr}(T)}{n}\mu(S^{n-1}). \tag{1.5.3}$$

Proof. Setting $\theta = e_i$ and $\theta = \frac{e_i+e_j}{\sqrt{2}}$ in (1.5.1) we get (1.5.2). On observing that if $T = (t_{ij})_{i,j=1}^n$ then $\langle \phi, T\phi\rangle = \sum_{i,j=1}^n t_{ij}\phi_i\phi_j$, we readily see that (1.5.2) implies (1.5.3). Finally, note that applying (1.5.3) with $T(\phi) = \langle \phi, \theta\rangle\theta$ we get (1.5.1). $\square$

1.5.1. John's position

Given a convex body K in $\mathbb{R}^n$, we consider the family $\mathcal{E}(K)$ of the *ellipsoids* which are contained in K. An ellipsoid in $\mathbb{R}^n$ is a convex body of the form

$$\mathcal{E} = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n \frac{\langle x, v_i\rangle^2}{\alpha_i^2} \leqslant 1 \right\},$$

where $\{v_i\}_{i\leqslant n}$ is an orthonormal basis of $\mathbb{R}^n$, and $\alpha_1, \dots, \alpha_n$ are positive reals (the directions and lengths of the semiaxes of $\mathcal{E}$, respectively). It is easy to check that $\mathcal{E} = T(B_2^n)$, where T is the linear transformation of $\mathbb{R}^n$ defined by $T(v_i) = \alpha_i v_i$, $i = 1, \dots, n$. Therefore, the volume of $\mathcal{E}$ is equal to

$$|\mathcal{E}| = |B_2^n| \prod_{i=1}^n \alpha_i.$$

The *volume ratio* of K is the quantity

$$vr(K) = \inf\Big\{ \Big(\frac{|K|}{|\mathcal{E}|}\Big)^{1/n} : \mathcal{E} \subseteq K \Big\},$$

where the infimum is taken over all the ellipsoids which are contained in K. One can show that there is a unique ellipsoid $\mathcal{E}$ of maximal volume which is contained in K. We will say that $\mathcal{E}$ is the *maximal volume ellipsoid* of K. Moreover, there is a unique ellipsoid $\mathcal{E}$ which contains K and has minimal volume (the *minimal volume ellipsoid* of K).

Assume that B_2^n is the maximal volume ellipsoid of K; then we say that K is in John's position. We will say that $x \in \mathbb{R}^n$ is a *contact point* of K and B_2^n if

$$\|x\|_2 = \|x\|_K = \|x\|_{K^\circ} = 1.$$

John's theorem describes the distribution of contact points on the unit sphere S^{n-1}.

THEOREM 1.5.3 (John). *If B_2^n is the maximal volume ellipsoid of the symmetric convex body K in $\mathbb{R}^n$, then there exist contact points $x_1, \ldots, x_m$ of K and B_2^n, and positive real numbers $c_1, \ldots, c_m$ such that*

$$x = \sum_{j=1}^{m} c_j \langle x, x_j \rangle x_j$$

for every $x \in \mathbb{R}^n$.

REMARK 1.5.4. Theorem 1.5.3 says that the identity operator I of $\mathbb{R}^n$ can be represented in the form

$$I = \sum_{j=1}^{m} c_j x_j \otimes x_j, \tag{1.5.4}$$

where $x_j \otimes x_j$ is the projection in the direction of x_j: $(x_j \otimes x_j)(x) = \langle x, x_j \rangle x_j$. Note that for every $x \in \mathbb{R}^n$

$$\|x\|_2^2 = \langle x, x \rangle = \sum_{j=1}^{m} c_j \langle x, x_j \rangle^2.$$

Also, if we choose $x = e_i$, $i = 1, \ldots, n$, where $\{e_i\}$ is the standard orthonormal basis of $\mathbb{R}^n$, we have

$$\begin{aligned} n = \sum_{i=1}^{n} \|e_i\|_2^2 &= \sum_{i=1}^{n} \sum_{j=1}^{m} c_j \langle e_i, x_j \rangle^2 = \sum_{j=1}^{m} c_j \sum_{i=1}^{n} \langle e_i, x_j \rangle^2 \\ &= \sum_{j=1}^{m} c_j \|x_j\|_2^2 = \sum_{j=1}^{m} c_j. \end{aligned}$$

Sketch of the proof of Theorem 1.5.3. We follow the presentation of K. Ball from [**39**]; Remark 1.5.4 shows that a necessary condition for a representation of the form (1.5.4) is that $\sum_{j=1}^{m} \frac{c_j}{n} = 1$. Our purpose is then to show that I/n is in the convex hull of the set of all matrices that have the form $x \otimes x$ for some contact point x of K and B_2^n. To this end, we define

$$\mathcal{C} = \{x \otimes x : \|x\|_2 = \|x\|_K = 1\},$$

and show that $I/n \in \mathrm{conv}(\mathcal{C})$. Note that $\mathrm{conv}(\mathcal{C})$ is a non-empty compact convex subset of $\mathbb{R}^{n^2}$.

Assume that $I/n \notin \mathrm{conv}(\mathcal{C})$. Using the Hahn-Banach theorem, one can prove the next lemma.

LEMMA 1.5.5. *If $I/n \notin \mathrm{conv}(\mathcal{C})$, there exist $s > 0$ and B symmetric with* $\mathrm{tr}(B) = 0$, *such that*

$$\langle B, x \otimes x \rangle \geqslant s$$

for every $x \otimes x \in \mathcal{C}$.

Then, we consider $\delta > 0$ small enough, and define the ellipsoid

$$\mathcal{E}_\delta = \{x \in \mathbb{R}^n : \langle (I + \delta B)x, x \rangle \leqslant 1\}.$$

The next step is to show that $\mathcal{E}_\delta \subseteq K$ if δ is small enough.

LEMMA 1.5.6. *There exists $\delta_0 > 0$ such that $\mathcal{E}_\delta \subseteq K$ for every $0 < \delta < \delta_0$.*

This leads to a contradiction. Choose $\delta > 0$ so small that $I + \delta B$ is positive definite and the ellipsoid $\mathcal{E}_\delta$ is contained in K. Since B_2^n is the maximal volume ellipsoid of K, we have $|\mathcal{E}_\delta| \leqslant |B_2^n|$. On the other hand, if $I + \delta B = S_\delta^2$, then

$$|\mathcal{E}_\delta| = |S_\delta^{-1}(B_2^n)| = |B_2^n|/\sqrt{\det(I+\delta B)}.$$

It follows that $\det(I + \delta B) \geqslant 1$. By the arithmetic-geometric means inequality

$$[\det(I+\delta B)]^{\frac{1}{n}} \leqslant \frac{\mathrm{tr}(I+\delta B)}{n} = 1 + \delta\frac{\mathrm{tr}(B)}{n} = 1,$$

because $\mathrm{tr}(B) = 0$. This means that we have equality in the arithmetic-geometric means inequality, and this implies that $I+\delta B$ is a multiple of the identity: $I+\delta B = \mu I$. But then, B is a multiple of the identity, and since $\mathrm{tr}(B) = 0$ we get $B = 0$.

This contradicts Lemma 1.5.5, because $\langle Bx, x\rangle \geqslant s > 0$ for all $x \otimes x \in \mathcal{C}$. Therefore, $I/n \in \mathrm{conv}(\mathcal{C})$. To finish the proof, we have to show that $\|x_j\|_2 = \|x_j\|_K = \|x_j\|_{K^\circ} = 1$ for all j. Since $x_j \in S^{n-1}$, we have

$$1 = \langle x_j, x_j\rangle \leqslant \|x_j\|_K\|x_j\|_{K^\circ} = \|x_j\|_{K^\circ}, \qquad j = 1, \dots, m.$$

On the other hand, at each x_j, K and B_2^n have the same supporting hyperplane with normal vector x_j. Therefore, for every $x \in K$ we have $\langle x, x_j\rangle \leqslant 1$, and by the symmetry of K, $|\langle x, x_j\rangle| \leqslant 1$. It follows that $\|x_j\|_K = \|x_j\|_{K^\circ} = \|x_j\|_2 = 1$ for all $j = 1, \dots, m$. □

Theorem 1.5.3 implies

$$\sum_{i=1}^m c_i\langle x_i, \theta\rangle^2 = 1$$

for every $\theta \in S^{n-1}$. In the terminology of Definition 1.5.1 the measure μ on S^{n-1} that gives mass c_i to the point x_i, $i = 1, \dots, m$, is isotropic. In this sense, John's position is an *isotropic position.* Conversely, one can show that if K is a symmetric convex body in $\mathbb{R}^n$ which contains the Euclidean unit ball B_2^n and if there exists an isotropic Borel measure μ on S^{n-1} which is supported by the contact points of K and B_2^n, then B_2^n is the maximal volume ellipsoid of K.

A well-known consequence of Theorem 1.5.3 (which is usually called *John's theorem*) states that if K is a symmetric convex body in $\mathbb{R}^n$ and $\mathcal{E}$ is the maximal volume ellipsoid of K, then $K \subseteq \sqrt{n}\mathcal{E}$. This is equivalent to the next proposition.

Theorem 1.5.7 (John). *If B_2^n is the maximal volume ellipsoid of K, then $K \subset \sqrt{n}B_2^n$.*

Proof. Consider the representation of the identity

$$x = \sum_{j=1}^m c_j\langle x, x_j\rangle x_j$$

of Theorem 1.5.3. We will use the fact that $\|x_j\|_K = \|x_j\|_{K^\circ} = \|x_j\|_2 = 1$, $j = 1, \dots, m$. For every $x \in K$ we have

$$\|x\|_2^2 = \sum_{j=1}^m c_j\langle x, x_j\rangle^2 \leqslant \sum_{j=1}^m c_j = n.$$

This shows that $\|x\|_2 \leqslant \sqrt{n}$. Therefore, $B_2^n \subseteq K \subseteq \sqrt{n}B_2^n$. □

John's theorem can be extended to the case of not necessarily symmetric convex bodies.

THEOREM 1.5.8. *Let K be a convex body in $\mathbb{R}^n$ such that B_2^n is the ellipsoid of maximal volume inscribed in K. We can find contact points $x_1,\dots,x_m$ of K and B_2^n, and positive reals $c_1,\dots,c_m$, such that: $\sum_{j=1}^m c_jx_j=0$ and*

$$I=\sum_{j=1}^m c_jx_j\otimes x_j.$$

Löwner's position is dual to John's position. We say that a convex body K is in Löwner's position if the ellipsoid of minimal volume containing K is the Euclidean unit ball B_2^n. By John's theorem, K is in Löwner's position if and only if $K\subseteq B_2^n$ and there exist $x_1,\dots,x_m\in \mathrm{bd}(K)\cap S^{n-1}$ and positive real numbers $c_1,\dots,c_m$ such that the measure μ on S^{n-1} which is supported by $\{x_1,\dots,x_m\}$ and gives mass c_j to $\{x_j\}$, $j=1,\dots,m$, is isotropic.

1.5.2. Dvoretzky-Rogers lemmas

Assume that B_2^n is the ellipsoid of maximal volume contained in the symmetric convex body K. Starting from John's representation of the identity $I=\sum_{j=1}^m c_jx_j\otimes x_j$ of Theorem 1.5.3, Dvoretzky and Rogers obtained precise information on the distribution of the contact points x_j on the Euclidean unit sphere. There are several results of this type, known as "Dvoretzky-Rogers lemmas". An example is given by Theorem 1.5.10 for which we sketch a proof.

LEMMA 1.5.9. *If B_2^n is the maximal volume ellipsoid of the symmetric convex body K then for every $T\in L(\mathbb{R}^n)$ there exists a contact point y of K and B_2^n which satisfies*

$$\langle y,Ty\rangle\geqslant\frac{\mathrm{tr}(T)}{n}.$$

Proof. We have

$$\mathrm{tr}\,T=\langle T,I\rangle=\sum_{j=1}^m c_j\langle T,x_j\otimes x_j\rangle.$$

Since $\sum_{j=1}^m c_j=n$, we may clearly find y among the x_j's which satisfies

$$\langle y,Ty\rangle=\langle T,y\otimes y\rangle\geqslant\frac{\mathrm{tr}(T)}{n}.$$

THEOREM 1.5.10 (Dvoretzky-Rogers). *If B_2^n is the maximal volume ellipsoid of the symmetric convex body K then there exists an orthonormal sequence $z_1,\dots,z_n$ in $\mathbb{R}^n$ such that*

$$\Big(\frac{n-i+1}{n}\Big)^{1/2}\leqslant\|z_i\|\leqslant\|z_i\|_2=1$$

for all $1\leqslant i\leqslant n$.

Proof. We define the z_i's inductively; z_1 can be any contact point of K and B_2^n. Assume that $z_1,\dots,z_k$ have been chosen for some $k<n$.

We set $F_k=\mathrm{span}\{z_1,\dots,z_k\}$. Then, $\mathrm{tr}(P_{F_k^\perp})=n-k$, and applying Lemma 1.5.9 we find a contact point y_{k+1} with

$$\|P_{F_k^\perp}y_{k+1}\|_2^2=\langle y_{k+1},P_{F_k^\perp}y_{k+1}\rangle\geqslant\frac{n-k}{n}.$$

It follows that $\|P_{F_k}y_{k+1}\|\leqslant\|P_{F_k}y_{k+1}\|_2\leqslant\sqrt{k/n}$.

We define $z_{k+1} = P_{F_k^\perp} y_{k+1}/\|P_{F_k^\perp} y_{k+1}\|_2$. Then,

$$1 = \|z_{k+1}\|_2 \geqslant \|z_{k+1}\| \geqslant |\langle y_{k+1}, z_{k+1}\rangle| = \left\|P_{F_k^\perp} y_{k+1}\right\|_2 \geqslant \left(\frac{n-k}{n}\right)^{1/2}.$$

The next corollary of Theorem 1.5.10 will play an important role in the proof of Dvoretzky theorem (see Section 1.10.1).

COROLLARY 1.5.11. *Assume that B_2^n is the maximal volume ellipsoid of the symmetric convex body K. If $k = \lfloor n/2\rfloor + 1$, then we can find orthonormal vectors $z_1, \dots, z_k$ such that*

$$\frac{1}{\sqrt{2}} \leqslant \|z_j\| \leqslant 1$$

for all $j = 1, \dots, k$.

1.5.3. Minimal mean width position

Let K be a convex body in $\mathbb{R}^n$ (without loss of generality we may assume that $0 \in \mathrm{int}\,(K)$). Recall that the mean width of K is the quantity

$$w(K) = \int_{S^{n-1}} h_K(x)\, d\sigma(x).$$

We say that K has *minimal mean width* if $w(K) \leqslant w(TK)$ for every $T \in SL(n)$.

We assume for simplicity that h_K is twice continuously differentiable (we then say that K is *smooth enough*) and we consider the measure ν_K on S^{n-1} with density h_K with respect to σ. The next theorem characterizes the minimal mean width position.

THEOREM 1.5.12 (Giannopoulos-Milman). *A smooth enough convex body K in $\mathbb{R}^n$ has minimal mean width if and only if*

$$\int_{S^{n-1}} h_K(x)\langle x, \theta\rangle^2 d\sigma(x) = \frac{w(K)}{n}$$

for every $\theta \in S^{n-1}$ (equivalently, if ν_K is isotropic). Moreover, this minimal mean width position is unique up to orthogonal transformations.

1.5.4. Minimal surface area position

Let K be a convex body of volume 1 in $\mathbb{R}^n$. As in the previous subsection, we consider the problem to find the minimum of the surface area $\partial(TK)$ over all $T \in SL(n)$. This minimum is attained for some T_0; we denote it by ∂_K and we call it the *minimal surface invariant* of K. We say that K has minimal surface area if $\partial(K) = \partial_K |K|^{\frac{n-1}{n}}$.

Recall the definition of the area measure σ_K of K from Section 1.4.3. It is defined on S^{n-1} and corresponds to the usual surface measure on K via the Gauss map: For every Borel $A \subseteq S^{n-1}$, we have

$$\sigma_K(A) = \nu\left(\{x \in \mathrm{bd}(K) : \text{ the outer normal to } K \text{ at } x \text{ is in } A\}\right),$$

where ν is the $(n-1)$-dimensional surface measure on K. We obviously have $\partial(K) = \sigma_K(S^{n-1})$.

A characterization of the minimal surface position through the area measure was given by Petty.

THEOREM 1.5.13 (Petty). *Let K be a convex body of volume 1 in $\mathbb{R}^n$. Then, $\partial(K) = \partial_K$ if and only if σ_K is isotropic. Moreover, this minimal surface area position is unique up to orthogonal transformations.*

1.6. Brascamp-Lieb inequality and its reverse form

1.6.1. Brascamp-Lieb inequality

The Brascamp-Lieb inequality estimates the norm of the multilinear operator $I : L^{p_1}(\mathbb{R}) \times \cdots \times L^{p_m}(\mathbb{R}) \to \mathbb{R}$ defined by

$$I(f_1, \ldots, f_m) = \int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j(\langle u_j, x\rangle)\, dx,$$

where $m \geqslant n$, $p_1, \ldots, p_m \geqslant 1$ with $\frac{1}{p_1} + \cdots + \frac{1}{p_m} = n$, and $u_1, \ldots, u_m \in \mathbb{R}^n$. Brascamp and Lieb proved that the norm of I is the supremum D of

$$\frac{I(g_1, \ldots, g_m)}{\prod_{j=1}^{m} \|g_j\|_{p_j}}$$

over all centered Gaussian functions $g_1, \ldots, g_m$, i.e. over all functions of the form $g_j(t) = e^{-\lambda_j t^2}$, $\lambda_j > 0$. This fact is a generalization of Young's convolution inequality $\|f * g\|_r \leqslant C_{p,q}\|f\|_p\|g\|_q$ for all $f \in L^p(\mathbb{R})$ and $g \in L^q(\mathbb{R})$, where $p, q, r \geqslant 1$ and $1/p + 1/q = 1 + 1/r$.

If we set $c_j = 1/p_j$ and replace f_j by $f_j^{c_j}$ then we can state the Brascamp-Lieb inequality in the following form.

THEOREM 1.6.1 (Brascamp-Lieb). *Let $m \geqslant n$, and let $u_1, \ldots, u_m \in \mathbb{R}^n$ and $c_1, \ldots, c_m > 0$ with $c_1 + \cdots + c_m = n$. Then,*

$$\int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j^{c_j}(\langle x, u_j\rangle)dx \leqslant D \prod_{j=1}^{m} \left(\int_{\mathbb{R}} f_j\right)^{c_j}$$

for all integrable functions $f_j : \mathbb{R} \to [0, \infty)$.

Direct computation of the ratio $I(g_1, \ldots, g_m)/\prod_{j=1}^{m} \|g_j\|_{p_j}$ for Gaussian functions $g_j(t) = e^{-\lambda_j t^2}$ shows that $D = 1/\sqrt{F}$ where

$$F = \inf\left\{\frac{\det\left(\sum_{j=1}^{m} c_j\lambda_j u_j \otimes u_j\right)}{\prod_{j=1}^{m} \lambda_j^{c_j}} \mid \lambda_j > 0\right\}.$$

1.6.2. Barthe's proof

A reverse form of Theorem 1.6.1 was proved by Barthe.

THEOREM 1.6.2 (Barthe). *Let $m \geqslant n$, $c_1, \ldots, c_m > 0$ with $c_1 + \cdots + c_m = n$, and $u_1, \ldots, u_m \in \mathbb{R}^n$. If $h_1, \ldots, h_m : \mathbb{R} \to [0, \infty)$ are measurable functions, we set*

$$K(h_1, \ldots, h_m) = \int_{\mathbb{R}^n}^{*} \sup\left\{\prod_{j=1}^{m} h_j^{c_j}(\theta_j) \mid \theta_j \in \mathbb{R}\ ,\ x = \sum_{j=1}^{m} \theta_j c_j u_j\right\}dx,$$

where $\int^*$ denotes the outer integral. Then,

$$\inf\left\{K(h_1, \ldots, h_m) \mid \int_{\mathbb{R}} h_j = 1\ ,\ j = 1, \ldots, m\right\} = \sqrt{F}.$$

Although the Brascamp-Lieb inequality and its reverse form do not play a central role in this book, we sketch Barthe's argument which is very elegant, short and related in spirit with arguments that appear in subsequent chapters. A remarkable fact is that it gives a new direct proof of the Brascamp-Lieb inequality as well.

The first observation is that

$$\inf\left\{K(h_1,\dots,h_m)\mid \int_{\mathbb{R}} h_j = 1\ ,\ j=1,\dots,m\right\}\leqslant \sqrt{F}.$$

This follows by direct computation with Gaussian functions. The main step in Barthe's argument is the following proposition.

Proposition 1.6.3 (Barthe). *Let $f_1,\dots,f_m:\mathbb{R}\to[0,\infty)$ and $h_1,\dots,h_m:\mathbb{R}\to[0,\infty)$ be integrable functions with*

$$\int_{\mathbb{R}} f_j(t)\,dt=\int_{\mathbb{R}} h_j(t)\,dt=1,\qquad j=1,\dots,m.$$

Then,

$$F\cdot I(f_1,\dots,f_m)\leqslant K(h_1,\dots,h_m).$$

Proof. We may assume that f_j,h_j are continuous and strictly positive, and also that F is finite and positive. The key idea is to use a transportation of measure argument, which resembles the proof of the Prékopa-Leindler inequality in Section 1.2b. For every $j=1,\dots,m$ we define $T_j:\mathbb{R}\to\mathbb{R}$ by the equation

$$\int_{-\infty}^{T_j(t)} h_j(s)ds=\int_{-\infty}^{t} f_j(s)ds.$$

Observe that each T_j is strictly increasing, 1-1 and onto, and

$$T_j'(t)h_j(T_j(t))=f_j(t),\qquad t\in\mathbb{R}.$$

Then we define $W:\mathbb{R}^n\to\mathbb{R}^n$ by

$$W(y)=\sum_{j=1}^m c_jT_j(\langle y,u_j\rangle)u_j.$$

We check that $J_W(y)=\sum_{j=1}^m c_jT_j'(\langle y,u_j\rangle)u_j\otimes u_j$ and this implies that $\langle [J_W(y)](v),v\rangle>0$ if $v\neq 0$, which shows that W is injective. We define

$$m(x)=\sup\left\{\prod_{j=1}^m h_j^{c_j}(\theta_j)\mid x=\sum_{j=1}^m \theta_jc_ju_j\right\}.$$

It is clear that

$$m(W(y))\geqslant\prod_{j=1}^m h_j^{c_j}(T_j(\langle y,u_j\rangle))$$

for every $y\in\mathbb{R}^n$, and hence

$$\begin{aligned}\int_{\mathbb{R}^n} m(x)dx&\geqslant\int_{W(\mathbb{R}^n)} m(x)dx\\&=\int_{\mathbb{R}^n} m(W(y))\,|\det J_W(y)|\,dy\\&\geqslant\int_{\mathbb{R}^n}\prod_{j=1}^m h_j^{c_j}(T_j(\langle y,u_j\rangle))\det\left(\sum_{j=1}^m c_jT_j'(\langle y,u_j\rangle)u_j\otimes u_j\right)\,dy.\end{aligned}$$

By the definition of the constant F we have

$$\det\left(\sum_{j=1}^m c_j T_j'(\langle y,u_j\rangle)u_j\otimes u_j\right)\geqslant F\cdot\prod_{j=1}^m\left(T_j'(\langle y,u_j\rangle)\right)^{c_j}.$$

So, we can write

$$\begin{aligned}\int_{\mathbb{R}^n} m(x)dx &\geqslant F\cdot\int_{\mathbb{R}^n}\prod_{j=1}^m h_j^{c_j}(T_j(\langle y,u_j\rangle))\cdot\prod_{j=1}^m\left(T_j'(\langle y,u_j\rangle)\right)^{c_j}dy\\ &= F\cdot\int_{\mathbb{R}^n}\prod_{j=1}^m f_j^{c_j}(\langle y,u_j\rangle)dy\\ &= F\cdot I(f_1,\dots,f_m).\end{aligned}$$

This proves that $F\cdot I(f_1,\dots,f_m)\leqslant K(h_1,\dots,h_m)$. □

Proof of Theorems 1.6.1 and 1.6.2. We have

$$\sup\left\{I(f_1,\dots,f_m)\mid\int_{\mathbb{R}} f_j=1\ ,\ j=1,\dots,m\right\}\geqslant D=\frac{1}{\sqrt{F}}.$$

On the other hand, Proposition 1.6.3 shows that

$$\begin{aligned}\frac{1}{\sqrt{F}} &\leqslant \sup\left\{I(f_1,\dots,f_m)\mid\int_{\mathbb{R}} f_j=1\right\}\\ &\leqslant \frac{1}{F}\cdot\inf\left\{K(h_1,\dots,h_m)\mid\int_{\mathbb{R}} h_j=1\right\}\leqslant\frac{1}{\sqrt{F}}.\end{aligned}$$

Then, we must have equality everywhere, and this ends the proof. □

1.6.3. Reverse isoperimetric inequality

The calculation of the constant $F=F(\{u_j\},\{c_j\})$ in Theorems 1.6.1 and 1.6.2 is not an easy task. An important observation of Ball is that its value is equal to 1 if the vectors u_j satisfy John's representation of the identity, i.e. if they behave like an orthogonal basis with weights equal to c_j.

Theorem 1.6.4 (Ball). *Let $u_1,\dots,u_m\in S^{n-1}$ and $c_1,\dots,c_m>0$ such that*

$$I=\sum_{j=1}^m c_j u_j\otimes u_j.$$

Then, the constant $F=F(\{u_j\},\{c_j\})$ in Theorems 1.6.1 and 1.6.2 is equal to 1.

A well-known application of the Brascamp-Lieb inequality in this context is Ball's reverse isoperimetric inequality. We ask for the best constant $\partial(n)$ for which every symmetric convex body K in $\mathbb{R}^n$ has a position $\tilde{K}$ satisfying

$$\partial(\tilde{K})\leqslant\partial(n)|\tilde{K}|^{(n-1)/n}.$$

The natural position of K is the minimal surface area position which was discussed in Section 1.5d. However, Ball's solution of the problem employs John's position. Assume that B_2^n is the maximal volume ellipsoid of K. Then,

$$\partial(K)=\lim_{t\to 0^+}\frac{|K+tB_2^n|-|K|}{t}\leqslant\lim_{t\to 0^+}\frac{|K+tK|-|K|}{t}=n|K|.$$

Then, Ball proves that among all bodies in John's position the cube has maximal volume.

THEOREM 1.6.5. *Let $Q_n = [-1,1]^n$ be the unit cube in $\mathbb{R}^n$. If K is a symmetric convex body in $\mathbb{R}^n$, and if K is in John's position, then $|K| \leqslant 2^n = |Q_n|$.*

For the proof we use John's representation of the identity, where the u_j's are contact points of K and B_2^n. Observe that

$$K \subseteq M := \{x : |\langle x, u_j\rangle| \leqslant 1,\ j = 1, \ldots, m\}.$$

Therefore,

$$\begin{aligned}|K| \leqslant |M| &= \int_{\mathbb{R}^n} \prod_{j=1}^m \mathbf{1}_{[-1,1]}^{c_j}(\langle x, u_j\rangle)dx \\ &\leqslant \prod_{j=1}^m \left(\int_{\mathbb{R}} \mathbf{1}_{[-1,1]}(t)dt\right)^{c_j} = 2^{\sum_{j=1}^m c_j} = 2^n,\end{aligned}$$

where we used the Brascamp-Lieb inequality and applied Theorem 1.6.4 and the fact that $\sum_{j=1}^m c_j = n$. $\square$

Theorem 1.6.5 shows that $\partial(K) \leqslant n|K| \leqslant 2n|K|^{(n-1)/n}$, and since K was arbitrary, $\partial(n) \leqslant 2n$. The example of the cube shows that, actually, $\partial(n) = 2n$.

Theorem 1.6.5 shows that the cube has maximal volume ratio among all symmetric convex bodies. In the general case, one can show that the simplex Δ_n is the extremal convex body (this was also proved by Ball).

1.7. Concentration of measure

It was for the purposes of geometric functional analysis that concentration of measure was initially understood to be very useful and was developed as a method; the starting point was V. Milman's proof of Dvoretzky theorem. As it was soon realized, it could be and it has been very well adapted to the needs of probability theory, asymptotic combinatorics and complexity as well. In this section we introduce the main examples of metric probability spaces that will be used in this book.

1.7.1. Metric probability spaces

Let (X, d) be a metric space. If μ is a probability measure on the Borel σ-algebra $\mathcal{B}(X)$ of (X, d), then the triple (X, d, μ) is called a *metric probability space.*

Typical examples of metric probability spaces include:

1. *The Euclidean sphere* S^{n-1}. We consider the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ equipped with the geodesic metric ρ defined as follows: if $x, y \in S^{n-1}$ then $\rho(x, y)$ is the convex angle $\widehat{xoy}$ in the plane determined by the origin o and x, y. The sphere S^{n-1} becomes a probability space with the unique rotationally invariant measure σ: for any Borel set $A \subseteq S^{n-1}$ we set

$$\sigma(A) := \frac{|C(A)|}{|B_2^n|},$$

where

$$C(A) := \{sx : x \in A \text{ and } 0 \leqslant s \leqslant 1\}.$$

One can check that ρ is indeed a metric, and $\|x-y\|_2 = 2\sin\frac{\rho(x,y)}{2}$, therefore

$$\frac{2}{\pi}\rho(x,y) \leqslant \|x-y\|_2 \leqslant \rho(x,y).$$

2. *Gauss space.* We consider the measure γ_n on $\mathbb{R}^n$ with density

$$g_n(x) = (2\pi)^{-n/2}e^{-\|x\|_2^2/2}.$$

In other words, if A is a Borel subset of $\mathbb{R}^n$ then

$$\gamma_n(A) = \frac{1}{(2\pi)^{n/2}}\int_A e^{-\|x\|_2^2/2}dx.$$

The metric probability space $(\mathbb{R}^n, \|\cdot\|_2, \gamma_n)$ is called the n-dimensional Gauss space.

The standard Gaussian measure γ_n has two important properties: it is a product measure, more precisely $\gamma_n = \gamma_1 \otimes \cdots \otimes \gamma_1$, and it is invariant under orthogonal transformations, that is, if $U \in O(n)$ and A is a Borel subset of $\mathbb{R}^n$ then $\gamma_n(U(A)) = \gamma_n(A)$.

3. *The discrete cube.* We consider the set $E_2^n = \{-1,1\}^n$, which we identify with the set of vertices of the unit cube $Q_n = [-1,1]^n$ in $\mathbb{R}^n$. We equip E_2^n with the uniform probability measure μ_n which gives mass 2^{-n} to each point, and with the Hamming metric

$$d_n(x,y) = \frac{1}{n}\text{card}\{i \leqslant n : x_i \neq y_i\} = \frac{1}{2n}\sum_{i=1}^n |x_i - y_i|.$$

Definition 1.7.1. Let (X,d,μ) be a metric probability space. For each nonempty $A \in \mathcal{B}(X)$ and any $t > 0$, the *t-extension* of A is the set

$$A_t = \{x \in X : d(x,A) < t\}.$$

The *concentration function* of (X,d,μ) is defined on $(0,\infty)$ by

$$\alpha_\mu(t) := \sup\{1 - \mu(A_t) : \mu(A) \geqslant 1/2\}.$$

The function α_μ is obviously decreasing and one can check that for every metric probability space (X,d,μ) one has

$$\lim_{t\to\infty}\alpha_\mu(t) = 0.$$

Roughly speaking, we say that we have measure concentration on the metric probability space (X,d,μ) if $\alpha_\mu(t)$ decreases fast to zero as $t \to \infty$. More precisely:

1. We say that μ has *normal concentration* on (X,d) if there exist constants $C, c > 0$ such that, for every $t > 0$,

$$\alpha_\mu(t) \leqslant Ce^{-ct^2}.$$

We shall say that a family (X_n, d_n, μ_n) is a *Lévy family* if for any $t > 0$

$$\alpha_{\mu_n}(t \cdot \text{diam}(X_n)) \to_{n\to\infty} 0$$

and we shall say that it is a *normal Lévy family* with constants c, C if for any $t > 0$

$$\alpha_{\mu_n}(t) \leqslant Ce^{-ct^2 n}.$$

We can check that this is the case for the examples we mentioned above: the sphere, the discrete cube and the Gauss space. Indeed, in the terminology of Definition 1.7.1, the following estimates hold:

(i) For the sphere (S^{n-1}, ρ, σ) one has

$$\alpha_\sigma(t) \leqslant \sqrt{\pi/8}\exp(-t^2 n/2).$$

(ii) For the Gauss space $(\mathbb{R}^n, \|\cdot\|_2, \gamma_n)$ one has

$$\alpha_{\gamma_n}(t) \leqslant \frac{1}{2}\exp(-t^2/2).$$

(iii) For the discrete cube (E_2^n, d_n, μ_n) one has

$$\alpha_{\mu_n}(t) \leqslant \frac{1}{2}\exp(-2t^2 n).$$

In addition,

(iv) For the family of the orthogonal groups $(SO(n), \rho_n, \mu_n)$ equipped with the Hilbert-Schmidt metric and the Haar probability measure one has

$$\alpha_{\rho_n}(t) \leqslant \sqrt{\pi/8}\exp(-t^2 n/8).$$

(v) All homogeneous spaces of $SO(n)$ inherit the property of forming Lévy families. In particular, any family of Stiefel manifolds W_{n,k_n} or any family of Grassman manifolds G_{n,k_n}, $n = 1, 2, \ldots$, $1 \leqslant k_n \leqslant n$, is a Lévy family with the same constants as $SO(n)$.

2. We say that μ has *exponential concentration* on (X, d) if there exist constants $C, c > 0$ such that, for every $t > 0$,

$$\alpha_\mu(t) \leqslant Ce^{-ct}.$$

1.7.2. Concentration of measure and Lipschitz functions

Many of the applications of measure concentration follow directly from the next theorem.

THEOREM 1.7.2. *Let (X, d, μ) be a metric probability space. If $f : X \to \mathbb{R}$ is a Lipschitz function with constant* 1*, i.e.* $|f(x) - f(y)| \leqslant d(x, y)$ *for all* $x, y \in X$*, then*

$$\mu\left(\{x \in X : |f(x) - \mathrm{med}(f)| > t\}\right) \leqslant 2\alpha_\mu(t),$$

where $\mathrm{med}(f)$ *is a Lévy mean of* f*.*

Note. Given an arbitrary function $g : X \to \mathbb{R}$, a real number $\mathrm{med}(g)$ is called a Lévy mean of g if we have

$$\mu(\{g \geqslant \mathrm{med}(g)\}) \geqslant 1/2 \text{ and } \mu(\{g \leqslant \mathrm{med}(g)\}) \geqslant 1/2.$$

Observe that there are cases in which there are more than one numbers having this property.

Proof of Theorem 1.7.2. We set $A = \{x : f(x) \geqslant \mathrm{med}(f)\}$ and $B = \{x : f(x) \leqslant \mathrm{med}(f)\}$. For every $y \in A_t$ there exists $x \in A$ with $d(x, y) \leqslant t$, and thus

$$f(y) = f(y) - f(x) + f(x) \geqslant -d(y, x) + \mathrm{med}(f) \geqslant \mathrm{med}(f) - t$$

because f is 1-Lipschitz. Similarly, if $y \in B_t$ then there exists $x \in B$ with $d(x, y) \leqslant t$, and thus

$$f(y) = f(y) - f(x) + f(x) \leqslant d(y, x) + \mathrm{med}(f) \leqslant \mathrm{med}(f) + t.$$

It follows that if $y \in A_t \cap B_t$ then $|f(x) - \mathrm{med}(f)| \leqslant t$. In other words,

$$\{x \in X : |f(x) - \mathrm{med}(f)| > t\} \subseteq (A_t \cap B_t)^c = A_t^c \cup B_t^c.$$

From the definition of the concentration function we have $\mu(A_t) \geqslant 1 - \alpha_\mu(t)$ and $\mu(B_t) \geqslant 1 - \alpha_\mu(t)$. It follows that

$$\mu\left(\{|f - \mathrm{med}(f)| > t\}\right) \leqslant (1 - \mu(A_t)) + (1 - \mu(B_t)) \leqslant 2\alpha_\mu(t),$$

as claimed. □

When the concentration function of the space decreases fast, Theorem 1.7.2 shows that 1-Lipschitz functions are "almost constant" on "almost all of the space". We will use this fact very often, in the following form: if $f : X \to \mathbb{R}$ is a Lipschitz function with constant $\|f\|_{\mathrm{Lip}}$, i.e. if $|f(x) - f(y)| \leqslant \|f\|_{\mathrm{Lip}} d(x, y)$ for all $x, y \in X$, then

$$\mu\left(\{x \in X : |f(x) - \mathrm{med}(f)| > t\}\right) \leqslant 2\alpha_\mu(t/\|f\|_{\mathrm{Lip}}),$$

where $\mathrm{med}(f)$ is a Lévy mean of f.

Converse statements are also valid.

Theorem 1.7.3. *Let (X, d, μ) be a metric probability space. Assume that for some $\eta > 0$ and some $t > 0$ one has*

$$\mu\left(\{x \in X : |f(x) - \mathrm{med}(f)| > t\}\right) \leqslant \eta$$

for every 1-Lipschitz function $f : X \to \mathbb{R}$. Then, $\alpha_\mu(t) \leqslant \eta$.

Proof. Let A be a Borel subset of X with $\mu(A) \geqslant 1/2$. We consider the function $f(x) = d(x, A)$. Then, f is 1-Lipschitz and $\mathrm{med}(f) = 0$ because f is non-negative and $\mu(\{x : f(x) = 0\}) \geqslant 1/2$. From the assumption we get

$$\mu(\{x \in X : d(x, A) > t\}) \leqslant \eta,$$

that is $1 - \mu(A_t) \leqslant \eta$. It follows that $\alpha_\mu(t) \leqslant \eta$. □

The next theorem shows that we can draw the same conclusion if we replace the Lévy mean by the expectation; this is often easier to compute.

Theorem 1.7.4. *Let (X, d, μ) be a metric probability space. Assume that there is some function $\alpha : [0, \infty) \to [0, \infty)$ such that, for every bounded 1-Lipschitz function $f : (X, d) \to \mathbb{R}$ and for every $t > 0$, one has*

$$\mu\Big(\Big\{x : f(x) \geqslant \int f d\mu + t\Big\}\Big) \leqslant \alpha(t).$$

Then, for every Borel set $A \subseteq X$ with $\mu(A) > 0$ and for every $t > 0$,

$$1 - \mu(A_t) \leqslant \alpha\big(\mu(A)t\big).$$

In particular,

$$\alpha_\mu(t) \leqslant \alpha(t/2), \qquad t > 0.$$

Proof. We fix $A \in \mathcal{B}(X)$ with $\mu(A) > 0$ and $t > 0$. We consider the function $f(x) = \min\{d(x, A), t\}$. Note that $\|f\|_{\mathrm{Lip}} \leqslant 1$ and

$$\int f \, d\mu \leqslant (1 - \mu(A))t.$$

From the assumption we have

$$\begin{aligned} 1 - \mu(A_t) = \mu(\{f \geqslant t\}) &\leqslant \mu\Big(\Big\{x : f(x) \geqslant \int f \, d\mu + \mu(A)t\Big\}\Big) \\ &\leqslant \alpha\big(\mu(A)t\big). \end{aligned}$$

In particular, if $\mu(A) \geqslant 1/2$ then we have $\int f\,d\mu \leqslant t/2$ and this gives

$$\begin{aligned}1-\mu(A_t) = \mu(\{f \geqslant t\}) &\leqslant \mu\Big(\Big\{x : f(x) \geqslant \int f\,d\mu + t/2\Big\}\Big)\\ &\leqslant \alpha(t/2).\end{aligned}$$

We thus conclude that $\alpha_\mu(t) \leqslant \alpha(t/2)$. □

1.7.3. Isoperimetric problems and concentration of measure

In this subsection we discuss the isoperimetric problem, which can be formulated for an arbitrary metric probability space. In the next subsections we will see that for the typical examples of metric probability spaces that we saw in Subsection 1.7.1 the solution to the isoperimetric problem is known.

Definition 1.7.5. Let (X, d) be a metric space and let μ be a (not necessarily finite) measure on the Borel σ-algebra $\mathcal{B}(X)$. The surface area (or *Minkowski content*) of a non-empty $A \in \mathcal{B}(X)$ is defined by

$$\mu^+(A) = \liminf_{t\to 0^+} \frac{\mu(A_t \setminus A)}{t},$$

where A_t is the t-extension of A. If $\mu(A) < \infty$ (which is certainly true if (X, d, μ) is a metric probability space) then

$$\mu^+(A) = \liminf_{t\to 0^+} \frac{\mu(A_t) - \mu(A)}{t}.$$

Given a metric probability space, one can now state the isoperimetric problem in one of the following ways:

(i) Given $0 < \alpha < 1$ and $t > 0$, find

$$\inf\{\mu(A_t) : A \in \mathcal{B}(X), \mu(A) \geqslant \alpha\}$$

and identify (if they exist) those sets A at which this infimum is attained.

(ii) Given $0 < \alpha < 1$ and $t > 0$, find

$$\inf\{\mu^+(A) : A \in \mathcal{B}(X), \mu(A) \geqslant \alpha\}$$

and identify (if they exist) those sets A at which this infimum is attained.

Obviously, there is no reason why the answer to the first question shouldn't vary with t. Nevertheless, in all classical examples, the minimizers turn out to be independent of t and quite symmetric subsets of X, therefore we can easily compute the measure of their t-extension as well as their surface area. Note that the examples of spaces that we discuss next, for which the solution to the isoperimetric problem is known, are also the cases which will be important for this book.

1.7.4. The spherical isoperimetric inequality

The isoperimetric problem for S^{n-1} is formulated as follows. Let $\alpha \in (0, 1)$ and $t > 0$. Among all Borel subsets A of the sphere which satisfy $\sigma(A) = \alpha$, determine the ones for which the measure $\sigma(A_t)$ of their t-extension is minimal.

The answer to this question is given by the next theorem.

THEOREM 1.7.6 (Lévy). *Let $\alpha \in (0,1)$ and let $B(x,r)$ be a ball of radius $r > 0$ in S^{n-1} such that $\sigma(B(x,r)) = \alpha$. Then, for every $A \subseteq S^{n-1}$ with $\sigma(A) = \alpha$ and every $t > 0$, we have*

$$\sigma(A_t) \geqslant \sigma(B(x,r)_t) = \sigma(B(x,r+t)).$$

In other words, for any given value of α and any $t > 0$, the spherical caps of measure α provide the solution to the isoperimetric problem. A proof of the spherical isoperimetric inequality can be given with spherical symmetrization and induction on the dimension. Let us consider the special case $\alpha = 1/2$. If $\sigma(A) = 1/2$ and $t > 0$, then we can estimate the size of A_t using the isoperimetric inequality: from Theorem 1.7.6 we have

$$\sigma(A_t) \geqslant \sigma\left(B(x, \frac{\pi}{2} + t)\right) \tag{1.7.1}$$

for every $t > 0$ and $x \in S^{n-1}$. Then, starting from (1.7.1) and computing the measure of a cap $B(x, \pi/2 + t)$ we obtain the following inequality.

THEOREM 1.7.7. *Let $A \subseteq S^{n+1}$ with $\sigma(A) = 1/2$ and let $t > 0$. Then,*

$$\sigma(A_t) \geqslant 1 - \sqrt{\pi/8}\exp(-t^2 n/2). \tag{1.7.2}$$

The proof of Theorem 1.7.7 is heavily based on the spherical isoperimetric inequality. However, in the applications, we do not really need the precise solution to the isoperimetric problem; we only need an inequality providing a similar estimate to that in (1.7.2) (even with constants that are "much" larger than $\sqrt{\pi/8}$ and 2, as long as they are independent of the dimension). It turns out that one can give a very simple proof of an analogous exponential estimate using the Brunn-Minkowski inequality. The key point is the following result of Arias-de-Reyna, Ball and Villa.

LEMMA 1.7.8 (Arias de Reyna-Ball-Villa). *Consider the probability measure $\mu(A) = |A|/|B_2^n|$ on the Euclidean unit ball B_2^n. If A, B are subsets of B_2^n with $\mu(A) \geqslant \alpha$, $\mu(B) \geqslant \alpha$, and if $\rho(A,B) = \inf\{\|a-b\|_2 : a \in A, b \in B\} = \rho > 0$, then*

$$\alpha \leqslant \exp(-\rho^2 n/8).$$

Proof. We may assume that A and B are closed. By the Brunn-Minkowski inequality, $\mu(\frac{A+B}{2}) \geqslant \alpha$. On the other hand, the parallelogram law shows that if $a \in A, b \in B$ then

$$\|a+b\|_2^2 = 2\|a\|_2^2 + 2\|b\|_2^2 - \|a-b\|_2^2 \leqslant 4 - \rho^2.$$

It follows that $\frac{A+B}{2} \subseteq \sqrt{1 - \frac{\rho^2}{4}} B_2^n$, hence

$$\mu\left(\frac{A+B}{2}\right) \leqslant \left(1 - \frac{\rho^2}{4}\right)^{n/2} \leqslant \exp(-\rho^2 n/8).$$

Proof of Theorem 1.7.7. Assume that $A \subseteq S^{n-1}$ with $\sigma(A) = 1/2$. Let $t > 0$ and define $B = (A_t)^c \subseteq S^{n-1}$. We fix $\lambda \in (0,1)$ and consider the subsets

$$\tilde{A} = \bigcup\{tA : \lambda \leqslant t \leqslant 1\} \quad \text{and} \quad \tilde{B} = \bigcup\{tB : \lambda \leqslant t \leqslant 1\}$$

of B_2^n. These are disjoint with distance $\simeq \lambda t$. Lemma 1.7.8 shows that $\mu(\tilde{B}) \leqslant \exp(-c\lambda^2 t^2 n/8)$, and since $\mu(\tilde{B}) = (1-\lambda^n)\sigma(B)$ we obtain

$$\sigma(A_t) \geqslant 1 - \frac{1}{1-\lambda^n}\exp(-c\lambda^2 t^2 n/8).$$

We conclude the proof by choosing $\lambda = 1/2$. $\square$

From Theorem 1.7.2 it follows that if $g : S^{n-1} \to \mathbb{R}$ is Lipschitz continuous, then

$$\sigma(\{\theta : |g(\theta) - \mathrm{med}(g)| \geqslant t\}) \leqslant 2\exp(-(n-1)t^2/2\|g\|_{\mathrm{Lip}}^2)$$

for every $t > 0$. We will often use the fact that, in this deviation inequality, one can replace the Lévy mean of g by the expectation of g (for a proof of this assertion in the setting of the Gaussian measure, see Theorem 15.1.11 and Corollary 15.1.12 in Chapter 15).

THEOREM 1.7.9. *If $g : S^{n-1} \to \mathbb{R}$ is a Lipschitz continuous function, then*

$$\sigma(\{\theta : |g(\theta) - \mathbb{E}_\sigma(g)| \geqslant t\}) \leqslant 2\exp(-(n-1)t^2/2\|g\|_{\mathrm{Lip}}^2) \tag{1.7.3}$$

for all $t > 0$.

Another useful fact is that if $g : \mathbb{R}^n \to \mathbb{R}$ is a norm then the Lévy mean $\mathrm{med}(g)$ and the expectation $\mathbb{E}_\sigma(g)$ of g on the sphere S^{n-1} are comparable: one has

$$\frac{1}{2}\mathrm{med}(g) \leqslant \mathbb{E}_\sigma(g) \leqslant c\,\mathrm{med}(g)$$

where $c > 0$ is an absolute constant.

1.7.5. Isoperimetric inequality in the Gauss space

The isoperimetric inequality in the "Gauss space" is the following statement.

THEOREM 1.7.10 (Borell, Sudakov-Tsirelson). *Let $\alpha \in (0,1)$ and $\theta \in S^{n-1}$ and let $H = \{x \in \mathbb{R}^n : \langle x, \theta\rangle \leqslant \lambda\}$ be a half-space in $\mathbb{R}^n$ with $\gamma_n(H) = \alpha$. Then, for every $t > 0$ and every Borel $A \subseteq \mathbb{R}^n$ with $\gamma_n(A) = \alpha$, we have*

$$\gamma_n(A_t) \geqslant \gamma_n(H_t).$$

COROLLARY 1.7.11. *If $\gamma_n(A) \geqslant 1/2$ then, for every $t > 0$,*

$$1 - \gamma_n(A_t) \leqslant \frac{1}{2}\exp(-t^2/2). \tag{1.7.4}$$

Proof. From Theorem 1.7.10 we know that

$$1 - \gamma_n(A_t) \leqslant 1 - \gamma_n(H_t)$$

where H is a half-space of measure $1/2$. Since γ_n is invariant under orthogonal transformations, we may assume that $H = \{x \in \mathbb{R}^n : x_1 \leqslant 0\}$, and then it follows that

$$1 - \gamma_n(H_t) = \frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-s^2/2}ds.$$

Differentiation shows that the function

$$F(x) = e^{x^2/2}\int_x^\infty e^{-s^2/2}ds$$

is decreasing on $[0, +\infty)$. The fact that $F(t) \leqslant F(0)$ completes the proof. □

As in the case of the sphere, the proof of the approximate isoperimetric inequality (1.7.4) requires knowing the exact solution of the Gaussian isoperimetric problem. However, there are also simple, direct proofs which do not assume the isoperimetric inequality in the Gauss space; an example is the following proof by Maurey that makes use of the Prékopa-Leindler inequality.

THEOREM 1.7.12 (Maurey). *Let A be a non-empty Borel subset of $\mathbb{R}^n$. Then,*

$$\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \leqslant \frac{1}{\gamma_n(A)},$$

where $d(x,A) = \inf\{\|x-y\|_2 : y \in A\}$. Therefore, if $\gamma_n(A) \geqslant \frac{1}{2}$ we have

$$1 - \gamma_n(A_t) \leqslant 2\exp(-t^2/4)$$

for every $t > 0$.

Proof. Consider the functions

$$f(x) = e^{d(x,A)^2/4} g_n(x)\ , \quad h(x) = \mathbf{1}_A(x) g_n(x)\ , \quad m(x) = g_n(x),$$

where g_n is the density of the gaussian measure γ_n. For every $x \in \mathbb{R}^n$ and $y \in A$ we see that

$$f(x)h(y) \leqslant \Big(m\Big(\frac{x+y}{2}\Big)\Big)^2,$$

using the parallelogram law and the fact that $d(x,A) \leqslant \|x-y\|_2$. Since $h(y) = 0$ whenever $y \notin A$, this implies that f, h, m satisfy the assumptions of the Prékopa-Leindler inequality with $\lambda = 1/2$. Therefore,

$$\left(\int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x)\right)\gamma_n(A) = \left(\int_{\mathbb{R}^n} f\right)\left(\int_{\mathbb{R}^n} h\right) \leqslant \left(\int_{\mathbb{R}^n} m\right)^2 = 1.$$

This proves the first assertion of the theorem. For the second one, observe that if $\gamma_n(A) \geqslant \frac{1}{2}$ then

$$e^{t^2/4}\gamma_n\big(\{x : d(x,A) \geqslant t\}\big) \leqslant \int_{\mathbb{R}^n} e^{d(x,A)^2/4} d\gamma_n(x) \leqslant \frac{1}{\gamma_n(A)} \leqslant 2.$$

This shows that $\gamma_n(A_t^c) \leqslant 2\exp(-t^2/4)$. □

1.7.6. Isoperimetric inequality in the discrete cube

The solution to the isoperimetric problem for E_2^n is given by by the d_n-balls (the so-called Hamming balls of E_2^n) in the case where N is the cardinality of some d_n ball. A combinatorial proof of this fact was given by Harper (the general minimizers for the isoperimetric problem are also known). Based on this information one can give an estimate for the concentration function of E_2^n.

THEOREM 1.7.13 (Harper). *If $\mu_n(A) \geqslant 1/2$ and $t > 0$, then*

$$\mu_n(A_t^c) \leqslant \frac{1}{2}\exp(-2t^2 n).$$

We will present a direct proof of an only slightly worse exponential estimate in Theorem 1.7.16. The proof is based on the following theorem of Talagrand.

THEOREM 1.7.14 (Talagrand). *Let A be a non-empty subset of E_2^n. We consider its convex hull* conv(A) *and for every $x \in E_2^n$ we define*

$$\phi_A(x) = \min\{\|x-y\|_2 : y \in \mathrm{conv}(A)\}.$$

Then,

$$\int_{E_2^n} \exp(\phi_A^2(x)/8) d\mu_n(x) \leqslant \frac{1}{\mu_n(A)}.$$

For every non-empty subset A of E_2^n, the function ϕ_A of Theorem 1.7.14 and the function

$$d_n(x,A) = \min\left\{\frac{1}{2n}\sum_{i=1}^n |x_i - y_i| : y \in A\right\}$$

which measures the distance from x to A are related as follows.

LEMMA 1.7.15. *For every non-empty $A \subseteq E_2^n$ and every $x \in E_2^n$,*

$$2\sqrt{n}\, d_n(x,A) \leqslant \phi_A(x).$$

Proof. Let $x \in E_2^n$. For every $y \in A$ we have

$$\langle x - y, x\rangle = \sum_{i=1}^n x_i(x_i - y_i) = 2n d_n(x,y) \geqslant 2n d_n(x,A). \tag{1.7.5}$$

From (1.7.5) we see that, for every $y \in \mathrm{conv}(A)$,

$$\sqrt{n}\|x-y\|_2 \geqslant \langle x-y, x\rangle \geqslant 2n d_n(x,A).$$

This proves the lemma. □

Combining the above we get the approximate isoperimetric inequality for E_2^n:

THEOREM 1.7.16. *Let $A \subseteq E_2^n$ with $\mu_n(A) \geqslant 1/2$. Then, for every $t > 0$, we have*

$$\mu_n(A_t) \geqslant 1 - 2\exp(-t^2 n/2).$$

Proof. If $x \notin A_t$, then $d_n(x,A) \geqslant t$ and Lemma 1.7.15 shows that $\phi_A(x) \geqslant 2t\sqrt{n}$. But, from Theorem 1.7.14 we have

$$e^{t^2n/2}\mu_n\big(\{x : \phi_A(x) \geqslant 2t\sqrt{n}\}\big) \leqslant \int_{E_2^n} \exp(\phi_A^2(x)/8)\, d\mu_n(x) \leqslant \frac{1}{\mu_n(A)} \leqslant 2,$$

and this gives

$$\mu_n(A_t^c) \leqslant \mu_n\big(\{x : \phi_A(x) \geqslant 2t\sqrt{n}\}\big) \leqslant 2\exp(-t^2n/2),$$

whence we are done. □

1.7.7. Kahane-Khintchine inequality

The Rademacher functions $r_i : [0,1] \to \mathbb{R}$, $i \geqslant 1$, are defined by

$$r_i(t) = \mathrm{sign}\, \sin(\pi 2^i t/2).$$

They are ± 1-valued (if we ignore a set of measure zero) independent random variables on $[0,1]$ and they form an orthonormal sequence in $L^2[0,1]$. An equivalent way to define them is to consider $E_2 = \prod_{i=1}^\infty \{-1,1\}$ endowed with the standard product measure and to define, for every $\epsilon = (\epsilon_i)_{i=1}^\infty$,

$$r_i(\epsilon) = \epsilon_i.$$

The classical Khintchine inequality states that for every $p > 0$ there exist constants $A_p, B_p > 0$ such that, for every $n \geqslant 1$ and any n-tuple of real numbers $a_1, \ldots, a_n$,

$$A_p\Big(\sum_{i=1}^n a_i^2\Big)^{1/2} \leqslant \Big(\int_{E_2^n}\Big|\sum_{i=1}^n a_i\epsilon_i\Big|^p d\mu_n(\epsilon)\Big)^{1/p} \leqslant B_p\Big(\sum_{i=1}^n a_i^2\Big)^{1/2}. \tag{1.7.6}$$

Since

$$\Big(\sum_{i=1}^{n} a_i^2\Big)^{1/2} = \Big(\int_{E_2^n} \Big|\sum_{i=1}^{n} a_i\epsilon_i\Big|^2 d\mu_n(\epsilon)\Big)^{1/2},$$

an equivalent way to state Khintchine's inequality is the following.

THEOREM 1.7.17 (Khintchine). *For every $p > 0$ there exist $A_p, B_p > 0$ such that for every $n \geqslant 1$ and any $a = (a_1, \dots, a_n) \in \ell_2^n$,*

$$A_p\Big\|\sum_{i=1}^{n} a_i\epsilon_i\Big\|_{L_2(E_2^n)} \leqslant \Big\|\sum_{i=1}^{n} a_i\epsilon_i\Big\|_{L_p(E_2^n)} \leqslant B_p\Big\|\sum_{i=1}^{n} a_i\epsilon_i\Big\|_{L_2(E_2^n)}. \tag{1.7.7}$$

Let A_p^*, B_p^* denote the best constants for which the statement of Theorem 1.7.17 is valid. From Hölder's inequality it is clear that $A_p^* = 1$ if $p \geqslant 2$ and $B_p^* = 1$ if $0 < p \leqslant 2$. The exact values of A_p^* and B_p^* have been determined by Szarek ($A_1^* = 1/\sqrt{2}$) and Haagerup (for all p). What is particularly important is to know the behavior of B_p^* as $p \to \infty$; the order of growth of B_p^* is $O(\sqrt{p})$ for large p.

Kahane's inequality generalizes Khintchine's inequality.

THEOREM 1.7.18 (Kahane). *There exists $C > 0$ such that for every normed space X, for any $n \geqslant 1$, for any $x_1, \dots, x_n \in X$ and any $p \geqslant 1$,*

$$\Big(\mathbb{E}\Big\|\sum_{i=1}^{n} \epsilon_i x_i\Big\|^p\Big)^{1/p} \leqslant C\sqrt{p}\,\mathbb{E}\Big\|\sum_{i=1}^{n} \epsilon_i x_i\Big\|. \tag{1.7.8}$$

1.8. Entropy estimates

1.8.1. Covering numbers

Let A and B be two convex bodies in $\mathbb{R}^n$. The *covering number* $N(A, B)$ of A by B is the least number of translates of B that are needed in order to cover A:

$$N(A, B) = \min\{N \in \mathbb{N} \mid \exists\, x_1, \dots, x_N \in \mathbb{R}^n : A \subseteq \cup(x_i + B)\}.$$

Sometimes, we require that the centers x_i belong to A; then we set

$$\overline{N}(A, B) = \min\{N \in \mathbb{N} \mid \exists\, x_1, \dots, x_N \in A : A \subseteq \cup(x_i + B)\}.$$

Some of the basic properties of covering numbers are listed below; we will be using them very often in this book:

(i) For all convex A, B and every invertible linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$ we have $N(A, B) = N(T(A), T(B))$.
(ii) For all convex A, B and C we have $N(A, B) \leqslant N(A, C)N(C, B)$.
(iii) For all convex A, B and C we have $N(A + C, B + C) \leqslant N(A, B)$.
(iv) For all convex A and B we have $\overline{N}(A, (A - A) \cap B) = \overline{N}(A, B)$. In particular we have that: for centrally symmetric A, $\overline{N}(A, 2A \cap B) = \overline{N}(A, B)$ and $N(A, 2(B \cap A)) \leq N(A, B)$.
(v) For all convex A we have $N(A, RB_2^n \cap (A - A)) = N(A, RB_2^n)$.

The *packing number* $M(A, B)$ of two centrally symmetric convex bodies A and B is the maximal cardinality of a B-separated set in A:

$$M(A, B) = \max\{N : \exists\, x_1, \dots, x_N \in A \text{ s.t. } \forall j \neq i, x_j \notin (x_i + B)\}.$$

Equivalently, we ask that if $i \neq j$ then $\|x_i - x_j\|_B > 1$. The packing number is closely related to the covering number of A by B: one can check that

$$M(A, 2B) \leqslant N(A, B) \leqslant \overline{N}(A, B) \leqslant M(A, B).$$

1.8.2. Sudakov inequality and its dual

Let K be a symmetric convex body in $\mathbb{R}^n$. It will be very useful to have sharp estimates for the covering numbers $N(K, tB_2^n)$ and $N(B_2^n, tK)$. A well-known bound for $N(K, tB_2^n)$ is given by *Sudakov inequality* (see [**481**]).

THEOREM 1.8.1 (Sudakov). *Let K be a symmetric convex body in $\mathbb{R}^n$. For every $t > 0$,*

$$\log N(K, tB_2^n) \leqslant cn \left(\frac{w(K)}{t}\right)^2,$$

where $c > 0$ is an absolute constant.

The simplest way to prove Sudakov inequality is from the so-called *dual Sudakov inequality* which was proved by Pajor and Tomczak-Jaegermann.

THEOREM 1.8.2 (Pajor-Tomczak). *Let K be a symmetric convex body in $\mathbb{R}^n$. For every $t > 0$,*

$$\log N(B_2^n, tK) \leqslant cn \left(\frac{w(K^\circ)}{t}\right)^2,$$

where $c > 0$ is an absolute constant.

We present a simple proof of Theorem 1.8.2 which is due to Talagrand; his argument makes use of the Gaussian measure γ_n. We need a simple lemma.

LEMMA 1.8.3. *Let K be a symmetric convex body in $\mathbb{R}^n$. For every $z \in \mathbb{R}^n$ we have $\gamma_n(K + z) \geqslant \exp(-\|z\|_2^2/2)\gamma_n(K)$.* □

Proof of Theorem 1.8.2. Let $x_1, \ldots, x_N$ be a maximal (tK)-separated set of points in B_2^n. Then, the sets $x_i + \frac{t}{2}K$ have disjoint interiors, and hence, for every $\lambda > 0$ the sets $\lambda x_i + \frac{\lambda t}{2}K$ have disjoint interiors. Since γ_n is a probability measure, we have

$$\sum_{i=1}^N \gamma_n\Big(\lambda x_i + \frac{\lambda t}{2}K\Big) = \gamma_n\Big(\bigcup_{i=1}^N \Big(\lambda x_i + \frac{\lambda t}{2}K\Big)\Big) \leqslant 1.$$

Note that $\|\lambda x_i\|_2 \leqslant \lambda$; then, Lemma 1.8.3 shows that

$$\gamma_n\Big(\lambda x_i + \frac{\lambda t}{2}K\Big) \geqslant \exp(-\lambda^2/2)\gamma_n\Big(\frac{\lambda t}{2}K\Big) \qquad i = 1, \ldots, N.$$

Consequently, for every $\lambda > 0$ we have that

$$N(B_2^n, tK) \leqslant N \leqslant \frac{\exp(\lambda^2/2)}{\gamma_n\left(\frac{\lambda t}{2}K\right)}.$$

It remains to choose $\lambda > 0$ in an optimal way: integration in polar coordinates shows that

$$\int_{\mathbb{R}^n} \|x\|_K \gamma_n(dx) \leqslant c\sqrt{n} \int_{S^{n-1}} \|\theta\|_K d\sigma(\theta) = c\sqrt{n}w(K^\circ),$$

and applying Markov's inequality we get

$$\gamma_n(\|x\|_K \geqslant \lambda t/2) \leqslant \frac{2}{\lambda t} \int_{\mathbb{R}^n} \|x\|_K \gamma_n(dx) \leqslant \frac{2c\sqrt{n}}{\lambda t} w(K^\circ),$$

or equivalently,

$$1-\gamma_n\Big(\frac{\lambda t}{2}K\Big)\leqslant \frac{2c\sqrt{n}}{\lambda t}w(K^\circ).$$

For $\lambda = 4c\sqrt{n}w(K^\circ)/t$ we get

$$\gamma_n\left(2c\sqrt{n}w(K^\circ)K\right)\geqslant \frac{1}{2},$$

and hence

$$N(B_2^n,tK)\leqslant 2\exp\left(8c^2nw^2(K^\circ)/t^2\right),$$

as claimed. □

Tomczak-Jaegermann observed that one can deduce Theorem 1.8.1 from Theorem 1.8.2 and vice versa.

THEOREM 1.8.4. *Let K be a symmetric convex body in $\mathbb{R}^n$. If*

$$B=\sup_{t>0}\, t\left(\log N(K,tB_2^n)\right)^{1/2},$$

and

$$A:=\sup_{t>0}\, t\left(\log \overline{N}(B_2^n,tK^\circ)\right)^{1/2}\leqslant c\sqrt{n}w(K),$$

then $B\leqslant 10A$. In particular, we have Sudakov inequality: for every $t>0$,

$$\log N(K,tB_2^n)\leqslant cn\left(\frac{w(K)}{t}\right)^2,$$

where $c>0$ is an absolute constant.

Proof. We first observe that $2K\cap\left(\frac{t^2}{2}K^\circ\right)\subseteq tB_2^n$. It follows that

$$\overline{N}(K,tB_2^n)\leqslant \overline{N}\Big(K,(2K)\cap(\frac{t^2}{2}K^\circ)\Big).$$

Next, we observe that

$$\overline{N}\Big(K,(2K)\cap(\frac{t^2}{2}K^\circ)\Big)=\overline{N}\Big(K,\frac{t^2}{2}K^\circ\Big).$$

Using the above we see that

$$\begin{aligned}N(K,tB_2^n)&\leqslant \overline{N}(K,tB_2^n)\leqslant \overline{N}\Big(K,\tfrac{t^2}{2}K^\circ\Big)\\ &\leqslant N(K,\tfrac{t^2}{4}K^\circ)\leqslant N(K,2tB_2^n)N\Big(B_2^n,\tfrac{t}{8}K^\circ\Big).\end{aligned}$$

We write

$$\begin{aligned}t^2\log N(K,tB_2^n)&\leqslant \frac{1}{4}(2t)^2\log N(K,2tB_2^n)+64(t/8)^2\log N\Big(B_2^n,\frac{t}{8}K^\circ\Big)\\ &\leqslant \frac{1}{4}(2t)^2\log N(K,2tB_2^n)+64A^2,\end{aligned}$$

and taking sup over all $t>0$ we arrive at $3B^2\leqslant 256A^2$.

We can now prove Sudakov inequality: for every $t>0$ we have

$$t^2\log N(K,tB_2^n)\leqslant 100A^2\leqslant cnw^2(K),$$

where $c>0$ is an absolute constant. □

Theorem 1.8.1 holds true for not necessarily symmetric convex bodies as well.

PROPOSITION 1.8.5. *Let K be a convex body in $\mathbb{R}^n$. For every $t > 0$,*

$$N(K, tB_2^n) \leqslant \exp\big(cn(w(K)/t)^2\big),$$

where $c > 0$ is an absolute constant.

Proof. Consider the difference body $K - K$ of K. Then,

$$\begin{aligned} w(K-K) &= \int_{S^{n-1}} h_{K-K}(u)d\sigma(u) = \int_{S^{n-1}} [h_K(u) + h_{-K}(u)]d\sigma(u) \\ &= \int_{S^{n-1}} [h_K(u) + h_K(-u)]d\sigma(u) = 2w(K). \end{aligned}$$

Since there is a translate of K which is contained in $K - K$, Sudakov's inequality gives

$$t^2 \log N(K, tB_2^n) \leqslant t^2 \log N(K-K, tB_2^n) \leqslant cnw^2(K-K) = 4cnw^2(K),$$

which proves the theorem. □

1.8.3. Duality of entropy

The duality of entropy numbers conjecture asserts that if X, Y are Banach spaces, if $T : X \to Y$ is a compact operator and if $N(T, t)$ denotes the covering number $N(T(B_X), tB_Y)$, then

$$a^{-1} \log N(T, bt) \leqslant \log N(T^*, t) \leqslant a \log N(T, b^{-1}t)$$

for every $t > 0$, where $a, b > 0$ are absolute constants, and T^* is the adjoint operator of T. This conjecture has been verified only under strong assumptions for both spaces X and Y. In the case where one of the two spaces is a Hilbert space, the conjecture is equivalent to the following statement about covering numbers of convex bodies: There exist two constants $\alpha, \beta > 0$ such that

$$\log N(B_2^n, \beta K^\circ) \leqslant \alpha \log N(K, B_2^n)$$

for every symmetric convex body K in $\mathbb{R}^n$.

This case was settled by Artstein-Avidan, Milman and Szarek.

THEOREM 1.8.6 (Artstein-Milman-Szarek). *There exist two absolute constants α and $\beta > 0$ such that for any dimension n and any symmetric convex body K in $\mathbb{R}^n$, one has*

$$N(B_2^n, \alpha^{-1}K^\circ)^{1/\beta} \leqslant N(K, B_2^n) \leqslant N(B_2^n, \alpha K^\circ)^\beta$$

Theorem 1.8.6 establishes a strong connection between the geometry of a set and its polar. Observe that since the theorem is true for any K, we actually have that, for any $t > 0$,

$$\beta^{-1} \log N(B_2^n, \alpha^{-1}tK^\circ) \leqslant \log N(K, tB_2^n) \leqslant \beta \log N(B_2^n, \alpha tK^\circ).$$

A weaker but general duality inequality has been proved by König and Milman. Using the reverse Santaló and Brunn-Minkowski inequalities they showed that

$$c^{-1} N(K_2^\circ, K_1^\circ)^{1/n} \leqslant N(K_1, K_2)^{1/n} \leqslant cN(K_2^\circ, K_1^\circ)^{1/n} \tag{1.8.1}$$

for every pair of symmetric convex bodies K_1 and K_2 in $\mathbb{R}^n$.

1.9. Gaussian and sub-Gaussian processes

1.9.1. Sub-Gaussian processes

Let (T,d) be a metric space and let $\mathcal{Y}=(Y_t)_{t\in T}$ be a family of real valued random variables, with indices from T, on a probability space $(\Omega,\mathcal{A},P)$. We say that the process $\mathcal{Y}=(Y_t)_{t\in T}$ is *sub-Gaussian* with respect to d if $\mathbb{E}(Y_t)=0$ for all $t\in T$ and, for all $t,s\in T$ and every $u>0$,

$$\mathbb{P}\left(|Y_t-Y_s|\geqslant u\right)\leqslant 2\exp\Big(-\frac{u^2}{d^2(t,s)}\Big).$$

A typical example is given by the discrete cube $E_2^n=\{-1,1\}^n$, equipped with the uniform probability measure. Write $\epsilon=(\epsilon_1,\dots,\epsilon_n)$ for the points of E_2^n and consider the *Rademacher functions* $r_i:E_2^n\to\{-1,1\}$, $1\leqslant i\leqslant n$, defined by $r_i(\epsilon)=\epsilon_i$.

For every $t=(t_1,\dots,t_n)\in T\subseteq\mathbb{R}^n$ we define

$$Y_t(\epsilon)=\langle t,\epsilon\rangle=t_1r_1(\epsilon)+\cdots+t_nr_n(\epsilon).$$

From Khintchine's inequality it follows that

$$\mathbb{P}\left(\epsilon\in E_2^n:|t_1r_1(\epsilon)+\cdots+t_nr_n(\epsilon)|\geqslant u\right)\leqslant 2\exp\big(-u^2/2(t_1^2+\cdots+t_n^2)\big)$$

for every $u>0$. This shows that $\mathcal{Y}=(Y_t)_{t\in\mathbb{R}^n}$ is sub-Gaussian with respect to the Euclidean metric.

A second example is given by Gaussian processes. We write g for a standard Gaussian random variable and $G=(g_1,\dots,g_n)$ for the standard Gaussian random vector in $\mathbb{R}^n$. The distribution of G is the Gaussian measure γ_n, with density $(2\pi)^{-n/2}\exp(-\|x\|_2^2/2)$.

Let T be a non-empty set. A family $\mathcal{Z}=(Z_t)_{t\in T}$ of real valued random variables on $(\Omega,\mathcal{A},P)$ is called a *Gaussian process* if, for any $a_1,\dots,a_m\in\mathbb{R}$ and any $Z_{t_1},\dots,Z_{t_m}\in\mathcal{Z}$, the linear combination $a_1Z_{t_1}+\cdots+a_mZ_{t_m}$ is a Gaussian random variable with mean 0. We may view $\mathcal{Z}$ as a subset of $L^2(\Omega)$, and then it induces on T the metric

$$d(t,s)=\|Z_t-Z_s\|_{L^2(\Omega)}.$$

By the definition of a Gaussian process, for every $t,s\in T$, Z_t-Z_s is a Gaussian random variable with mean 0 and variance $\mathbb{E}(Z_t-Z_s)^2=d^2(t,s)$. Consequently, for every $u>0$ we have

$$\mathbb{P}\left(|Z_t-Z_s|\geqslant u\right)=\frac{2}{d(t,s)\sqrt{2\pi}}\int_u^\infty\exp\Big(-\frac{r^2}{2d^2(t,s)}\Big)\,dr\leqslant 2\exp\Big(-\frac{u^2}{d^2(t,s)}\Big),$$

which implies that $\mathcal{Z}$ is sub-Gaussian with respect to the metric d it induces to T.

EXAMPLES 1.9.1. (i) If $g_1,\dots,g_N$ are independent standard Gaussian random variables on $(\Omega,\mathcal{A},P)$, then $\mathcal{Z}=\{g_1,\dots,g_N\}$ is a Gaussian process.

(ii) Consider n independent standard Gaussian random variables $g_1,\dots,g_n$. For every non-empty $T\subseteq\mathbb{R}^n$ we define a process $\mathcal{Z}=(Z_t)_{t\in T}$, by

$$Z_t(\omega)=\langle t,G(\omega)\rangle=\Big\langle t,\sum_{i=1}^n g_i(\omega)e_i\Big\rangle=\sum_{i=1}^n\langle t,e_i\rangle g_i(\omega),$$

where $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $\mathbb{R}^n$ and $G = (g_1, \ldots, g_n)$. Then, $\mathcal{Z}$ is a Gaussian process and the induced metric is the Euclidean metric on $\mathbb{R}^n$: for all $t, s \in T$,

$$d(t,s) = \|Z_t - Z_s\|_{L^2(\Omega)} = \|t-s\|_2.$$

Definition 1.9.2. Let (T,d) be a metric space and let $\mathcal{Y} = (Y_t)_{t\in T}$ be a sub-Gaussian process with respect to d. We define

$$\mathbb{E}\sup_{t\in T} Y_t = \sup\big\{\mathbb{E}\max_{t\in F} Y_t : F \subseteq T, 0 < |F| < \infty\big\}.$$

An important question is to obtain sharp upper bounds for the expectation $\mathbb{E}\sup Y_t$ in terms of the geometry of (T,d); in the next two subsections we discuss a number of important related results, which will be used several times in this book.

Remark 1.9.3. A basic observation, which connects this question with convex geometric analysis, is that if we consider a convex body K in $\mathbb{R}^n$ and the Gaussian process $\mathcal{Z} = (Z_t)_{t\in K}$, where $Z_t(\omega) = \langle t, G(\omega)\rangle$, that was defined in Example 1.9.1 (ii), then

(1.9.1)

$$\begin{aligned}\mathbb{E}\sup_{t\in K} Z_t &= \mathbb{E}\sup_{t\in K}\langle t, G\rangle = \mathbb{E}\, h_K(G)\\ &= \frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n} h_K(x)e^{-|x|^2/2}dx \simeq \sqrt{n}\int_{S^{n-1}} h_K(\theta)\sigma(d\theta) \simeq \sqrt{n}w(K).\end{aligned}$$

We just used the fact that the distribution of G is the standard Gaussian measure on $\mathbb{R}^n$ and integration in polar coordinates.

1.9.2. Metric entropy — the case of Gaussian processes

Let (T,d) be a metric space. For every $\varepsilon > 0$ we define

$$N(T,d,\varepsilon) = \min\left\{N : \text{ there exist } t_1, \ldots, t_N \in T : T \subseteq \bigcup_{i=1}^N B(t_i,\varepsilon)\right\},$$

where $B(t,\varepsilon) = \{s \in T : d(t,s) < \varepsilon\}$. The function $\varepsilon \mapsto \log N(T,d,\varepsilon)$ is the *metric entropy function* of T.

Consider as an example the Gaussian process $\mathcal{Z} = \{g_1, \ldots, g_N\}$, $N \geqslant 2$. We easily check that $\|g_i - g_j\|_2 = \sqrt{2}$ if $i \neq j$, and hence, $N(\varepsilon) = N$ if $0 < \varepsilon \leqslant \sqrt{2}$ and $N(\varepsilon) = 1$ if $\varepsilon > \sqrt{2}$. Also, using the fact that g_i are independent we may check that

$$\mathbb{E}\max_{1\leqslant i\leqslant N} g_i \simeq \sqrt{\log N}.$$

Let $\mathcal{Z} = (Z_t)_{t\in T}$ be a Gaussian process. We view T as a metric space with the induced metric d. The next theorem gives upper and lower bounds for $\mathbb{E}\sup Z_t$ in terms of the metric entropy function of (T,d).

Theorem 1.9.4 (Sudakov-Dudley). *There exist constants $c_1, c_2 > 0$ with the following property: if $\mathcal{Z} = (Z_t)_{t\in T}$ is a Gaussian process and d is the induced metric, then*

$$c_1\sup_{\varepsilon>0}\varepsilon\sqrt{\log N(T,d,\varepsilon)} \leqslant \mathbb{E}\sup_{t\in T} Z_t \leqslant c_2\int_0^\infty \sqrt{\log N(T,d,\varepsilon)}d\varepsilon.$$

The left hand side inequality is Sudakov's inequality, while the right hand side inequality is Dudley's inequality. In the example of $\mathcal{Z} = \{g_1, \dots, g_N\}$, both bounds give the right order of $\mathbb{E}\sup g_i$. The proof of Sudakov's inequality is based on a classical comparison lemma of Slepian. We will discuss Dudley's inequality in the more general context of sub-Gaussian processes.

THEOREM 1.9.5 (Slepian). *If* $(X_1, \dots, X_N)$ *and* $(Y_1, \dots, Y_N)$ *are two* N*-tuples of Gaussian random variables with mean* 0 *which satisfy the condition*

$$\|X_i - X_j\|_2 \leqslant \|Y_i - Y_j\|_2$$

for all $i \neq j$, *then*

$$\mathbb{E}\max_{i\leqslant N} X_i \leqslant \mathbb{E}\max_{i\leqslant N} Y_i.$$

Proof of Sudakov's inequality. We use Slepian's lemma as follows: Let $\mathcal{Z} = (Z_t)_{t\in T}$ be a Gaussian process and let d be the induced metric. Given $\varepsilon > 0$ we consider a subset $\{t_1, \dots, t_N\}$ of T which is maximal with respect to the condition "$d(t,s) \geqslant \varepsilon$ if $t \neq s$". Then $T \subseteq \bigcup_{i=1}^N B(t_i, \varepsilon)$, which implies $N(T, d, \varepsilon) \leqslant N$.

If $\delta = \min \|Z_{t_i} - Z_{t_j}\|_2$, we consider the N-tuple $\big(\frac{\delta g_1}{\sqrt{2}}, \dots, \frac{\delta g_N}{\sqrt{2}}\big)$, where g_i are independent standard Gaussian random variables. If $i \neq j$ then

$$\left\|\frac{\delta g_i}{\sqrt{2}} - \frac{\delta g_j}{\sqrt{2}}\right\|_2 = \delta \leqslant \|Z_{t_i} - Z_{t_j}\|_2,$$

so we can apply Slepian's lemma. It follows that

$$\mathbb{E}\sup_{t\in T} Z_t \geqslant \mathbb{E}\max_{i\leqslant N} Z_{t_i} \geqslant \frac{\delta}{\sqrt{2}}\mathbb{E}\max_{i\leqslant N} g_i \geqslant c_1\varepsilon\sqrt{\log N}.$$

Thus, $\mathbb{E}\sup_{t\in T} Z_t \geqslant c_1 \sup_{\varepsilon>0} \varepsilon\sqrt{\log N(T,d,\varepsilon)}$. □

1.9.3. Dudley's bound for sub-Gaussian processes

Dudley's inequality is more generally valid for sub-Gaussian processes. The proof uses a successive approximation argument which we briefly describe:

Proof of Dudley's inequality. We assume that (T, d) is a metric space and $\mathcal{Y} = (Y_t)_{t\in T}$ is a process such that $\mathbb{E}(Y_t) = 0$ for all $t \in T$ and, for all $t, s \in T$ and every $u > 0$,

$$\mathbb{P}\left(|Y_t - Y_s| \geqslant u\right) \leqslant 2\exp\Big(-\frac{u^2}{d^2(t,s)}\Big).$$

We consider a non-empty finite subset F of T and fix $t_0 \in F$. We set $R = \max\{d(t, t_0) : t \in F\}$ and $r_k = R/2^k$ for all $k \geqslant 0$.

We define $A_0 = \{t_0\}$ and for every $k \geqslant 1$ we find $A_k \subseteq F$ with cardinality $|A_k| = N(F, d, r_k)$ such that $F \subseteq \bigcup_{t\in A_k} B(t, r_k)$. Finally, for every $t \in F$ and $k \geqslant 0$ we choose $\pi_k(t) \in A_k$ with the property $d(t, \pi_k(t)) \leqslant r_k$. Since F is finite, for every $t \in F$ we eventually have $\pi_k(t) = t$. Note also that

$$d(\pi_k(t), \pi_{k-1}(t)) \leqslant r_k + r_{k-1} = 3r_k.$$

For every $t \in F$ we write

$$Y_t - Y_{t_0} = \sum_{k=1}^{\infty} \big(Y_{\pi_k(t)} - Y_{\pi_{k-1}(t)}\big)$$

and, using the fact that $\mathbb{E}(Y_{t_0}) = 0$,

$$\mathbb{E}\max_{t\in F} Y_t = \mathbb{E}\max_{t\in F}(Y_t - Y_{t_0}) = \int_0^\infty P\left(\max_{t\in F}(Y_t - Y_{t_0}) \geqslant u\right) du.$$

We fix $\alpha_k > 0$ (which will be suitably chosen) with $S := \sum \alpha_k < \infty$ and set $B_k = \{(w,z) \in A_k \times A_{k-1} : d(w,z) \leqslant 3r_k\}$. Using the sub-Gaussian assumption we write

$$\begin{aligned} P\left(\max_{t\in F}(Y_t - Y_{t_0}) \geqslant uS\right) &\leqslant P\left(\sum_{k=1}^\infty \max_{t\in F}(Y_{\pi_k(t)} - Y_{\pi_{k-1}(t)}) \geqslant \sum_{k=1}^\infty u\alpha_k\right) \\ &\leqslant \sum_{k=1}^\infty P\left(\max_{(w,z)\in B_k}(Y_w - Y_z) \geqslant u\alpha_k\right) \\ &\leqslant \sum_{k=1}^\infty \exp(-u^2\alpha_k^2/9r_k^2)|A_k|\cdot|A_{k-1}|. \end{aligned}$$

We now choose $\alpha_k = 3r_k\sqrt{\log(2^k|A_k|^2)}$. For every $u \geqslant 1$ and every k we have

$$\exp(-u^2\alpha_k^2/9r_k^2)|A_k|\cdot|A_{k-1}| \leqslant |A_k|^2\left(2^k|A_k|^2\right)^{-u^2} \leqslant 2^{-u^2k},$$

and hence

$$P\left(\max_{t\in F}(Y_t - Y_{t_0}) \geqslant uS\right) \leqslant \sum_{k=1}^\infty 2^{-u^2k} \leqslant c2^{-u^2}.$$

This shows that

$$\begin{aligned} \mathbb{E}\max_{t\in F} Y_t &= S\int_0^\infty P\left(\max_{t\in F}(Y_t - Y_{t_0}) \geqslant uS\right) du \\ &\leqslant S + S\int_1^\infty c2^{-u^2} du \leqslant c'S. \end{aligned}$$

Going back to the definition of S we see that

$$\begin{aligned} S &= \sum_{k=1}^\infty \frac{3R}{2^k}\left(k\sqrt{\log 2} + \sqrt{\log N(F,d,R/2^k)}\right) \\ &\simeq \sum_{k=1}^\infty \frac{R}{2^k}\sqrt{\log N(F,d,R/2^k)} \\ &\simeq \int_0^\infty \sqrt{\log N(F,d,\varepsilon)}d\varepsilon. \end{aligned}$$

Since $N(F,d,\varepsilon) \leqslant N(T,d,\varepsilon)$, the proof is complete. □

1.9.4. Majorizing measures

Dudley's bound is not always sharp, as one can see by the following example: Consider an infinite sequence $\{g_n\}$ of independent standard Gaussian random variables, fix $a = (a_n) \in \ell_2$ and define the ellipsoid

$$\mathcal{E} = \left\{t = (t_n) \in \ell_2 : \sum_{n=1}^\infty t_n^2/a_n^2 \leqslant 1\right\}.$$

If we set $Z_t = \sum_n t_n g_n$, then $\mathcal{Z} = (Z_t)_{t\in\mathcal{E}}$ is a Gaussian process and

$$\mathbb{E}\sup_{t\in\mathcal{E}} Z_t \simeq \left(\sum_{n=1}^{\infty} a_n^2\right)^{1/2} < \infty.$$

On the other hand, one can choose $a \in \ell_2$ so that "Dudley's integral" will diverge.

Starting with the argument that we presented in the previous subsection one may check that the following (genuinely better) version of Dudley's inequality can be obtained along the same lines (see [**491**, Section 1.2]).

THEOREM 1.9.6. *Let (T, d) be a metric space and let $\{Y_t\}_{t\in T}$ be a sub-Gaussian process with respect to d. Assume that $\{T_n\}_{n\geqslant 0}$ is a sequence of subsets of T such that $|T_0| = 1$ and $|T_n| \leqslant 2^{2^n}$ for all $n \geqslant 1$. Then,*

$$\mathbb{E}\sup_{t\in T} Y_t \leqslant C \sup_{t\in T} \sum_{n=0}^{\infty} 2^{n/2} d(t, T_n),$$

where $C > 0$ is an absolute constant.

Talagrand's majorizing measure theorem, which we describe below, shows that the upper bound of Theorem 1.9.6 provides the correct estimate for $\mathbb{E}\sup_{t\in T} Z_t$ for every Gaussian process $\mathcal{Z} = (Z_t)_{t\in T}$.

Let (T, d) be a metric space and let $\{\mathcal{A}_n\}_{n=0}^{\infty}$ be an increasing sequence of partitions of T; the term *increasing* means that $\mathcal{A}_{n+1}$ is a refinement of $\mathcal{A}_n$ for every $n \geqslant 0$. We say that $\{\mathcal{A}_n\}$ is *admissible* if $\mathcal{A}_0 = \{T\}$ and $|\mathcal{A}_n| \leqslant 2^{2^n}$ for every $n \geqslant 1$. Given a partition P of T, for any $t \in T$ we denote by $P(t)$ the set from P which contains t.

Now, let $\{Y_t\}_{t\in T}$ be a sub-Gaussian process with respect to d. Let $\{\mathcal{A}_n\}$ be an admissible sequence of partitions of T. Given n, consider a subset T_n of T which contains exactly one point from each set in the partition $\mathcal{A}_n$. Then, Theorem 1.9.6 shows that

$$\mathbb{E}\sup_{t\in T} Y_t \leqslant C \sup_{t\in T} \sum_{n=0}^{\infty} 2^{n/2} d(t, T_n) \leqslant C \sup_{t\in T} \sum_{n=0}^{\infty} 2^{n/2} \mathrm{diam}(\mathcal{A}_n(t)).$$

The next theorem shows that in the case of Gaussian processes this bound is optimal.

THEOREM 1.9.7 (Talagrand). *There exists an absolute constant $c_0 > 0$ with the following property: if $\{Z_t\}_{t\in T}$ is a Gaussian process, then there exists an admissible sequence $\{\mathcal{A}_n\}$ of partitions of T such that*

$$\mathbb{E}\sup_{t\in T} Z_t \geqslant c_0 \sup_{t\in T} \sum_{n=0}^{\infty} 2^{n/2} \mathrm{diam}(\mathcal{A}_n(t)).$$

In other words,

$$\mathbb{E}\sup_{t\in T} Z_t \simeq \inf_{\{\mathcal{A}_n\}} \sup_{t\in T} \sum_{n=0}^{\infty} 2^{n/2} \mathrm{diam}(\mathcal{A}_n(t)).$$

A direct consequence of the above is the following comparison theorem.

THEOREM 1.9.8 (Talagrand). *Let $\mathcal{Z} = (Z_t)_{t\in T}$ be a Gaussian process and let d be the induced metric. If the process $\mathcal{Y} = (Y_t)_{t\in T}$ is sub-Gaussian with respect to d, then*

$$\mathbb{E}\sup_{t\in T} Y_t \leqslant C \cdot \mathbb{E}\sup_{t\in T} Z_t,$$

where $C > 0$ is an absolute constant.

1.10. Dvoretzky type theorems

Dvoretzky theorem states that every high-dimensional normed space has a subspace of "large dimension" which is C-isomorphic to a Euclidean space, where C is an absolute constant. We use the terminology "Dvoretzky-type theorems" to refer to a wide family of results concerning the existence of large nice substructures inside normed spaces of sufficiently high dimension. One of the most crucial and important aspects of the theory is to find concrete estimates for the different parameters we are interested in, such as the dimension of the substructures in relation to the dimension of the whole space; there are many theorems which provide such estimates (or in some cases even asymptotic formulas) depending on the various parameters which appear in this type of results. Another important topic is to find the optimal dependences.

The starting point for the original Dvoretzky theorem has been a lemma of Dvoretzky and Rogers which shows that, given a symmetric convex body K whose maximal volume ellipsoid is B_2^n, we can find a k-dimensional subspace E of $\mathbb{R}^n$, with $k \simeq \sqrt{n}$, such that $B_2^n \cap E \subseteq K \cap E \subseteq 2Q_k$, where we write Q_k for the unit cube in E with respect to a suitable coordinate system. Therefore, we have

THEOREM 1.10.1 (Dvoretzky-Rogers). *Assume that B_2^n is the maximal volume ellipsoid of the symmetric convex body K. There exist $k \simeq \sqrt{n}$ and orthonormal vectors $z_1, \ldots, z_k$ in $\mathbb{R}^n$ such that for all $a_1, \ldots, a_k \in \mathbb{R}$,*

$$\frac{1}{2}\max_{i\leqslant k} |a_i| \leqslant \left\| \sum_{i=1}^k a_i z_i \right\| \leqslant \left(\sum_{i=1}^k a_i^2 \right)^{1/2}.$$

Grothendieck asked whether it is possible to replace Q_k by $B_2^n \cap E$ in the above statement, and still have k increase to infinity with n. Dvoretzky theorem provides an affirmative answer to this question. The best known version is the following

THEOREM 1.10.2 (Dvoretzky-Milman). *Let X be an n-dimensional normed space and $\varepsilon \in (0,1)$. There exist an integer $k \geqslant c\varepsilon^2 \log n$ and a k-dimensional subspace F of X which satisfies $d(F, \ell_2^k) \leqslant 1 + \varepsilon$.*

In geometric terms this can be stated as follows: if K is a symmetric convex body in $\mathbb{R}^n$, then for every $\varepsilon \in (0,1)$ we can find $k \geqslant c\varepsilon^2 \log n$, a subspace $F \in G_{n,k}$ and an ellipsoid $\mathcal{E}$ in F so that

$$\mathcal{E} \subseteq K \cap F \subseteq (1+\varepsilon)\mathcal{E}.$$

The example of ℓ_∞^n shows that the logarithmic dependence of k on n is the best possible if we fix small values of ε. The exact way that all three parameters, n, ε and k, are related to each other has not been settled yet. It seems reasonable that ℓ_∞^n represents the worst case, and if this proved to be true it would imply that, for fixed k and ε, every n-dimensional normed space has a k-dimensional subspace which is $(1+\varepsilon)$-isomorphic to ℓ_2^k, provided that $n \geqslant c(k)\varepsilon^{-\frac{k-1}{2}}$. The problem is of great interest even for small values of k.

1.10.1. Proof of Dvoretzky theorem

Vitali Milman's proof of Theorem 1.10.2 utilizes the concentration of measure on S^{n-1}. We start with an n-dimensional normed space X, and we assume without loss of generality that B_2^n is the ellipsoid of maximal volume inscribed in the unit ball K of X. Observe that the function $r : S^{n-1} \to \mathbb{R}$ defined by $r(x) = \|x\|$, where $\|\cdot\|$ is the norm of X, is Lipschitz continuous with constant 1. If L_r is the Lévy median of r, Theorems 1.7.7 and 1.7.2 show that for every $t \in (0,1)$

$$\sigma\left(x \in S^{n-1} : |r(x) - L_r| \geqslant tL_r\right) \leqslant \exp\bigl(-ct^2L_r^2 n\bigr),$$

where $c > 0$ is an absolute constant. The idea is that, since the function $r(x) = \|x\|$ is almost constant and equal to L_r on a subset of the sphere whose measure is practically equal to 1, one can extract a subsphere on the whole of which r will be almost equal to L_r; this is done by a discretization argument via nets of spheres.

THEOREM 1.10.3 (Milman). *Let $X = (\mathbb{R}^n, \|\cdot\|)$ and assume that $\|x\| \leqslant \|x\|_2$ for all $x \in \mathbb{R}^n$. For every $\varepsilon \in (0,1)$ we can find $k \geqslant c\varepsilon^2 L_r^2 n$ and a k-dimensional subspace F of $\mathbb{R}^n$ such that*

$$(1+\varepsilon)^{-1/2} L_r \|x\|_2 \leqslant \|x\| \leqslant L_r (1+\varepsilon)^{1/2} \|x\|_2$$

for every $x \in F$. □

If $Y = (F, \|\cdot\|)$, it is clear that $d(Y, \ell_2^k) \leqslant 1+\varepsilon$, and what remains to do is to give a lower bound for L_r. To this end, it is easier to work with the expectation

$$M = M(X) = \int_{S^{n-1}} \|x\|\, d\sigma(x),$$

of the norm on the sphere, and then a rather simple computation, based on measure concentration, shows that $L_r \simeq M$.

Finally, we make full use of the fact that B_2^n is the ellipsoid of maximal volume inscribed in K. By the Dvoretzky-Rogers lemma (Corollary 1.5.11), we can find an orthonormal basis $\{v_1, \ldots, v_n\}$ with $\|v_i\| \geqslant 1/2$ for all $i \leqslant n/2$. One may check that

$$\begin{aligned} M &= \int_{S^{n-1}} \Bigl\| \sum_{i=1}^n a_i v_i \Bigr\|\, d\sigma(a) = \int_{S^{n-1}} \int_{E_2^n} \Bigl\| \sum_{i=1}^n \epsilon_i a_i v_i \Bigr\|\, d\epsilon d\sigma(a) \\ &\geqslant \int_{S^{n-1}} \max_{1\leqslant i\leqslant n} \|a_i v_i\|\, d\sigma(a) \geqslant \frac{1}{2} \int_{S^{n-1}} \max_{1\leqslant i\leqslant n/2} |a_i|\, d\sigma(a) \geqslant c\sqrt{\log n/n}, \end{aligned}$$

where $c > 0$ is an absolute constant. Going back to Theorem 1.10.3 we conclude the proof of Theorem 1.10.2. □

1.10.2. The critical dimension

Let $X = (\mathbb{R}^n, \|\cdot\|)$ be a normed space and let a, b be the smallest positive constants so that $a^{-1}\|x\|_2 \leqslant \|x\| \leqslant b\|x\|_2$ for all $x \in \mathbb{R}^n$. The proof of Dvoretzky theorem in the previous section shows that a subspace E of X with $\dim E = \lfloor c\varepsilon^2 n (M/b)^2 \rfloor$ is $(1+\varepsilon)$-Euclidean with high probability. This inspires the following definition:

DEFINITION 1.10.4. *Let X be an n-dimensional normed space. We set $k(X)$ to be the largest positive integer $k \leqslant n$ for which*

$$\nu_{n,k}\left(\left\{E_k \in G_{n,k} : \frac{1}{2}M\|x\|_2 \leqslant \|x\| \leqslant 2M\|x\|_2, x \in E_k\right\}\right) \geqslant 1 - \frac{k}{n+k}.$$

In other words, $k(X)$ is the largest possible dimension $k \leqslant n$ such that the distance of most of the k-dimensional subspaces of X from the Euclidean space is at most 4. Note that the presence of M in the definition corresponds to the right normalization, since the average of $M(E)$ over $G_{n,k}$ is equal to M for all $1 \leqslant k \leqslant n$.

Starting from the proof of Dvoretzky theorem and using the equivalence $L_r \simeq M$ one can check that $k(X) \geqslant cn(M/b)^2$. Milman and Schechtman observed that an inverse inequality also holds true.

THEOREM 1.10.5 (Milman-Schechtman). *For every n-dimensional normed space X one has*

$$k(X) \leqslant 8n(M/b)^2.$$

Proof. We fix orthogonal subspaces $E_1, \ldots, E_t$ of dimension $k(X)$ such that $E = \sum_{i=1}^t E_i$ has dimension $n \geqslant m > n - k(X)$, and we write $\mathbb{R}^n = \sum_{i=1}^t E_i + E^\perp$. We may also expand $E^\perp$ to a $k(X)$-dimensional subspace E_{t+1} of $\mathbb{R}^n$ in such a way that $\dim(E_t \cap E_{t+1}) = k(X) + m - n$. By the definition of $k(X)$, most orthogonal images of each E_i are 4-Euclidean, and we can find $U \in O(n)$ such that

$$\frac{1}{2}M\|x\|_2 \leqslant \|x\| \leqslant 2M\|x\|_2, \qquad x \in U(E_i)$$

for all $i = 1, \ldots, t, t+1$. Every $x \in \mathbb{R}^n$ can be written in the form $x = \sum_{i=1}^t x_i + y$, where $x_i \in U(E_i)$ and $y \in U(E^\perp) \subset U(E_{t+1})$. Since the x_i are orthogonal, we get

$$\|x\| \leqslant 2M\sum_{i=1}^t \|x_i\|_2 + 2M\|y\|_2 \leqslant 2M\sqrt{t+1}\|x\|_2.$$

This shows that $b \leqslant 2M\sqrt{t+1}$, and since $t = \lfloor n/k(X)\rfloor$ we conclude that $k(X) \leqslant 8n(M/b)^2$. □

Combining all the above, we arrive at an *asymptotic formula* for $k(X)$:

$$k(X) \simeq n(M/b)^2 \tag{1.10.1}$$

for every n-dimensional normed space X. In the case of the classical spaces ℓ_p^n, $1 \leqslant p \leqslant \infty$, we can use this formula to compute the order of $k_p := k(\ell_p^n)$ as a function of p and n.

THEOREM 1.10.6. *If $1 \leqslant p \leqslant 2$ then $k_p \simeq n$, whereas if $2 < q < \infty$ then $c_1 n^{2/q} \leqslant k_q \leqslant c_2(q)n^{2/q}$, with $c_2(q) \simeq q$.*

When $p = \infty$, we have $k_p \geqslant c\log n$ from Dvoretzky theorem. This estimate turns out to be sharp.

THEOREM 1.10.7. *$k_\infty \simeq \log n$.*

The idea of the proof of Theorem 1.10.7 gives a more general result.

PROPOSITION 1.10.8. *If P is a polytope with m facets in $\mathbb{R}^k$, and if $B_2^k \subseteq P \subseteq aB_2^k$, then*

$$m \geqslant \exp\Big(\frac{k}{2a^2}\Big).$$

Proof. We can write

$$P = \{x \in \mathbb{R}^k : \langle x, v_j\rangle \leqslant 1\ ,\ j \leqslant m\}$$

for some vectors $v_j \in \mathbb{R}^k$. Since $B_2^k \subseteq P$, we must have $\|v_j\|_2 \leqslant 1$ for every $j = 1, \dots, m$. From our other assumption, that $P \subseteq aB_2^k$, it follows that for every $\theta \in S^{n-1}$ there exists $j \leqslant m$ for which $\langle \theta, v_j\rangle \geqslant 1/a$.

We set $u_j = v_j/\|v_j\|_2$, $j = 1, \dots, m$. Since $\|v_j\|_2 \leqslant 1$,

$$\{\theta \in S^{k-1} : \langle \theta, v_j\rangle \geqslant 1/a\} \subseteq \{\theta \in S^{k-1} : \langle \theta, u_j\rangle \geqslant 1/a\},$$

and hence

$$S^{k-1} \subset \bigcup_{j=1}^{m} \left\{\theta \in S^{k-1} : \langle \theta, u_j\rangle \geqslant 1/a\right\}. \tag{1.10.2}$$

For each j, $\{\theta \in S^{k-1} : \langle \theta, u_j\rangle \geqslant 1/a\}$ is a cap in S^{k-1}, centered at u_j and with angular radius $2\arcsin\frac{1}{2a}$. Using the next lemma we can estimate its measure.

LEMMA 1.10.9. *For every $u \in S^{k-1}$ and for every $\varepsilon \in (0,1)$ we set $C(u,\varepsilon) := \{\theta \in S^{k-1} : \langle u, \theta\rangle \geqslant \varepsilon\}$. Then,*

$$\sigma(C(u,\varepsilon)) \leqslant \exp(-\varepsilon^2 k/2).$$

Proof. The measure σ of $C(u,\varepsilon)$ is equal to the percentage of B_2^k which is occupied by the *spherical cone* which corresponds to $C(u,\varepsilon)$. Observe that this cone is contained in a Euclidean ball of radius $(1-\varepsilon^2)^{1/2}$, and hence

$$\sigma(C(u,\varepsilon)) \leqslant (1-\varepsilon^2)^{k/2} \leqslant \exp(-\varepsilon^2 k/2),$$

as claimed. □

Now, applying (1.10.2) and Lemma 1.10.9 with $\varepsilon = 1/a$, we get

$$1 = \sigma(S^{k-1}) \leqslant m\sigma(C(u,\varepsilon)) \leqslant m\exp(-k/2a^2),$$

and we complete the proof of Proposition 1.10.8. □

Proof of Theorem 1.10.7. We assume that for some $k \in \mathbb{N}$ there exists a k-dimensional subspace of ℓ_∞^n such that $d(F, \ell_2^k) \leqslant 4$. Then, there exists an ellipsoid $\mathcal{E}$ in F such that $\mathcal{E} \subset Q_n \cap F \subset 4\mathcal{E}$. The cube Q_n has $2n$ facets, which implies that $P := Q_n \cap F$ has $m \leqslant 2n$ facets. Thus, applying a linear transformation, we find a polytope $P_1 = T(P) \subset \mathbb{R}^k$ with m facets, which satisfies

$$B_2^k \subset P_1 \subset 4B_2^k.$$

Proposition 1.10.8 shows that $2n \geqslant m \geqslant \exp(k/32)$, which gives

$$k \leqslant 32\log(2n).$$

It follows that $k_\infty \leqslant 32\log(2n)$, and hence $k_\infty \simeq \log n$. □

Recall that, if $X = (\mathbb{R}^n, \|\cdot\|)$ is an n-dimensional normed space, then the dual norm is defined by

$$\|x\|_* = \sup\{|\langle x, y\rangle| : \|y\| \leqslant 1\}.$$

If $a^{-1}\|x\|_2 \leqslant \|x\| \leqslant b\|x\|_2$ for all x, it is clear that $b^{-1}\|x\|_2 \leqslant \|x\|_* \leqslant a^{-1}\|x\|_2$ for all x. Thus, if we define

$$k_*(X) = k(X^*) \quad \text{and} \quad M^*(X) = M(X^*),$$

we have

$$k_* \simeq n(M^*/a)^2.$$

On the other hand, a trivial application of the Cauchy-Schwarz inequality shows that

$$MM^* \geqslant \left(\int_{S^{n-1}} \|x\|_*^{1/2}\|x\|^{1/2}d\sigma(x)\right)^2 \geqslant \left(\int_{S^{n-1}} |\langle x,x\rangle|^{1/2}d\sigma(x)\right)^2 = 1.$$

This gives

$$kk_* \geqslant cn^2\frac{(MM^*)^2}{(ab)^2} \geqslant \frac{cn^2}{(ab)^2}.$$

Using John's theorem, we can always bring the unit ball of X to a position such that $ab \leqslant \sqrt{n}$. This immediately proves the next fact.

THEOREM 1.10.10. *For every n-dimensional normed space X there exists a Euclidean structure for which one has*

$$k(X)k_*(X) \geqslant cn,$$

where $c > 0$ is an absolute constant.

1.10.3. Volume ratio and Kashin's theorem

Let K be a symmetric convex body in $\mathbb{R}^n$. Recall that the *volume ratio* of K is the quantity

$$vr(K) = \inf\left\{\left(\frac{|K|}{|\mathcal{E}|}\right)^{1/n} : \mathcal{E} \subseteq K\right\},$$

where the infimum is taken over all the ellipsoids which are contained in K. It is easy to check that the volume ratio is invariant under invertible linear transformations of $\mathbb{R}^n$.

EXAMPLE 1.10.11. Let K be a symmetric convex body in $\mathbb{R}^n$. If $\|\cdot\|$ is the norm induced by K, then

$$\int_{\mathbb{R}^n} e^{-\|x\|^p}dx = \int_{\mathbb{R}^n}\int_{\|x\|}^\infty pt^{p-1}e^{-t^p}dtdx = \int_0^\infty pt^{p-1}e^{-t^p}|\{x : \|x\| \leqslant t\}|dt$$
$$= |K|\int_0^\infty pt^{n+p-1}e^{-t^p}dt = |K|\Gamma\left(\frac{n}{p}+1\right).$$

If we choose $K = B_p^n$, $1 \leqslant p < \infty$, we see that

$$\int_{\mathbb{R}^n} e^{-\|x\|^p}dx = \left(2\int_0^\infty e^{-t^p}dt\right)^n = [2\Gamma(1/p+1)]^n.$$

Therefore,

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

Observe that if $1 \leqslant p \leqslant 2$ then the maximal volume ellipsoid of B_p^n is $n^{\frac{1}{2}-\frac{1}{p}}B_2^n$. It follows that

$$vr(B_p^n) = \frac{2\Gamma(\frac{1}{p}+1)[\Gamma(\frac{n}{2}+1)]^{\frac{1}{n}}}{n^{\frac{1}{2}-\frac{1}{p}}[\Gamma(\frac{n}{p}+1)]^{\frac{1}{n}}\sqrt{\pi}} \leqslant C,$$

where $C > 0$ is an absolute constant. We say that the unit balls of ℓ_p^n, $1 \leqslant p \leqslant 2$ have *uniformly bounded volume ratio.*

The next theorem asserts that, if a body K has bounded volume ratio, then the space X_K contains subspaces F of dimension proportional to n which have bounded Banach-Mazur distance from $\ell_2^{\dim F}$. This fact was first proved by Kashin in the case of ℓ_1^n, and later Szarek and Tomczak-Jaegermann introduced the notion of volume ratio and proved the following

THEOREM 1.10.12 (Kashin, Szarek-Tomczak). *Let K be a symmetric convex body in $\mathbb{R}^n$ such that $B_2^n \subseteq K$ and $|K| = \alpha^n |B_2^n|$ for some $\alpha > 1$. Given $1 \leqslant k \leqslant n$, a random subspace $E \in G_{n,k}$ satisfies the following:*

$$B_E \subseteq K \cap E \subseteq (c\alpha)^{\frac{n}{n-k}} B_E,$$

with probability greater than $1 - e^{-n}$, where $c > 0$ is an absolute constant.

Proof. Since $B_2^n \subseteq K$, we have $\|x\| \leqslant \|x\|_2$ for every $x \in \mathbb{R}^n$. Let $k \leqslant n$. We may write

$$\int_{G_{n,k}} \int_{S_E} \|x\|^{-n} d\sigma_E(x) d\nu_{n,k}(E) = \int_{S^{n-1}} \|x\|^{-n} d\sigma(x) = \frac{|K|}{|B_2^n|} = \alpha^n.$$

From Markov's inequality, the measure of the set of $E \in G_{n,k}$ which satisfy

$$\int_{S_E} \|x\|^{-n} d\sigma_E(x) \leqslant (e\alpha)^n$$

is greater than $1 - e^{-n}$. Let E be such a subspace. Then, applying Markov's inequality again, we see that for every $r \in (0,1)$

$$\sigma_E\{x \in S_E : \|x\| < r\} \leqslant (er\alpha)^n. \tag{1.10.3}$$

We will use the following simple lemma.

LEMMA 1.10.13. *If $x \in S^{k-1}$ and $0 < t < 1$, then $\sigma_k(B(x,t)) \geqslant (t/3)^k$.*

Proof. There exists a t-net $\mathcal{N}$ of S^{k-1} of cardinality $|\mathcal{N}| \leqslant \left(1 + \frac{2}{t}\right)^k$. Since $S^{k-1} \subseteq \bigcup_{x \in \mathcal{N}} B(x,t)$ we must have $\sigma_k(B(x,t))|\mathcal{N}| \geqslant 1$. This implies that

$$\sigma_k(B(x,t)) \geqslant \left(\frac{t}{t+2}\right)^k \geqslant \left(\frac{t}{3}\right)^k,$$

and the lemma is proved. □

We now return to the proof of Theorem 1.10.12. Let $E \in G_{n,k}$ satisfy (1.10.3). Fix $x \in S_E$. From the lemma we see that if $(er\alpha)^n < (r/6)^k$ then

$$C(x, r/2) \cap \{y \in S_E : \|y\| \geqslant r\} \neq \emptyset.$$

Then, we may find $y \in S_E$ with $\|x - y\|_2 \leqslant r/2$ and $\|y\| \geqslant r$. By the triangle inequality we get

$$\|x\| \geqslant \|y\| - \|x - y\| \geqslant r - \|x - y\|_2 \geqslant r/2.$$

Since $x \in S_E$ was arbitrary, this shows that

$$B_E \subseteq K \cap E \subseteq \frac{2}{r} B_E.$$

It remains to choose an optimal r: we want

$$e^n 6^k \alpha^n r^{n-k} < 1,$$

which gives $r_{\max} = (6e\alpha)^{-\frac{n}{n-k}}$. □

REMARK 1.10.14. Theorem 1.10.12 says, for example, that if $1 \leqslant p \leqslant 2$ and $\lambda \in (0,1)$, then ℓ_p^n has subspaces F of dimension $k = \lfloor \lambda n \rfloor + 1$ with $d(F, \ell_2^k) \leqslant C_1^{\frac{1}{1-\lambda}}$, where $C_1 > 0$ is an absolute constant. Of course the estimate is bad when $\lambda \to 1$, but the distance remains uniformly bounded as long as, say, $\lambda \leqslant 1/2$.

The next result is a "global formulation" of the volume ratio theorem.

THEOREM 1.10.15. *Let K be a convex body in $\mathbb{R}^n$ such that $B_2^n \subseteq K$ and $|K| = \alpha^n |B_2^n|$ for some $\alpha > 1$. There exists $U \in O(n)$ with the property*

$$B_2^n \subset K \cap U(K) \subset c\alpha^2 B_2^n,$$

where $c > 0$ is an absolute constant.

Proof. Note that

$$\|x\|_{K\cap U(K)} = \max\{\|Ux\|, \|x\|\} \geqslant \frac{\|Ux\| + \|x\|}{2}$$

for all $U \in O(n)$ and $x \in \mathbb{R}^n$. Since $B_2^n \subset K \cap U(K)$ for every $U \in O(n)$, the theorem will follow if we find $U \in O(n)$ such that

$$N_U(\theta) := \frac{\|U\theta\| + \|\theta\|}{2} \geqslant \frac{1}{c\alpha^2}$$

for all $\theta \in S^{n-1}$. We have

$$\begin{aligned}
\int_{O(n)} \int_{S^{n-1}} \frac{1}{\|U\theta\|^n \|\theta\|^n} d\sigma(\theta) d\nu(U) &= \int_{S^{n-1}} \left(\int_{O(n)} \frac{1}{\|U\theta\|^n} d\nu(U) \right) \frac{1}{\|\theta\|^n} d\sigma(\theta) \\
&= \int_{S^{n-1}} \left(\int_{S^{n-1}} \frac{1}{\|\phi\|^n} d\sigma(\phi) \right) \frac{1}{\|\theta\|^n} d\sigma(\theta) \\
&= \left(\int_{S^{n-1}} \frac{1}{\|\theta\|^n} d\sigma(\theta) \right)^2 \\
&= \alpha^{2n}.
\end{aligned}$$

Therefore, we can find $U \in O(n)$ which satisfies

$$\int_{S^{n-1}} \left(\frac{2}{\|U\theta\| + \|\theta\|} \right)^{2n} d\sigma(\theta) \leqslant \int_{S^{n-1}} \frac{1}{\|U\theta\|^n \|\theta\|^n} d\sigma(\theta) \leqslant \alpha^{2n}.$$

Let $\theta \in S^{n-1}$ and set $N_U(\theta) = t$. If $\phi \in S^{n-1}$ and $\|\theta - \phi\|_2 \leqslant t$, then the fact that N_U is a norm with Lipschitz constant 1 gives

$$N_U(\phi) \leqslant N_U(\theta) + N_U(\phi - \theta) \leqslant t + \|\phi - \theta\|_2 \leqslant 2t.$$

On the other hand, $\sigma(B(\theta, t)) \geqslant (t/3)^n$, and hence

$$\left(\frac{t}{3}\right)^n \frac{1}{(2t)^{2n}} \leqslant \sigma(B(\theta,t)) \frac{1}{(2t)^{2n}} \leqslant \int_{S^{n-1}} \left(\frac{1}{N_U(\phi)} \right)^{2n} d\sigma(\phi) \leqslant \alpha^{2n}.$$

It is now clear that $t \geqslant 1/(c\alpha^2)$ for some absolute constant $c > 0$. This completes the proof. □

REMARK 1.10.16. It is worth observing that the proofs of the two theorems proceed along the same lines. This is an instance of a much more general principle: local statements (like Theorem 1.10.12), which describe the geometry of proportional sections of a convex body K, frequently have analogous global statements (like Theorem 1.10.15), which relate K to its orthogonal images.

In the case of ℓ_1^n we get a very interesting application of Theorem 1.10.15.

THEOREM 1.10.17. *There exist vectors* $y_1,\dots,y_{2n}\in S^{n-1}$ *such that*

$$c\sqrt{n}\|x\|_2 \leqslant \sum_{j=1}^{2n}|\langle x,y_j\rangle| \leqslant 2\sqrt{n}\|x\|_2$$

for every $x\in\mathbb{R}^n$*, where* $c>0$ *is an absolute constant.*

Proof. The maximal volume ellipsoid of B_1^n is $n^{-1/2}B_2^n$, and its volume ratio is bounded by an absolute constant $C>0$. From the proof of Theorem 1.10.15, we can find $U\in O(n)$ with the property

$$2\sqrt{n}\geqslant \|\theta\|_1+\|U\theta\|_1\geqslant \frac{\sqrt{n}}{C_1^2}$$

for every $\theta\in S^{n-1}$, where $C_1>0$ is an absolute constant. We set $y_i=e_i$ and $y_{n+i}=U^*(e_i)$, $i=1,\dots,n$. Then,

$$2\sqrt{n}\geqslant \sum_{j=1}^{2n}|\langle \theta,y_j\rangle|\geqslant \frac{\sqrt{n}}{C_1^2}$$

for every $\theta\in S^{n-1}$, and this completes the proof. □

1.11. The ℓ-position and Pisier's inequality

1.11.1. ℓ-position

Let X be an n-dimensional normed space, and let α be a norm on the space $L(\ell_2^n,X)$ of linear operators $u:\ell_2^n\to X$. The *trace dual norm* is defined on $L(X^*,\ell_2^n)$ by

$$\alpha^*(v)=\sup\{\mathrm{tr}(vu):\alpha(u)\leqslant 1\}.$$

The next lemma of Lewis applies to any pair of trace dual norms.

THEOREM 1.11.1. *For any norm* α *on* $L(\ell_2^n,X)$*, there exists* $u:\ell_2^n\to X$ *such that* $\alpha(u)=1$ *and* $\alpha^*(u^{-1})=n$*.*

The ℓ*-norm* on $L(\ell_2^n,X)$ was defined by Figiel and Tomczak-Jaegermann: Let $\{g_1,\dots,g_n\}$ be a sequence of independent standard Gaussian random variables on some probability space, and let $\{e_1,\dots,e_n\}$ be the standard orthonormal basis of $\mathbb{R}^n$. If $u:\ell_2^n\to X$, the ℓ-norm of u is defined by

$$\ell(u)=\left(\mathbb{E}\,\Big\|\sum_{i=1}^n g_iu(e_i)\Big\|^2\right)^{1/2}.$$

A standard computation gives

$$\ell(u)\simeq\sqrt{n}w(u^*(K^\circ)), \tag{1.11.1}$$

where K is the unit ball of X. This formula connects the ℓ-norm to the mean width. It is more instructive to replace the Gaussians by the Rademacher functions $r_i:E_2^n\to\{-1,1\}$ defined by $r_i(\epsilon)=\epsilon_i$, where $E_2^n=\{-1,1\}^n$ is viewed as a probability space with the uniform measure. An inequality of Maurey and Pisier shows that

$$\ell(u) \simeq \left(\int_{E_2^n} \big\| \sum_{i=1}^n r_i(\epsilon) u(e_i) \big\|^2 d\epsilon \right)^{1/2}$$

up to a $\sqrt{\log n}$-term.

Consider the Walsh functions $w_A(\varepsilon) = \prod_{i\in A} r_i(\varepsilon)$, where $A \subseteq \{1, \dots, n\}$. It is not hard to see that every function $f : E_2^n \to X$ is uniquely represented in the form

$$f(\epsilon) = \sum_A w_A(\epsilon) x_A,$$

for some $x_A \in X$. The space of all functions $f : E_2^n \to X$ becomes a Banach space with the norm

$$\|f\|_{L_2(X)} = \left(\int_{E_2^n} \|f(\epsilon)\|^2 d\epsilon \right)^{1/2}$$

The *Rademacher projection* $R_n : L_2(X) \to L_2(X)$ is the operator sending $f = \sum w_A x_A$ to the function $R_n f := \sum_{i=1}^n r_i x_{\{i\}}$. Write $\mathrm{Rad}(X)$ for the operator norm of R_n. Figiel and Tomczak-Jaegermann proved the following.

THEOREM 1.11.2 (Figiel-Tomczak). *Let X be an n-dimensional normed space. There exists $u : \ell_2^n \to X$ such that*

$$\ell(u)\ell((u^{-1})^*) \leqslant n\mathrm{Rad}(X).$$

Let us briefly sketch the proof. From Theorem 1.11.1 we can find an isomorphism $u : \ell_2^n \to X$ such that $\ell(u)\ell^*(u^{-1}) = n$. On the other hand,

$$\ell\left((u^{-1})^*\right) = \left(\int_{E_2^n} \big\| \sum_{i=1}^n r_i(\epsilon)(u^{-1})^*(e_i) \big\|_*^2 d\epsilon \right)^{1/2}.$$

There exists a function $\phi : E_2^n \to X$, which can be represented in the form $\phi = \sum_A w_A x_A$ and has norm $\|\phi\|_{L_2(X)} = 1$, such that

$$\ell((u^{-1})^*) = \big\langle \sum_{i=1}^n r_i (u^{-1})^*(e_i), \phi \big\rangle = \sum_{i=1}^n \langle (u^{-1})^*(e_i), x_{\{i\}} \rangle.$$

If we define $v : \ell_2^n \to X$ by $v(e_i) = x_{\{i\}}$, we easily check that

$$\ell((u^{-1})^*) = \mathrm{tr}(u^{-1}v) \leqslant \ell^*(u^{-1})\ell(v).$$

Finally, observing that

$$\ell(v) = \|R_n(\phi)\|_{L_2(X)} \leqslant \mathrm{Rad}(X)\|\phi\|_{L_2(X)} = \mathrm{Rad}(X),$$

we get

$$\ell(u)\ell((u^{-1})^*) \leqslant \ell(u)\ell^*(u^{-1})\mathrm{Rad}(X) = n\mathrm{Rad}(X).$$

1.11.2. Pisier's inequality and the MM^*-estimate

Pisier gave a sharp estimate for $\mathrm{Rad}(X)$ in terms of the Banach-Mazur distance $d(X,\ell_2^n)$.

THEOREM 1.11.3 (Pisier). *Let X be an n-dimensional normed space. Then,*

$$\mathrm{Rad}(X) \leqslant c\log[d(X,\ell_2^n)+1],$$

where $c>0$ is an absolute constant.

Combined with the results of Lewis, Figiel and Tomczak-Jaegermann, Theorem 1.11.3 leads to the following statement (where we also use relation (1.11.1)).

THEOREM 1.11.4 (MM^*-estimate). *Let K be a symmetric convex body in $\mathbb{R}^n$. There exists a position $\tilde{K}$ of K for which*

$$w(\tilde{K})w(\tilde{K}^\circ) \leqslant c\log[d(X_K,\ell_2^n)+1],$$

where $c>0$ is an absolute constant.

Computing the volume of $\tilde{K}$ in polar coordinates and using Hölder's inequality, we check that $w(\tilde{K}^\circ)^{-1} \leqslant c_2\sqrt{n}|\tilde{K}|^{1/n}$. It follows that

$$w(\tilde{K}) \leqslant c\sqrt{n}\log n|\tilde{K}|^{1/n}.$$

Normalizing the volume we obtain the following *reverse Urysohn inequality.*

THEOREM 1.11.5. *If K is a symmetric convex body in $\mathbb{R}^n$, there exists a linear image $\tilde{K}$ of K with volume $|\tilde{K}|=1$ and mean width*

$$w(\tilde{K}) \leqslant c\sqrt{n}\log n,$$

where $c>0$ is an absolute constant.

Moreover, a simple argument based on the Rogers-Shephard inequality shows that we can remove the assumption of symmetry.

1.12. Milman's low M^*-estimate and the quotient of subspace theorem

1.12.1. Low M^*-estimate

Milman's low M^*-estimate is the first step towards a general theory of sections and projections of symmetric convex bodies in $\mathbb{R}^n$ onto subspaces with dimension proportional to n. In geometric terms, it says that for fixed $\lambda\in(0,1)$, the diameter of a random $\lfloor\lambda n\rfloor$-dimensional section of a body K is controlled by its mean width

$$w(K)=\int_{S^{n-1}} h_K(x)\,d\sigma(x)$$

up to a function depending only on λ.

THEOREM 1.12.1 (Milman). *There exists a function $f:(0,1)\to[0,\infty)$ with the following property: for every $\lambda\in(0,1)$ and every n-dimensional normed space X, a random subspace $H\in G_{n,\lfloor\lambda n\rfloor}$ satisfies*

$$\|x\| \geqslant \frac{f(\lambda)}{M^*}\|x\|_2 \quad \textit{for every } x\in H,$$

where $M^*:=w(K_X)$ is the mean width of the unit ball K_X of X.

The precise dependence on λ has been established in a series of papers. Theorem 1.12.1 was originally proved by Milman, who also gave a second proof using the isoperimetric inequality on S^{n-1}, with a bound of the form $f(\lambda) \geqslant c(1-\lambda)$. Pajor and Tomczak-Jaegermann later showed that one can take $f(\lambda) \geqslant c\sqrt{1-\lambda}$. Finally, Gordon proved that the theorem holds true with

$$f(\lambda) \geqslant \sqrt{1-\lambda}\left(1 + O\left(\frac{1}{(1-\lambda)n}\right)\right).$$

Geometrically speaking, Theorem 1.12.1 says that for a random $\lfloor \lambda n \rfloor$-dimensional section of K_X we have

$$K_X \cap E \subseteq \frac{w(K_X)}{f(\lambda)} B_2^n \cap E,$$

that is, a random section does not have the same diameter as K_X but rather has radius $w(K_X)$, which is roughly the level r at which half of the supporting hyperplanes of rB_2^n cut the body K_X.

The dual formulation of the theorem has an interesting geometric interpretation too. A random λn-dimensional projection of K_X contains a ball of radius of the order of $1/M$. More precisely, for a random $E \in G_{n,\lfloor \lambda n \rfloor}$ we have

$$P_E(K_X) \supseteq \frac{f(\lambda)}{M} B_2^n \cap E.$$

Sketch of proof. We sketch Milman's proof of the inequality which gives linear dependence on λ. Consider the set $A = \{y \in S^{n-1} : \|y\|_* \leqslant 2M^*\}$. We obviously have $\sigma(A) \geqslant \frac{1}{2}$.

CLAIM 1.12.2. For every $\lambda \in (0,1)$ there exists a subspace E of dimension $k = \lfloor \lambda n \rfloor$ such that

$$E \cap S^{n-1} \subseteq A_{(\frac{\pi}{2}-\delta)},$$

where A_ε is the ε-extension of A on the sphere and $\delta \geqslant c(1-\lambda)$.

Proof. We have $\sigma(A_{\pi/4}) \geqslant 1 - c\sqrt{n}\int_0^{\pi/4} \sin^{n-2} t\, dt$, and integration over $G_{n,k}$ shows that a random $E \in G_{n,k}$ satisfies

$$\sigma_k(A_{\pi/4} \cap E) \geqslant 1 - c\sqrt{n}\int_0^{\pi/4} \sin^{n-2} t dt.$$

On the other hand, for every $x \in S^{n-1} \cap E$ we have

$$\sigma_k\Big(B\Big(x, \frac{\pi}{4} - \delta\Big)\Big) \simeq \sqrt{k}\int_0^{\frac{\pi}{4}-\delta} \sin^{k-2} t dt.$$

This implies that, if

$$\sqrt{\lambda}\int_0^{\frac{\pi}{4}-\delta} \sin^{k-2} t dt \simeq \int_0^{\frac{\pi}{4}} \sin^{n-2} t dt, \tag{1.12.1}$$

then $A_{\pi/4} \cap B(x, \frac{\pi}{4} - \delta) \neq \emptyset$, and hence $x \in A_{\frac{\pi}{2}-\delta}$. Analyzing condition (1.12.1), we see that we can choose $\delta \geqslant c(1-\lambda)$. □

We complete the proof of Theorem 1.12.1 as follows. Let $x \in S^{n-1} \cap E$. There exists $y \in A$ such that

$$\sin\delta \leqslant |\langle x, y\rangle| \leqslant \|y\|_*\|x\| \leqslant 2M^*\|x\|,$$

and since $\sin\delta \geqslant \frac{2}{\pi}\delta \geqslant c'(1-\lambda)$, the theorem follows. □

1.12.2. Quotient of subspace theorem

Milman's *quotient of subspace theorem* states the following: by performing two operations in an n-dimensional space, first selecting a subspace and then a quotient of it, we can always arrive at a new space of dimension proportional to n which is close to being Euclidean (independently of n).

In order to interpret this in the language of convex bodies, observe that if K is the unit ball of $X = (\mathbb{R}^n, \|\cdot\|)$ and if $G \subseteq E \subseteq X$ then E/G is isometrically isomorphic to the subspace $F := E \cap G^{\perp}$ equipped with the norm induced by $P_F(K \cap E)$. We write $QS(X)$ for the class of all quotient spaces of subspaces of X; a space $Y \in QS(X)$ is of the form E/G where $G \subset E \subset X$. It is useful to note that $QS(X)$ is the same as the class $SQ(X)$ of all subspaces of quotient spaces of X. Indeed, if $F \subset E \subset \mathbb{R}^n$ one sees that $P_F(K \cap E) = (P_{F+E^{\perp}}(K)) \cap F$. This implies the following very useful property: if $Y \in QS(X)$ then every subspace or quotient space of Y also belongs to $QS(X)$ and $QS(Y) \subseteq QS(X)$. Thus, every iteration of the operation of choosing a quotient of a subspace leads to an element of $QS(X)$.

THEOREM 1.12.3 (Milman). *Let X be an n-dimensional normed space. For every $1 \leqslant k \leqslant n$ there exists $Y \in QS(X)$ with $\dim(Y) = n-k$ and*

$$d(Y, \ell_2^{n-k}) \leqslant C\frac{n}{k}\log\Big(\frac{Cn}{k}\Big), \tag{1.12.2}$$

where $C > 0$ is an absolute constant.

Geometrically, the quotient of subspace theorem asserts that for every centrally symmetric convex body K in $\mathbb{R}^n$ and any $\alpha \in (0,1)$ we can find subspaces $F \subseteq E$ with $\dim(F) \geqslant \alpha n$ and an ellipsoid $\mathcal{E}$ in F such that

$$\mathcal{E} \subset P_F(K \cap E) \subset c(1-\alpha)^{-1}|\log(c(1-\alpha))|\mathcal{E}.$$

The proof of the theorem is based on the low M^*-estimate and an iteration procedure which makes essential use of the ℓ-position.

Proof. We may assume that K_X is in ℓ-position: then, by Theorem 1.11.4 we have $M(X)M^*(X) \leqslant c\log[d(X, \ell_2^n) + 1]$.

Step 1. Let $\lambda \in (0,1)$. We show that there exist a subspace E of X with $\dim(E) \geqslant \lambda n$ and a subspace F of E^* with $\dim(F) = k \geqslant \lambda^2 n$, such that $d(F, \ell_2^k) \leqslant c(1-\lambda)^{-1}\log[d(X, \ell_2^n) + 1]$.

The proof of this fact follows from a double application of the low M^*-estimate. By Theorem 1.12.1 a random λn-dimensional subspace E of X satisfies

$$\frac{c_1\sqrt{1-\lambda}}{M^*(X)}\|x\|_2 \leqslant \|x\| \leqslant b\|x\|_2, \qquad x \in E.$$

Moreover, since this is true for a random $E \in G_{n,\lambda n}$, we may also assume that $M(E) \leqslant c_2 M(X)$. Repeating the same argument for E^*, we may find a subspace F of E^* with $\dim(F) = k \geqslant \lambda^2 n$ and

$$\frac{c_3\sqrt{1-\lambda}}{M(X)}\|x\|_2 \leqslant \frac{c_1\sqrt{1-\lambda}}{M^*(E^*)}\|x\|_2 \leqslant \|x\|_{E^*} \leqslant \frac{M^*(X)}{c_1\sqrt{1-\lambda}}\|x\|_2$$

for every $x \in F$. Since K_X is in ℓ-position, we obtain

$$d(F, \ell_2^k) \leqslant c_4(1-\lambda)^{-1}M(X)M^*(X) \leqslant c(1-\lambda)^{-1}\log[d(X, \ell_2^n) + 1].$$

Step 2. Denote by $QS(X)$ the class of all quotient spaces of a subspace of X and define a function $f:(0,1)\to[0,\infty)$ by

$$f(\alpha)=\inf\{d(F,\ell_2^k):F\in QS(X),\dim F\geqslant \alpha n\}.$$

Then, what we have really proved in Step 1 is the estimate

$$f(\lambda^2\alpha)\leqslant c(1-\lambda)^{-1}\log f(\alpha).$$

An iteration lemma allows us to deduce that

$$f(\alpha)\leqslant c(1-\alpha)^{-1}|\log(1-\alpha)|,$$

and thus conclude the proof. □

1.13. Bourgain-Milman inequality and the M-position

1.13.1. The Bourgain-Milman inequality

In Subsection 1.3.4 we saw that, for every symmetric convex body K in $\mathbb{R}^n$, the volume product $s(K)=|K||K^\circ|$ is less than or equal to the volume product $s(B_2^n)$; this is the Blaschke-Santaló inequality. In the opposite direction, a well-known conjecture of Mahler states that $s(K)\geqslant s(Q_n)=4^n/n!$ for every symmetric convex body K, where $Q_n=[-1,1]^n$ is the n-dimensional cube. This has been verified for some classes of bodies, e.g. zonoids and 1-unconditional bodies but in general Mahler's question remains open. However, the Bourgain–Milman inequality does provide an "affirmative" answer to it in the asymptotic sense: for every symmetric convex body K in $\mathbb{R}^n$, the n-th root $s(K)^{1/n}$ of the volume product is of the order of $1/n$. Note that for many applications, some of which we will see in the rest of this book, knowing the order of the n-th root of the volume product suffices.

THEOREM 1.13.1 (Bourgain-Milman). *There exists an absolute constant $c>0$ such that*

$$\left(\frac{s(K)}{s(B_2^n)}\right)^{1/n}\geqslant c$$

for every symmetric convex body K in $\mathbb{R}^n$.

The original proof of the Bourgain-Milman inequality (also called "reverse Santaló inequality") employed a dimension descending procedure which was based on Milman's quotient of subspace theorem. Later, Milman offered a second approach, introducing an "isomorphic symmetrization" technique. This symmetrization scheme, which we describe below, is closer to classical convexity, much more geometric in nature, and does preserve dimension unlike the procedure in the original proof of the reverse Santaló inequality; however, it is a symmetrization scheme which is in many ways different from the classical symmetrizations. At each step of the inductive procedure Milman's proof employs, none of the natural parameters of the body is being preserved, but the ones which are of interest remain under control. The MM^*-estimate is again crucial for the proof.

Since $s(K)$ is an affine invariant, we may start from a position of K which satisfies the inequality $M(K)M^*(K)\leqslant c\log[d(X_K,\ell_2^n)+1]$. We may also normalize so that $M(K)=1$. We define

$$\lambda_1=M^*(K)a_1\quad\text{and}\quad\lambda_1'=M(K)a_1$$

for some $a_1 > 1$ which will be suitably chosen, and consider a new body

$$K_1 := \mathrm{conv}\left(\left(K \cap \lambda_1 B_2^n\right) \cup \frac{1}{\lambda_1'} B_2^n\right).$$

Sudakov's inequality and elementary properties of covering numbers show that

$$|K_1| \geqslant |K \cap \lambda_1 B_2^n| \geqslant |K|/N(K, \lambda_1 B_2^n) \geqslant |K| \exp(-cn/a_1^2).$$

In an analogous way, using the dual Sudakov inequality one can show that

$$|K_1| \leqslant |\mathrm{conv}(K \cup (1/\lambda_1') B_2^n)| \leqslant |K| \exp(cn/a_1^2).$$

By the way K_1 is defined, one can apply the same reasoning with K_1°, and finally obtain that

$$\exp(-c/a_1^2) \leqslant \left(\frac{s(K_1)}{s(K)}\right)^{1/n} \leqslant \exp(c/a_1^2).$$

By construction, for the new body K_1 we have $d(X_{K_1}, \ell_2^n) \leqslant M(K)M^*(K)a_1^2$ and, since $s(K_1)$ is a linear invariant, we may also assume that K_1 is in (a suitably normalized) ℓ-position, so that $M(K_1)M^*(K_1) \leqslant c \log[d(X_{K_1}, \ell_2^n)+1]$ and $M(K_1) = 1$. If we set $\lambda_2 = M^*(K_1)a_2$, $\lambda_2' = M(K_1)a_2$ and define

$$K_2 = \mathrm{conv}\left((K_1 \cap \lambda_2 B_2^n) \cup \frac{1}{\lambda_2'} B_2^n\right)$$

for some $a_2 > 1$, we obtain

$$\exp(-c/a_2^2) \leqslant \left(\frac{s(K_2)}{s(K_1)}\right)^{1/n} \leqslant \exp(c/a_2^2).$$

We now iterate this procedure, choosing $a_1 = \log n$, $a_2 = \log\log n, \ldots, a_t = \log^{(t)} n$ – the t-iterated logarithm of n, and stop the procedure at the first t for which $a_t < 2$. It is easy to check that $d(X_{K_t}, \ell_2^n) \leqslant C$, therefore

$$\frac{1}{C} \leqslant \left(\frac{s(K_t)}{s(B_2^n)}\right)^{1/n} \leqslant C.$$

On the other hand,

$$c_1 \leqslant \exp\left(-c\left(\frac{1}{a_1^2} + \cdots + \frac{1}{a_t^2}\right)\right) \leqslant \left(\frac{s(K)}{s(K_t)}\right)^{1/n} \leqslant \exp\left(c\left(\frac{1}{a_1^2} + \cdots + \frac{1}{a_t^2}\right)\right),$$

which proves the theorem (observe that the series $\frac{1}{a_1^2} + \cdots + \frac{1}{a_t^2} + \cdots$ remains bounded by an absolute constant, irrespective of the final number of steps). □

1.13.2. M-position

The existence of an "M-ellipsoid" associated with any centered convex body K in $\mathbb{R}^n$ was first proved by Milman.

Theorem 1.13.2 (Milman). *There exists an absolute constant $c > 0$ such that the following holds: given a symmetric convex body K in $\mathbb{R}^n$ we can find an ellipsoid $\mathcal{E}_K$ satisfying $|K| = |\mathcal{E}_K|$ and*

$$\frac{1}{c}|\mathcal{E}_K + T|^{1/n} \leqslant |K + T|^{1/n} \leqslant c|\mathcal{E}_K + T|^{1/n}, \qquad (1.13.1)$$
$$\frac{1}{c}|\mathcal{E}_K^\circ + T|^{1/n} \leqslant |K^\circ + T|^{1/n} \leqslant c|\mathcal{E}_K^\circ + T|^{1/n}$$

for every convex body T in $\mathbb{R}^n$.

Sketch of proof. We use the same sequence of bodies as in the proof of Theorem 1.13.1 that we described above. For every s, we check that

$$\exp(-cn/a_s^2) \leqslant \frac{|K_s + T|}{|K_{s-1} + T|} \leqslant \exp(cn/a_s^2),$$

for every convex body T, and the same holds true with K_s°. After t steps, we arrive at a body K_t whose geometric distance from an ellipsoid $\mathcal{E}$ is bounded by an absolute constant c; if we normalize so that $|K_t| = |\mathcal{E}|$, then K_t and $\mathcal{E}$ satisfy estimates like the ones in (1.13.1). Our volume estimates show that $|K_t|^{1/n} \simeq |K|^{1/n}$ up to an absolute constant. If we define $\mathcal{E}_K = \rho\mathcal{E}$ where $\rho > 0$ is such that $|\mathcal{E}_K| = |K|$, then $\rho \simeq 1$ and the result follows. □

The existence of M-ellipsoids can be equivalently established by introducing the *M-position* of a convex body. To any given symmetric convex body K in $\mathbb{R}^n$ we can apply a linear transformation and find a position $\tilde{K} = u_K(K)$ of volume $|\tilde{K}| = |K|$ such that (1.13.1) is satisfied with $\mathcal{E}_K$ a multiple of B_2^n. This is the so-called M-position of K. It follows then that for every pair of convex bodies K_1 and K_2 in $\mathbb{R}^n$ and for all $t_1, t_2 > 0$,

$$|t_1\tilde{K}_1 + t_2\tilde{K}_2|^{1/n} \leqslant c' \left(t_1|\tilde{K}_1|^{1/n} + t_2|\tilde{K}_2|^{1/n} \right), \tag{1.13.2}$$

where $c' > 0$ is an absolute constant, and that (1.13.2) remains true if we replace $\tilde{K}_1$ or $\tilde{K}_2$ (or both) by their polars. This statement is Milman's *reverse Brunn-Minkowski inequality.*

To define the M-position of a convex body, we can alternatively use covering numbers. Recall that the covering number $N(A, B)$ of a body A by a second body B is the least integer N for which there exist N translates of B whose union covers A. In a similar way as above, we can show that there exists an absolute constant $\beta > 0$ such that every centered convex body K in $\mathbb{R}^n$ has a linear image $\tilde{K}$ which satisfies $|\tilde{K}| = |B_2^n|$ and

$$\max\{N(\tilde{K}, B_2^n), N(B_2^n, \tilde{K}), N(\tilde{K}^\circ, B_2^n), N(B_2^n, \tilde{K}^\circ)\} \leqslant \exp(\beta n). \tag{1.13.3}$$

We say that a convex body K which satisfies (1.13.3) is in M-position with constant β. If K_1 and K_2 are two such convex bodies, there is a standard way to show that they and their polar bodies satisfy the reverse Brunn–Minkowski inequality.

Pisier has proposed a different approach to these results, which allows one to acquire a whole family of M-ellipsoids along with more detailed information on how the corresponding covering numbers behave. The precise statement is as follows.

THEOREM 1.13.3 (Pisier). *For every $0 < \alpha < 2$ and every symmetric convex body K in $\mathbb{R}^n$, there exists a linear image $\tilde{K}$ of K such that*

$$\max\{N(\tilde{K}, tB_2^n), N(B_2^n, t\tilde{K}), N(\tilde{K}^\circ, tB_2^n), N(B_2^n, t\tilde{K}^\circ)\} \leqslant \exp\left(\frac{c(\alpha)n}{t^\alpha}\right)$$

for every $t \geqslant 1$, where $c(\alpha)$ is a constant depending only on α, with $c(\alpha) = O\big((2-\alpha)^{-\alpha/2}\big)$ as $\alpha \to 2$.

We say that a body $\tilde{K}$ which satisfies the conclusion of Theorem 1.13.3 is in *M-position of order α* (or *α-regular M-position*).

We close this section with a useful observation about the M-position.

Proposition 1.13.4. *If K is in M-position with constant β then, given any $\lambda \in (0,1)$, a random orthogonal projection $P_E(K)$ onto a $\lfloor \lambda n \rfloor$-dimensional subspace E has volume ratio bounded by a constant $C(\beta,\lambda)$.*

Proof. Note that (1.13.3) implies

$$|\mathrm{conv}(K^\circ \cup B_2^n)|^{1/n} \leqslant C|B_2^n|^{1/n},$$

where C depends on β. In other words, $W = \mathrm{conv}(K^\circ \cup B_2^n)$ has bounded volume ratio, and thus Theorem 1.10.12 shows that, for a random $E \in G_{n,[\lambda n]}$,

$$K^\circ \cap E \subseteq W \cap E \subseteq C(\beta,\lambda) B_E.$$

By duality, this means that $P_E(K)$ contains a ball rB_E of radius $r \geqslant 1/C(\beta,\lambda)$. Since

$$|P_E(K)| \leqslant N(P_E(K), B_E)|B_E| \leqslant N(K, B_2^n)|B_E| \leqslant \exp(\beta n)|B_E|,$$

we arrive at a bound of the desired order for $\big(|P_E(K)|/|rB_E|\big)^{1/\lfloor \lambda n \rfloor}$. □

1.14. Notes and references

Basic references

A thorough exposition of the theory of convex bodies can be found in the classical monographs by Schneider [**463**] and Gruber [**237**]. The books of Bonnesen and Fenchel [**94**], Burago and Zalgaller [**125**], Gardner [**191**] and Groemer [**228**] are very useful sources of additional information from different perspectives.

The asymptotic theory of finite dimensional normed spaces is presented in the books by V. Milman and Schechtman [**387**], Pisier [**430**] and Tomczak-Jaegermann [**493**]. Additional information can be found in the survey articles of Ball [**39**], Giannopoulos and V. Milman [**206**] and [**207**], Lindenstrauss [**329**], Lindenstrauss and V. Milman [**330**], V. Milman [**383**] and Pisier [**428**].

Convex bodies and mixed volumes

The tools that we will use from the theory of mixed volumes are described in Section 1.4. We refer to the books of Schneider [**463**] and Gruber [**237**] for proofs, historical information and detailed references. Minkowski's theory of mixed volumes appeared mainly in [**390**] and [**392**]; one can trace its roots in the works of Steiner [**478**] and Brunn [**122**], [**123**]. The theory of area measures was developed by Alexandrov in a series of works (see also Fenchel and Jessen [**172**]). In particular, the Alexandrov-Fenchel inequality appears in [**5**] and [**6**] (Fenchel sketched an alternative proof in [**171**]). Minkowski's existence theorem appears in [**390**] (see also [**7**]).

Brunn-Minkowski inequality

The Brunn-Minkowski inequality has its origin in the work of Brunn [**122**], [**123**] who discovered it in dimensions $n = 2, 3$. Minkowski established the n-dimensional case and characterized the case of equality in [**391**]. Blaschke gave a proof using Schwarz symmetrization in [**68**]. Lusternik [**342**] first extended the inequality to the class of compact sets. An alternative proof of Lusternik's result was obtained by Henstock and Macbeath in [**254**], and by Hadwiger and Ohmann [**244**], who also clarified Lusternik's conditions for equality. Several applications, analogues and variants of the Brunn-Minkowski inequality are discussed at length in the very informative survey paper of Gardner [**190**].

Theorem 1.2.3 is usually attributed to Prékopa and Leindler; it was proved in [**320**] and [**435**] (see also [**434**]). The survey article of Das Gupta [**155**] provides detailed information on the historical background and on related results by other groups of authors. The proof that we present in the text is more or less similar to the one given in [**430**].

The proof of the existence of Knothe's map is from [**289**].

Applications of the Brunn-Minkowski inequality

The inequality of Rogers and Shephard on the volume of the difference body was proved in [**443**]. The slightly different proof of Theorem 1.3.1 that we describe is due to Chakerian [**136**]. Other variants of the proof as well as extensions of this result can be found in [**444**], [**445**]. Theorem 1.3.2 is from [**386**].

Borell's Lema 1.3.3 appears in [**96**] and holds true in the more general setting of log-concave probability measures. It will play a very important role in this book, and we will discuss several of its applications.

The Blaschke-Santaló inequality was proved by Blaschke [**69**] in dimension $n = 3$ and by Santaló [**451**] in all dimensions. The simple proof that we describe appears in the article [**364**] of Meyer and Pajor, and in the PhD Thesis of Ball [**30**] (see also [**365**] for the not necessarily symmetric case). A well-known conjecture of Mahler states that, conversely, $s(K) \geqslant 4^n/n!$ for every centrally symmetric convex body K (with one of the minimizers being the n-dimensional cube), while in the not necessarily symmetric case one would expect that $s(K) \geqslant (n+1)^{n+1}/(n!)^2$ (with the minimum being attained at an n-dimensional regular simplex). The Bourgain-Milman inequality (which is presented in Section 1.13.1) verifies this conjecture in the asymptotic sense: for every centrally symmetric convex body K in $\mathbb{R}^n$ one has $s(K)^{1/n}$ is of the order of $1/n$.

Urysohn's inequality appears in [**494**].

Classical positions of convex bodies

John's theorem appears in [**260**]; the representation of the identity as a sum of rank one projections defined by contact points is from the same paper. Our sketch of the proof follows Ball's presentation in [**39**]. The isotropic characterization of John's position is due to Ball, see [**38**].

The Dvoretzky-Rogers lemma was proved in [**162**] and was used in the proof of the fact that every infinite dimensional Banach space X contains an unconditionally convergent series that is not absolutely convergent. It was the starting point for a question of Grothendieck that led to Dvoretzky theorem and it is used in the proof of Dvoretzky theorem.

Theorem 1.5.12 is due to Giannopoulos and Milman [**204**]. Theorem 1.5.13 is due to Petty [**422**] (see also [**215**] for a second proof and some applications to sharp inequalities for the volume of projection bodies and their polars). For a comparison of various classical positions of convex bodies see [**453**], [**351**] and [**352**].

Brascamp-Lieb inequality and its reverse form

The original proof of the Brascamp-Lieb inequality [**114**] was based on a general rearrangement inequality of Brascamp, Lieb and Luttinger[**115**] which states that if f^* is the symmetric decreasing rearrangement of a Borel measurable function f with level sets of finite measure, then

$$I(f_1, \dots, f_m) \leqslant I(f_1^*, \dots, f_m^*).$$

Then, Brascamp and Lieb used a generalized version of this inequality for functions of several variables and the fact that radial functions in high dimensions behave like Gaussian functions. The proof of the inequality and of its reverse form that we present in the text is due to Barthe [**52**] (see also [**53**] and [**51**]).

Theorem 1.6.4, the normalized form of the Brascamp-Lieb inequality, is due to Ball (see e.g. [**34**]) who applied it to obtain sharp volume estimates. Theorem 1.6.5 and the reverse isoperimetric inequality are also due to Ball [**37**]. Note that the reverse Brascamp-Lieb inequality plays a similar role if one is interested in dual statements: for example consider the outer volume ratio $\mathrm{ovr}(K) = \inf \big(|\mathcal{E}|/|K|\big)^{1/n}$, where the infimum is taken over all ellipsoids containing K. Then, $\mathrm{ovr}(K) \leqslant \mathrm{ovr}(\Delta_n)$ for every convex body K in $\mathbb{R}^n$. In the symmetric case the extremal body is the cross-polytope (the unit ball of ℓ_1^n). For a proof of these results see [**51**]).

Concentration of measure

General references on concentration, from various viewpoints, are the books of Ledoux [**316**], Ledoux and Talagrand [**318**], Gromov [**231**], and the articles of Ball [**39**], Gromov [**230**], Milman [**379**], [**381**], [**383**], Schechtman [**454**].

The solution of the isoperimetric problem on the sphere is given by Paul Lévy in [**321**] and by Schmidt [**459**]. For a proof using spherical symmetrization see [**174**]. Lemma 1.7.8, which leads to a very simple proof of the approximate isoperimetric inequality for the sphere, is due to Arias de Reyna, Ball and Villa [**20**].

The Gaussian isoperimetric inequality was discovered by Sudakov-Tsirelson [**483**] who used the isoperimetric theorem on the sphere and the observation that projections of uniform measures on N-dimensional spheres of radius $\sqrt{N}$ when projected to $\mathbb{R}^n$ approximate Gaussian measure as $N \to \infty$ (this is known as "Poincaré's lemma"; see [**480**] and [**313**]). The same result was also proved by Borell [**98**] who also obtained a Brunn-Minkowski inequality in Gauss space, and by Erhard [**168**] who developed a rearrangement of sets argument in Gauss space. Bobkov [**75**] proved an isoperimetric inequality on the discrete cube from which he also derived the Gaussian isoperimetric inequality (see also [**73**] and [**74**]). Theorem 1.7.12, which establishes the approximate isoperimetric inequality in Gauss space as a direct application of the Prékopa-Leindler inequality, is due to Maurey [**354**].

The isoperimetric problem for the discrete cube was solved by Harper in [**248**]. Theorem 1.7.14 is due to Talagrand [**488**].

Khintchine's appears in [**270**] and it was first stated in this form by Littlewood [**332**]. The best constants A_p^* and B_p^* were found by Szarek [**484**] who showed that $A_1^* = 1/\sqrt{2}$, and by Haagerup [**243**] who determined the best constants for all p. Kahane's inequality appears in [**265**] with constant proportional to p as $p \to \infty$. The optimal dependence $\sqrt{p}$ is due to Kwapien see [**304**].

Entropy estimates

Sudakov's inequality [**481**] in its original form gives a lower bound for the expectation of the supremum of a Gaussian process; this form is presented in Section 1.7. Theorem 1.8.1 is a direct application to the covering numbers of a convex body that follows from Sudakov's inequality once the geometric translation is done. The original proof of the dual statement, Theorem 1.8.2, is due to Pajor and Tomczak-Jaegermann [**404**]. The argument that allows one to pass from Sudakov's inequality to its dual and vice versa is due to Tomczak-Jaegermann [**492**] and is based on Theorem 1.8.4. We first present Talagrand's proof of the dual Sudakov inequality (see [**318**]) and following a similar route we deduce Sudakov's inequality using Tomczak's argument.

The duality of entropy theorem (Theorem 1.8.6) is due to Artstein-Avidan, Milman and Szarek [**24**]. The inequality of König and Milman was proved in [**299**].

Gaussian and sub-Gaussian processes

Theorem 1.9.4 combines the bounds of Sudakov and Dudley. Dudley proved the upper bound in [**158**] and conjectured the lower bound that was later proved by Sudakov in [**481**]. Theorem 1.9.5 was proved in [**468**] (see also [**481**], [**27**] and [**173**]).

Theorem 1.9.7 and Theorem 1.9.8 are due to Talagrand (see [**491**] and [**487**] for the original proof).

Dvoretzky theorem

Theorem 1.10.2 appears in [**160**] and [**161**]. Milman's proof (with the estimate $n(M/b)^2$) is from [**374**]. The definition of the critical dimension and Theorem 1.10.5 are from [**388**].

The "volume-ratio theorem" was first proved by Kashin [**269**] in the case of ℓ_1^n, and later Szarek and Tomczak-Jaegermann introduced the notion of volume ratio (in [**485**] and [**486**]) and proved Theorem 1.10.12.

The ℓ-position and Pisier's inequality

Theorem 1.11.1 appears in [**323**]. The ℓ-norm was introduced by Figiel and Tomczak-Jaegermann in [**175**] and Theorem 1.11.2 appears in the same paper. The first proof of Pisier's inequality (Theorem 1.11.3) appeared in [**426**]; see also [**427**], [**428**] and [**430**].

Low M^*-estimate and Milman's quotient of subspace theorem

Milman's first proof of the low M^*-estimate was using Urysohn's inequality and appears in [**375**]. Milman's second proof from [**376**], which makes use of the isoperimetric inequality on S^{n-1}, is the one that we sketch in the text. Afterwards, Pajor and Tomczak-Jaegermann obtained the asymptotically optimal version in [**405**]. Finally, Gordon [**222**] proved that the theorem holds true with

$$f(\lambda) \geqslant \sqrt{1-\lambda}\left(1 + O\left(\frac{1}{(1-\lambda)n}\right)\right).$$

The quotient of subspace theorem is due to V. Milman [**377**].

Bourgain-Milman inequality and M-position

Mahler's conjecture appears in [**349**] and [**350**] in connection with some questions from the geometry of numbers. The conjecture has been verified for some classes of bodies: for the class of 1-unconditional convex bodies by Saint-Raymond [**449**] (the equality cases were clarified by Meyer [**362**] and Reisner [**440**]) and for the class of zonoids by Reisner in [**438**] and [**439**]. A short proof of Mahler's conjecture for zonoids was also given in [**223**].

The reverse Brunn-Minkowski inequality was proved by Milman in [**378**]. The Bourgain-Milman inequality is from [**112**]. Milman introduced the method of isomorphic symmetrization in [**380**]. The proof of the reverse Santaló inequality that we present in this section comes from that paper.

Kuperberg's proof of the reverse Santaló inequality appears in [**302**]. The proof of Nazarov can be found in [**398**].

Theorem 1.13.3 is from [**429**]. See also Pisier's book [**430**, Chapter 7].