CHAPTER 4

Convex Geometry and Functional Analysis

Keith Ball
Department of Mathematics, University College London, London, England
E-mail: kmb@math.ucl.ac.uk

Contents
Introduction ..................................................... 163
1. Convolution inequalities in convex geometry ........................................ 163
   1.1. The Brascamp–Lieb inequality ...................................... 164
   1.2. The reverse isoperimetric inequality ................................... 169
   1.3. The reverse Brascamp–Lieb inequality ................................. 171
   1.4. Higher dimensional projections ..................................... 173
2. Volumes of sections of convex bodies ..................................... 174
   2.1. Vaaler’s Theorem and its relatives .................................... 177
   2.2. Upper bounds for the volumes of sections of cubes ...................... 181
3. Plank problems ................................................. 182
   3.1. Bang’s lemma ............................................ 185
   3.2. The affine plank theorem ......................................... 185
   3.3. Nazarov’s solution to the coefficient problem .............................. 189
   3.4. Related results and open problems .................................... 192
References ..................................................... 193

HANDBOOK OF THE GEOMETRY OF BANACH SPACES, VOL. 1
Edited by William B. Johnson and Joram Lindenstrauss
© 2001 Elsevier Science B.V. All rights reserved
Introduction

This article describes three topics that lie at the intersection of functional analysis, harmonic analysis, probability theory and convex geometry. The first section consists principally of applications of harmonic analysis to convex geometry. The second section uses probabilistic inequalities to estimate volumes of convex bodies. The last section links an old problem in geometry, the plank problem, with some of the fundamental principles in functional analysis, and with the coefficient problem in classical harmonic analysis.

1. Convolution inequalities in convex geometry

The classical isoperimetric inequality states that among bodies of a given volume in $\mathbb{R}^n$, the Euclidean balls have least surface area. This principle seems to have been recognised, at least in two dimensions, by the ancients, and by the end of the last century, there were a number of proofs which worked in arbitrary dimension. Some time in the 1930’s, a reverse isoperimetric problem was formulated for convex bodies. Some care is needed in the formulation, since even convex bodies can have large surface area and small volume. The most natural way to pose the reverse problem is to consider affine-equivalence classes of convex bodies. The following solution to the problem appeared in [2] and [4].

**Theorem 1 (Reverse isoperimetric inequality).** If $C$ is a convex body in $\mathbb{R}^n$ and $T$ is a regular solid simplex in $\mathbb{R}^n$ then $C$ has an affine image $\tilde{C}$ for which

$$\text{vol}(\tilde{C}) = \text{vol}(T), \quad \text{but} \quad \text{vol}(\partial \tilde{C}) = \text{vol}(\partial T).$$

If $C$ is centrally symmetric, $T$ may be replaced by a cube $Q$.

The proof of Theorem 1 uses a sharp convolution inequality coming from harmonic analysis. There has been quite a lot of recent work dealing with other links between harmonic analysis/PDE and convex geometry: some years ago Jerison found an analogue of Minkowski’s existence theorem for harmonic measure and later he, Cafarelli and Lieb found a way to extend much of the classical Brunn–Minkowski theory, to capacity: but these developments will not be described in the present article.

It may not come as a surprise that harmonic analysis is linked with isoperimetric problems: after all, the classical isoperimetric inequality in $\mathbb{R}^n$ has an extension to compactly supported functions on $\mathbb{R}^n$ as the simplest Sobolev inequality:

$$c_n \|f\|_{n-1} \leq \|\nabla f\|_1,$$

for an appropriate constant $c_n$. This inequality belongs squarely within harmonic analysis. However, the relationship between the reverse isoperimetric problem and harmonic analysis is a bit more mysterious: in particular, a crucial role is played in the proofs by Gaussian densities.
It makes sense at this point to explain which affine image is used in the proofs of the assertions in Theorem 1. Of course, for any convex body, there is an affine image which is optimal: an affine image of the given volume, whose surface area is least. This affine image was characterised by Petty in the 60s, [35]; (some consequences of this characterisation were recently studied by Giannopoulos and Papadimitrakis, [22]). But the characterisation didn’t seem to be of much use for tackling the reverse isoperimetric problem. However, as explained in the article [26, Section 8], each convex body includes an unique ellipsoid of largest volume; this, maximal ellipsoid, was characterised by Fritz John. The affine image of a body $C$, that will be used to prove the theorem, is the one whose ellipsoid of largest volume is the standard Euclidean unit ball. If $C$ has the Euclidean ball as its maximal ellipsoid, then, according to John’s theorem, there are Euclidean unit vectors $(u_i)_1^n$ on the boundary of $C$ and positive numbers $(c_i)_1^n$, for which

$$
\sum_i c_i u_i = 0 \quad \text{(1)}
$$

and

$$
\sum_i c_i u_i \otimes u_i = I_n, \quad \text{(2)}
$$

where $I_n$ is the identity on $\mathbb{R}^n$.

The principal tool used in the proof of Theorem 1 is an inequality of Brascamp and Lieb [10] which among other things, is a sharp form of Young’s inequality for convolutions. Over the last 20 years there has been a steady growth in understanding of the best constants in classical inequalities like the Hausdorff–Young and Young inequalities. The paper of Lieb [30] contains very general results of this kind and describes the developments leading up to them. Recently, a delightful (and delightfully short), new proof of the Brascamp–Lieb inequality was found by Barthe [9]. He also proved a reverse analogue of the inequality which had been conjectured by the present author and which, like the forward inequality, has applications to the study of convex bodies. It should be remarked that Barthe’s new proof is sufficiently simple that he was able to determine cases of equality in his Theorems and in the applications to geometry.

The body of this chapter is divided up as follows. The Brascamp–Lieb inequality is discussed in Section 1.1 together with Barthe’s new proof. The proof of the reverse isoperimetric inequality is given in Section 1.2. Section 1.3 contains the reverse inequality and a brief example of how it may be used and Section 1.4 discusses a higher-dimensional version of these inequalities.

1.1. The Brascamp–Lieb inequality

Young’s inequality for convolution on a locally-compact group states that if $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{s}$ then for functions $f$ and $g$ in the appropriate spaces

$$
\|f * g\|_s \leq \|f\|_p \|g\|_q,
$$

where $\sigma$ is the identity on $\mathbb{R}^n$. 
On compact groups, where constant functions belong to $L_p$ spaces, the inequality is sharp: but on $\mathbb{R}$ it is not (except for special values of $p$ and $q$). In the 70’s Beckner and Brascamp and Lieb independently showed that the extremal functions in Young’s inequality are Gaussian densities: $x \mapsto e^{-ax^2}$. Using the duality between $L_p$ spaces, it is possible restate Young’s inequality as follows: if $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$ then

$$f \ast g \ast h(0) \leq \|f\|_p \|g\|_q \|h\|_r.$$

(The $L_2$ norm of $f \ast g$ is realised as an integral against a function in $L_r$.)

The left hand side of this inequality is

$$\int \int f(x)g(y-x)h(-y) \, dx \, dy. \quad (3)$$

This can be rewritten as

$$\int_{\mathbb{R}^2} f((x, e_1))g((x, e_2))h((x, e_3)) \, dx,$$

where $e_1 = (1, 0)$, $e_2 = (-1, 1)$ and $e_3 = (0, -1)$. The inequality of Brascamp and Lieb states that if $v_1, v_2, \ldots, v_m$ are vectors in $\mathbb{R}^n$, $f_1, f_2, \ldots, f_m$ are functions on the line and $(p_i)$ are numbers between 1 and $\infty$ satisfying $\sum \frac{1}{p_i} = n$ then the ratio

$$\frac{\int_{\mathbb{R}^n} \prod_{i=1}^m f_i((x, v_i)) \, dx}{\prod_{i=1}^m \|f_i\|_{p_i}}$$

is “maximised” by appropriate Gaussian densities. (There are degenerate cases in which the maximum is not attained.)

The value of the maximum is not easily computed since the $a_i$ are the solutions of nonlinear equations in the $p_i$ and $v_i$. This apparently unpleasant problem evaporates in the context of convex geometry: the inequality has a normalised form, introduced in [2], which “matches” Fritz John’s Theorem.

**Theorem 2 (Brascamp–Lieb).** Suppose $(u_i)$ is a sequence of unit vectors in $\mathbb{R}^n$ and $(c_i)$ is a sequence of positive numbers satisfying the John condition

$$\sum c_i u_i \otimes u_i = I_n.$$

If $f_i : \mathbb{R} \to [0, \infty)$ are measurable then

$$\int_{\mathbb{R}^n} \prod_{i=1}^m f_i((x, u_i))^{c_i} \, dx \leq \prod_{i=1}^m \left( \int f_i \right)^{c_i}. \quad (4)$$

In this version, the weights $c_i$ play the role of $1/p_i$. If one takes traces of the operators in the John condition one obtains, $\sum c_i = n$. The use of the John condition has magically
guaranteed that the best constant is 1. Moreover, the replacing of \( f_i^{p_i} \), by \( f_i \), ensures that the extremal functions are identical Gaussian densities:

\[
f_i(t) = e^{-t^2}
\]

for example. It is perhaps instructive to see why this is. The Fritz John condition, can be written

\[
\sum c_i \langle x, u_i \rangle^2 = |x|^2
\]

for each vector \( x \) in \( \mathbb{R}^n \). So with \( f_i(t) = e^{-t^2} \) for each \( i \),

\[
\int_{\mathbb{R}^n} \prod_{i=1}^{m} f_i(\langle x, u_i \rangle)^{c_i} \, dx = \int_{\mathbb{R}^n} e^{-|x|^2} \, dx = \left( \int_{\mathbb{R}} e^{-t^2} \, dt \right)^n = \prod_{i=1}^{m} \left( \int f_i \right)^{c_i}.
\]

The rest of this section is devoted to the recent proof of Theorem 2 found by F. Barthe. First a simple lemma.

**Lemma 1.** Let \((u_i)\) and \((c_i)\) be vectors and weights satisfying the Fritz John condition (2)

\[
\sum_{i} c_i u_i \otimes u_i = I_n.
\]

(1) If \( x = \sum c_i \theta_i u_i \), then \( |x|^2 \leq \sum c_i \theta_i^2 \).

(2) For any operator \( T \) on \( \mathbb{R}^n \),

\[
\det T \leq \prod_{i} \| Tu_i \|^{c_i}.
\]

(3) For any sequence \((\alpha_i)\) of positive numbers

\[
\det \left( \sum c_i \alpha_i u_i \otimes u_i \right) \geq \prod \alpha_i^{c_i}.
\]

**Remark.** The third statement is exactly what is needed to check that identical Gaussian densities are extremal among Gaussian densities, for the inequality of Theorem 2.

**Proof.** The first statement is just the Cauchy–Schwarz inequality.

For the second, it may be assumed that \( T \) is positive semi-definite and symmetric: say,

\[
T = \sum_{j} \lambda_j e_j \otimes e_j
\]

for some orthonormal basis \((e_j)\). Now, for each \( i \),

\[
\| Tu_i \|^2 = \sum_{j} \lambda_j^2 \langle u_i, e_j \rangle^2.
\]
and since $\sum_j (u_i, e_j)^2 = 1$, this is at least
\[ \prod_j \lambda_j^{2(u_i, e_j)^2}. \]

Hence
\[ \|Tu_i\| \geq \prod_j \lambda_j^{(u_i, e_j)^2}. \]

Raise this inequality to the power $c_i$ and take the product.

The third inequality is just the dual of the second. For each sequence of $\alpha_i$, there is an operator $T$ of determinant 1 for which
\[
\left( \det \left( \sum c_i \alpha_i u_i \otimes u_i \right) \right)^{1/n} = \frac{1}{n} \text{tr} \left( \sum c_i \alpha_i Tu_i \otimes Tu_i \right)
= \frac{1}{n} \sum c_i \alpha_i \|Tu_i\|^2
\geq \left( \prod_i \alpha_i \right)^{c_i/n} \left( \prod_i \|Tu_i\|^2 \right)^{c_i/n}
\]
and this is at least
\[ \left( \prod_i \alpha_i \right)^{c_i/n} \]
by part (2).

Now for Theorem 2.

**Proof.** Let $(u_i)$ and $(c_i)$ be vectors and weights satisfying the John condition, and $(f_i)$ a sequence of functions on $\mathbb{R}$ with
\[ \int f_i = 1 \]
for each $i$. The aim is to show that
\[ \int_{\mathbb{R}^n} \prod_{i=1}^m f_i((u_i, x))^{c_i} \, dx \leq 1. \]
Let $g$ be the function given by $g(s) = e^{-\pi s^2}$, chosen so that $\int g = 1$. Assume that each $f_i$ is strictly positive and smooth. For each $i$, define $T_i : \mathbb{R} \to \mathbb{R}$ by

$$\int_{-\infty}^{t} f_i(s) \, ds = \int_{-\infty}^{T_i(t)} g(s) \, ds.$$

Then, for each $t$ and $i$,

$$f_i(t) = g(T_i(t)).T_i'(t).$$

So

$$\int \prod f_i(\langle u_i, x \rangle)^{c_i} \, dx$$

$$= \int \prod g(T_i(\langle u_i, x \rangle))^{c_i} \prod T_i'(\langle u_i, x \rangle)^{c_i} \, dx$$

$$\leq \int \prod g(T_i(\langle u_i, x \rangle))^{c_i} \det \left( \sum c_i T_i'(\langle u_i, x \rangle) u_i \otimes u_i \right) \, dx$$

$$= \int \exp \left( -\pi \sum c_i (T_i(\langle u_i, x \rangle))^2 \right) \det \left( \sum c_i T_i'(\langle u_i, x \rangle) u_i \otimes u_i \right) \, dx$$

where the inequality follows from part (3) of the lemma.

The determinant in the last expression is the Jacobian $J(y, x)$ for the $n$-dimensional change of variables given by

$$y = \sum c_i T_i(\langle u_i, x \rangle).u_i.$$

With $y$ given by this equation, the first part of the lemma shows that

$$|y|^2 \leq \sum c_i (T_i(\langle u_i, x \rangle))^2,$$

and hence

$$e^{-\pi |y|^2} \geq \exp \left( -\pi \sum c_i (T_i(\langle u_i, x \rangle))^2 \right).$$

So the integral is at most

$$\int e^{-\pi |y|^2} J(y, x) \, dx = 1.$$  \qed

As was mentioned in the introduction, Barthe obtained conditions for equality in the above result, but these will not be discussed here.
1.2. The reverse isoperimetric inequality

The main purpose in this section is to describe the proofs of the two estimates in Theorem 1. Both the symmetric and the general case of the theorem are deduced from volume-ratio estimates. The volume-ratio of a convex body $K$ in $\mathbb{R}^n$, whose ellipsoid of maximal volume is $E$, is the number

$$\left( \frac{\text{vol}(K)}{\text{vol}(E)} \right)^{1/n}.$$ 

(See the article [21] in this collection.) Most of the effort goes into proving the following.

**Theorem 3 (Volume ratios).** The simplex has largest volume ratio among convex bodies of a given dimension, while among symmetric bodies, the cube is extremal.

Let's see how the volume-ratio estimate implies the reverse isoperimetric inequality, for example, in the case of the cube. The aim is to show that if the symmetric body $C$ has the standard Euclidean ball, $B^2_2$, as its maximal ellipsoid, then,

$$\text{vol}(\partial C) \leq 2^n \text{vol}(C)^{(n-1)/n}.$$ 

The volume ratio estimate guarantees that

$$\text{vol}(C) \leq 2^n$$

since this is the volume of the cube whose maximal ellipsoid is $B^2_2$. Now, from the fact that $C$ contains $B^2_2$, it follows that

$$\text{vol}(\partial C) = \lim_{\varepsilon \to 0} \frac{\text{vol}(C + \varepsilon B^2_2) - \text{vol}(C)}{\varepsilon} \leq \lim_{\varepsilon \to 0} \frac{\text{vol}(C + \varepsilon C) - \text{vol}(C)}{\varepsilon} = \text{vol}(C) \lim_{\varepsilon \to 0} \frac{(1 + \varepsilon)^n - 1}{\varepsilon} = n \text{vol}(C) \left( \text{vol}(C) \right)^{1/n} \left( \text{vol}(C) \right)^{(n-1)/n} \leq 2n \left( \text{vol}(C) \right)^{(n-1)/n}$$

as required. The argument for the simplex is similar.

The problem, then, is to prove Theorem 3.

**Proof.** For the symmetric case, the result is almost immediate from the normalised Brascamp–Lieb inequality and John's Theorem. If the maximal volume ellipsoid of $C$ is
$B^n_2$, then, since the vectors $u_i$ given by John's Theorem are points of contact between the boundaries of $C$ and the Euclidean ball, the set $C$ is included into the set $K = K(C)$.

$$\{x \in \mathbb{R}^n : |\langle x, u_i \rangle| \leq 1, \text{ for all } i \}.$$ 

So it suffices to estimate the volume of $K$. Now if $f_i$ is the indicator of the interval $[-1, 1]$, for each $i$, then the indicator $1_K$ of $K$ is exactly $\prod f_i(|\langle x, u_i \rangle| c_i)$. By Theorem 2,

$$\text{vol}(K) \leq \prod_1^m \left( \int f_i \right)^{c_i} = \prod_1^m 2^{c_i} = 2^n.$$ 

The volume ratio estimate for general bodies needs a bit more. A simplex whose maximal ellipsoid is $B^n_2$, has volume

$$\frac{n^{n/2}(n + 1)^{(n+1)/2}}{n!}.$$ 

So the aim is to show that if $B^n_2$ is the maximal ellipsoid in $C$, then the volume of $C$ is at most this number. As usual, let $u_i$ and $c_i$ be the unit vectors and the weights guaranteed by John's Theorem. In this case, the most we can say is that $C$ is included into

$$K = \{x \in \mathbb{R}^n : \langle x, u_i \rangle \leq 1, \text{ for all } i \}.$$ 

The absence of absolute values in the definition of this set makes it impossible to apply the Brascamp–Lieb inequality directly. The clue as to what to do is provided by the fact that a regular simplex of dimension $n$ sits more naturally (as a section of an orthant) in $\mathbb{R}^{n+1}$ than it does in $\mathbb{R}^n$.

Regard $\mathbb{R}^{n+1}$ as $\mathbb{R}^n \times \mathbb{R}$. For each $i$ define a unit vector $v_i$ in $\mathbb{R}^{n+1}$ by

$$v_i = \left( -\frac{\sqrt{n}}{\sqrt{n} + 1} u_i, \frac{1}{\sqrt{n} + 1} \right)$$ 

and a weight $d_i = \frac{n+1}{n} c_i$. Using the two conditions from John's Theorem, it is not too hard to check that

$$\sum d_i v_i \otimes v_i = I_{n+1}.$$ 

For each $i$, let

$$f_i(t) = \begin{cases} e^{-t} & \text{if } t \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and let $F$ be the function on $\mathbb{R}^{n+1}$,

$$F(y) = \prod f_i(|\langle y, v_i \rangle| d_i).$$
By Theorem 2 the integral of $F$, over $\mathbb{R}^{n+1}$ is at most 1.

Now, let $y = (x, r)$ be a point of $\mathbb{R}^{n+1}$. For each $i$,

$$\langle y, u_i \rangle = \frac{\sqrt{n}}{\sqrt{n+1}} \langle x, u_i \rangle + \frac{1}{\sqrt{n+1}} r.$$ 

Since $\sum c_i u_i = 0$, $\sum c_i \langle x, u_i \rangle = 0$ as well. Hence, if $r < 0$ there is at least one $i$ for which $\langle y, u_i \rangle$ is negative. This will ensure that $F(y) = 0$. On the other hand, if $r \geq 0$, the inner products $\langle y, u_i \rangle$ are all non-negative, precisely if

$$\langle x, u_i \rangle \leq \frac{1}{\sqrt{n}} r$$

for every $i$. The function $F$ is thus supported on a cone whose vertex is 0 and whose cross-section at distance $r$ from the vertex is $\frac{r}{\sqrt{n}} K$. At a point $(x, r)$, on this cross-section, the function’s value is

$$\exp \left( -\sum d_i \left( -\frac{\sqrt{n}}{\sqrt{n+1}} \langle x, u_i \rangle + \frac{1}{\sqrt{n+1}} r \right) \right)$$

and since $\sum c_i u_i = 0$ and $\sum d_i = n + 1$, this reduces to $e^{-\sqrt{n+1} r}$. The integral of $F$ is therefore

$$\int_0^\infty e^{-\sqrt{n+1} r} \left( \frac{r}{\sqrt{n}} \right)^n \frac{n! \text{vol}(K)}{n^{n/2}(n+1)^{(n+1)/2}}.$$

This completes the proof. \(\square\)

The article [4] includes a number of other applications of the Brascamp–Lieb inequality. In particular, it is shown that if $1 \leq p < \infty$, the space $\ell_p^n$ has largest volume-ratio among all $n$-dimensional subspaces of $L_p$. This fact is of most interest for $p = 1$.

### 1.3. The reverse Brascamp–Lieb inequality

At the time that the article [4] was written, the author conjectured a reverse form of the Brascamp–Lieb inequality. This conjecture was proved by Barthe [9]. Indeed, his proof of the forward inequality, described above, also gives the reverse inequality, if the roles of the $f_i$ and $g$ are interchanged. In fact, as Barthe pointed out, the two arguments can even be combined into one. The reverse inequality is the following.

**Theorem 4 (Barthe).** Suppose $(u_i)$ is a sequence of unit vectors in $\mathbb{R}^n$ and $(c_i)$ is a sequence of positive numbers satisfying the John condition

$$\sum c_i u_i \otimes u_i = I_n.$$
Suppose \( f_i : \mathbb{R} \to [0, \infty) \) and \( F : \mathbb{R} \to [0, \infty) \) are measurable, and that

\[
F(x) \geq \prod_{i} f_i(\theta_i)^{c_i}, \quad \text{whenever } x = \sum_{i} c_i \theta_i u_i.
\]

Then

\[
\int_{\mathbb{R}^n} F(x) \, dx \geq \prod_{i=1}^{m} \left( \int f_i \right)^{c_i}.
\]

The reverse inequality also has applications to convex geometry. Barthe used it to prove an extremal property of the simplex with respect to mean width. The following estimate for “outer volume ratio” provides a simpler illustration of its use.

**Theorem 5.** If \( C \) is a symmetric convex body in \( \mathbb{R}^n \), then \( K \) is included in an ellipsoid, \( \mathcal{E} \) with

\[
\frac{\text{vol}(\mathcal{E})}{\text{vol}(K)} \leq \frac{n}{2^n / n!}.
\]

**Proof.** The dual form of John’s Theorem guarantees that if the ellipsoid of minimal volume containing \( C \) is the ball \( B_2^n \), then there are unit vectors \( (u_i) \) in \( C \), and weights \( c_i \) satisfying the John condition (2). The problem is to show that the absolutely convex hull of these \( (u_i) \) has volume at least that of the \( \ell_1^n \) ball; namely \( 2^n / n! \). Let \( K \) be the absolutely convex hull and let \( \| \cdot \| \) be the norm corresponding to \( K \). By using polar coordinates it is easy to check that

\[
\text{vol}(K) = \frac{1}{n!} \int_{\mathbb{R}^n} e^{-\|x\|} \, dx.
\]

For any \( x \),

\[
\|x\| = \min \left\{ \sum_{i} |\lambda_i| : x = \sum_{i} \lambda_i u_i \right\} = \min \left\{ \sum_{i} c_i |\theta_i| : x = \sum_{i} c_i \theta_i u_i \right\}.
\]

Hence, if \( f_i(t) = e^{-|t|} \) for each \( i \),

\[
e^{-\|x\|} = \max \left\{ \prod_{i} f_i(\theta_i)^{c_i} : x = \sum_{i} c_i \theta_i u_i \right\}.
\]

By Barthe’s Theorem,

\[
\int_{\mathbb{R}^n} e^{-\|x\|} \, dx \geq 2^n
\]

as required. ☐
1.4. Higher dimensional projections

In his paper [30], Lieb produced a greatly generalised version of the Brascamp–Lieb inequality, which, in the normalised form, reads:

**Theorem 6 (Lieb).** If \((P_i)\) are orthogonal projections on \(\mathbb{R}^n\) and \((c_i)\) are positive numbers satisfying \(\sum c_i P_i = I_n\), and for each \(i\), \(f_i\) is a non-negative function on the image of \(P_i\), then

\[
\int_{\mathbb{R}^n} \prod_{i=1}^m f_i(P_i x)^{c_i} \, dx \leq \prod_{i=1}^m \left( \int f_i \right)^{c_i}.
\]

This higher-dimensional version also has a simpler proof, recently found by Barthe [9]. The basic idea is the same as for the 1-dimensional case: to use a homeomorphism of Euclidean space to represent a given density in terms of the Gaussian. But on higher dimensional spaces, there is a wide choice of such homeomorphisms and the extra problem is to find one which is appropriate to the inequality. Barthe does this by invoking a theorem of Brenier [11] (see also McCann [32]). The results of Brenier, McCann and others are part of an active current investigation of the transportation of measures.

The special case in which \(m = n\) and the \(P_i\) are the 1-codimensional projections onto the coordinate hyperplanes in \(\mathbb{R}^n\), goes back (essentially) to Loomis and Whitney. They proved that if \(K\) is compact in \(\mathbb{R}^n\) then

\[
\text{vol}(K) \leq \prod_{1}^{n} \text{vol}(P_i K)^{1/(n-1)}.
\]

This follows from Lieb’s Theorem applied to the functions

\[
f_i(x) = 1_{P_i K}(x),
\]

whose product on \(\mathbb{R}^n\) is equal to 1 at each point of \(K\). Although their result is specific to indicator functions, their argument also gives the case of general functions. The reason that Lieb’s theorem is much more difficult, lies not in the passage from sets to functions, but in the fact that his projections \(P_i\) do not commute with one another.

The Loomis–Whitney inequality already looks a bit like an isoperimetric inequality since it estimates the volume of \(K\) in terms of the volumes of 1-codimensional sets derived from \(K\). In fact, a simple generalisation of the Loomis–Whitney inequality is the main technical tool in Gagliardo’s proof of the Sobolev Embedding Theorem, which goes back to the late 50s. One can begin to see how Loomis–Whitney could be used, by noticing that it immediately implies a version of the isoperimetric inequality (without the right constant). If \(K\) is compact and convex, then by Loomis–Whitney,

\[
\text{vol}(K)^{(n-1)/n} \leq \prod_{1}^{n} \text{vol}(P_i K)^{1/n} \leq \frac{1}{n \cdot \sum_{1}^{n} \text{vol}(P_i K)}.
\]
Now average the inequality over all choices of coordinate systems. The result is

$$\text{vol}(K)^{(n-1)/n} \leq \text{Average}(\text{vol}(P_K)),$$

where the average is taken over all orthogonal projections of codimension 1. The last expression is proportional to the surface area of $K$ (with a constant of proportionality depending upon $n$, but not upon $K$).

2. Volumes of sections of convex bodies

In the previous section it was explained that the simplex has the largest volume ratio of any convex body, and that the cube has the largest of any symmetric body. For a convex body whose maximal ellipsoid is known, these volume ratio estimates automatically provide upper bounds for the volume of the body. It turns out that such upper bounds are pretty accurate, as long as the body has "few" faces, relative to the dimension. (The precise meaning of "few" will become clear later.) This section of the article describes some lower bounds for the volumes of bodies with few faces, which complement the upper bounds of the last section and indicate how sharp they are.

The oldest discovery in the Geometry of Numbers is Minkowski’s Box Theorem which guarantees that every symmetric convex body in $\mathbb{R}^n$, of volume at least $2^n$, contains a non-zero point of the integer lattice $\mathbb{Z}^n$. An immediate consequence is that if $A = (a_{ij})$ is an $n \times n$ matrix of determinant at most 1, then there is a sequence $(z_j)_{j=1}^n$ of integers, not all 0, so that for each $i$,

$$\left| \sum_i a_{ij} z_j \right| \leq 1.$$

A natural question, which arises in applications to number theory, asks what happens if the matrix is not square. A very natural conjecture, which was publicised by A. Good, went as follows.

**Conjecture 1.** Let $A = (a_{ij})$ be an $m \times n$ matrix satisfying

$$\det(A^* A) \leq 1. \quad (4)$$

Then there is a sequence $(z_j)_{j=1}^n$ of integers, not all 0, so that for each $i$,

$$\left| \sum_i a_{ij} z_j \right| \leq 1.$$

The problem is to show that if $A$ is an $m \times n$ matrix satisfying (4) then the set

$$K = \left\{ x : \sum_j a_{ij} x_j \leq 1, \text{ for all } i \right\}$$
has volume at least $2^n$. The inequality (4) states precisely that the map from $\mathbb{R}^n$ to $\mathbb{R}^m$ given by $A$ does not increase $n$-dimensional volume. The image of $K$ under this map, is the $n$-dimensional slice of the cube $[-1, 1]^m$ by the subspace $A(\mathbb{R}^n)$. Thus, as Good himself pointed out, the conjecture can be restated: for each $m$ and each $n \leq m$, every $n$-dimensional slice through the centre of the cube $[-1, 1]^m$ has volume at least $2^n$. For $n = m - 1$, this conjecture was proved by Hensley [24] and for arbitrary $n$, by Vaaler [38].

Yet another formulation of Vaaler’s Theorem reads as follows:

**Theorem 7 (Vaaler).** Suppose $(v_i)^m_1$ is a sequence of vectors in $\mathbb{R}^n$ and we set

$$r^2 = \frac{1}{n} \sum_{i=1}^{m} |v_i|^2.$$ 

Then

$$\text{vol}\left( \{ x : \langle x, v_i \rangle \leq 1, \text{ for all } i \} \right)^{1/n} \geq \frac{2}{r}.$$

Some years after Vaaler’s result, several authors, independently, found an analogue, in which the maximum length of the $v_i$ is controlled, instead of the sum of the squares of the lengths. It states:

**Theorem 8.** There is a constant $\delta > 0$ (independent of everything) so that if $(v_i)^m_1$ is a sequence of unit vectors in $\mathbb{R}^n$, then

$$\text{vol}\left( \{ x : \langle x, v_i \rangle \leq 1, \text{ for all } i \} \right)^{1/n} \geq \frac{\delta}{\sqrt{\log(1 + \frac{m}{n})}}.$$ 

This statement can be extracted from some results of Carl, and this was done by Carl and Pajor [14]: it was proved independently by Gluskin [23] and by Barany and Furédi [7]; and it was rediscovered by Bourgain, Lindenstrauss and Milman [12]. In fact Gluskin’s method gives something a bit stronger: it is not necessary to assume a “uniform” estimate on the lengths of the vectors.

**Theorem 9 (Gluskin).** If $(v_k)$ is a sequence of vectors in $\mathbb{R}^n$ satisfying

$$|v_k| \leq \frac{1}{2 \sqrt{\log(1 + \frac{k}{n})}}$$

for each $k$, then

$$\text{vol}\left( \{ x : \langle x, v_i \rangle \leq 1, \text{ for all } i \} \right)^{1/n} \geq \frac{1}{2}.$$
It should be remarked immediately, that if one replaces the expressions \( \log(1 + \frac{k}{n}) \) by \( \log(1 + k) \), then the statement becomes almost trivial. The improvement given by the theorem is of interest if the number of vectors is fairly small compared to the dimension: especially if it is no more than a fixed multiple of the dimension. This is also the range in which Vaaler’s Theorem is most useful (and is typically used). Gluskin’s result (in the form stated above) includes Vaaler’s, apart from the value of the “constant” involved. To see why, observe that if

\[
\frac{1}{n} \sum |v_k|^2 \leq 1
\]

and the \( v_k \) are in decreasing order of norm, then

\[
|v_k| \leq \sqrt{\frac{n}{k}} \leq \frac{1}{\sqrt{\log(1 + \frac{k}{n})}}.
\]

Strangely enough, the proofs of Theorems 7 and 9, like the proofs of the volume ratio estimates of the last chapter, use comparisons between Lebesgue measure and an appropriate Gaussian measure. In the case of Theorem 9, the crucial property of Gaussian measure is captured by Sidak’s Lemma. A *symmetric slab* in \( \mathbb{R}^n \) is a set of the form

\[
\{ x \in \mathbb{R}^n : |\langle x, u \rangle| \leq t \}
\]

for some vector \( u \) and positive number \( t \).

**Lemma 2 (Sidak).** If \( \gamma \) is standard Gaussian measure on \( \mathbb{R}^n \) and \( (S_i) \) is a sequence of symmetric slabs then

\[
\gamma \left( \bigcap S_i \right) \geq \prod \gamma (S_i).
\]

Sidak’s Lemma (and more especially its proof) solves a special case of a well-known and rather intriguing problem: the correlation problem for Gaussian measure.

**Question 1.** Is it true that if \( \gamma \) is standard Gaussian measure on \( \mathbb{R}^n \) and \( K \) and \( L \) are symmetric convex bodies in \( \mathbb{R}^n \), then

\[
\gamma (K \cap L) \geq \gamma (K) \gamma (L)?
\]

For a discussion of this problem see, e.g., [36].

In his article on the central sections of the cube, Hensley not only showed that the volumes of the 1-codimensional sections are at least 1, but also that they are at most 5, independent of the dimension. At first sight this seems very surprising; but in a second article [25] Hensley showed that all convex bodies have a similar property; each convex body (after an appropriate linear map), has 1-codimensional sections with “almost constant”
volume. This fact is of considerable importance in a number of situations. Hensley also asked for the sharp upper bound for the volumes of 1-codimensional sections of the cube, conjecturing the value $\sqrt{2}$. The present author proved this in [1], and in a later paper [2] found the optimal upper bound for $k$-codimensional sections, $(\sqrt{2})^k$. These results will not be proved here, but in the short Section 2.2, a general upper bound for volumes of sections of the cube will be deduced from the Brascamp–Lieb inequality of the last chapter.

An unexpected byproduct of the upper bound, $\sqrt{2}$, for 1-codimensional sections of the cube, was a “concrete” solution of the so-called Busemann–Petty problem. In [13], these authors asked the following question.

QUESTION 2. Suppose that $K$ and $L$ are symmetric convex bodies in some Euclidean space with the property that, for each 1-codimensional subspace $H$,

$$\text{vol}(H \cap K) \leq \text{vol}(H \cap L).$$

Does it follow that

$$\text{vol}(K) \leq \text{vol}(L)?$$

At the time, there was a widespread belief that the answer should be yes: among other things, it was known that if two such bodies have sections of equal volume, then they are the same body. So it came as something of a surprise, when Larman and Rogers [29] constructed a random symmetric perturbation of the Euclidean ball in $\mathbb{R}^{12}$, whose slices all had smaller volume than those of a ball of equal volume. However, with the knowledge that the unit cube has sections of volume at most $\sqrt{2}$, the problem loses some of its mystique. When the dimension is large, the Euclidean ball of volume 1 has 1-codimensional sections of volume about $\sqrt{c}$. Thus, for large enough dimension, a cube and an Euclidean ball of slightly smaller volume, provide a counterexample for the Busemann–Petty problem. Following this observation, Giannopoulos [19] pointed out that, in fact, there are extremely simple concrete counterexamples, in dimensions as low as 7. It is now known that the Busemann–Petty problem has a negative answer if and only if the dimension is at least 5. The positive answer in dimension 3 was provided by Gardner [16] and that in dimension 4, by Zhang [39]. Recently Koldobsky [28] has developed a new approach to the problem which greatly simplifies and clarifies matters. This led to a definitive unified solution to the problem in all dimensions, which will appear in [18]. (The topic was also dealt with at some length, in the book by Gardner [17].)

This chapter is organised as follows. Vaaler’s Theorem, Sidak’s Lemma and Gluskin’s Theorem are proved in the first section. The second section gives a brief account of some upper bounds for volumes of sections of cubes.

2.1. Vaaler’s Theorem and its relatives

As explained in the introduction all the results in this section depend upon comparisons between Lebesgue and Gaussian measures. The comparison method was studied in detail by Kanter [27]. It begins with the following definition.
If $\mu$ and $\nu$ are two probability measures on some Euclidean space $\mathbb{R}^n$, say that $\mu$ is more peaked than $\nu$ if for every symmetric convex set $K$ in $\mathbb{R}^n$,

$$
\mu(K) \geq \nu(K).
$$

In order to prove Vaaler's Theorem, the aim will be to show that Lebesgue measure on the cube is more peaked than an appropriate Gaussian measure. To begin with, let $\mu$ be Lebesgue measure on the interval $[-\frac{1}{2}, \frac{1}{2}]$ and let $\nu$ be the unique Gaussian probability on $\mathbb{R}$ whose density

$$
f(t) = e^{-\pi t^2}
$$
satisfies $f(0) = 1$. Clearly in this case, $\mu$ is more peaked than $\nu$. Kanter's basic lemma, shows that peaking-order is preserved under the formation of product measures, at least in the presence of some restriction. This restriction involves a notion of unimodality: several such notions have been considered. The most versatile seems to be the one used here, which was introduced by Barthe [8].

A probability measure on $\mathbb{R}^n$ will be called unimodal if it has a density $f$, which can be expressed as the increasing limit of a sequence of functions, each of which is a positively weighted sum

$$
\sum p_i 1_{K_i}
$$
of characteristic functions, $1_{K_i}$, of symmetric convex sets. Clearly, Lebesgue measure on the cube and Gaussian measures are unimodal.

**Lemma 3 (Kanter).** If $\mu_1$ and $\mu_2$ are probabilities on $\mathbb{R}^n$, with $\mu_1$ being more peaked than $\mu_2$, and $\nu$ is a unimodal probability on $\mathbb{R}^m$, then $\mu_1 \otimes \nu$ is more peaked than $\mu_2 \otimes \nu$.

**Proof.** The problem is to show that for every symmetric convex $K$ in $\mathbb{R}^{n+m}$,

$$
\mu_1 \otimes \nu(K) \geq \mu_2 \otimes \nu(K).
$$

By the unimodality of $\nu$, it may be assumed that the density of $\nu$ is the characteristic function of a symmetric convex set $C$. If $\tilde{K}$ is the intersection of $K$ with the cylinder $\mathbb{R}^n \times C$, then for any probability $\mu$ on $\mathbb{R}^n$,

$$
\mu \otimes \nu(K) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} 1_K(x, y) 1_C(y) dy \, d\mu(x) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} 1_{\tilde{K}}(x, y) dy \, d\mu(x).
$$

It is a standard consequence of the Brunn–Minkowski inequality that the function $g$, defined on $\mathbb{R}^n$ by

$$
g(x) = \int_{\mathbb{R}^m} 1_{\tilde{K}}(x, y) dy
$$
has a concave logarithm. Also, $g$ is an even function, because the set $\tilde{K}$ is symmetric. From this it follows that $g$ can be approximated by positive combinations of the characteristic functions of symmetric convex sets in $\mathbb{R}^n$. The fact that $\mu_1$ is more peaked than $\mu_2$ now ensures that

$$\int_{\mathbb{R}^n} g(x) \, d\mu_1(x) \geq \int_{\mathbb{R}^n} g(x) \, d\mu_2(x),$$

as required. $\square$

By applying Kanter’s Lemma repeatedly it is easy to see that if $\mu$ and $\nu$ are the uniform and Gaussian probabilities on $\mathbb{R}$, described above, then

$$\mu \otimes \ldots \otimes \mu$$

is more peaked on $\mathbb{R}^n$ than

$$\nu \otimes \ldots \otimes \nu.$$

From this one gets Vaaler’s Theorem as follows.

**Proof.** The task is to show that if $H$ is a $k$-dimensional subspace of $\mathbb{R}^n$ and $Q$ is the unit cube $[-\frac{1}{2}, \frac{1}{2}]^n$ then the $k$-dimensional volume of $H \cap Q$ is at least 1. As explained above, if $K$ is any symmetric convex body in $\mathbb{R}^n$, then $K \cap Q$ has volume at least as large as

$$\int_K e^{-\pi |x|^2} \, dx.$$

By approximating the subspace $H$ by very thin convex “tubes”, one obtains that the $k$-dimensional volume of $H \cap Q$ is at least the integral over $H$ (with respect to $k$-dimensional measure) of $e^{-\pi |x|^2}$; and this is 1. $\square$

The reader is invited to check that the two other formulations of Vaaler’s Theorem, described in the introduction to this chapter, follow from the statement just proved. An application of Kanter’s Lemma, to the volumes of sections of $\ell_p^n$ balls other than cubes appears in the article of Meyer and Pajor [33].

It was mentioned in the introduction, that the proof of Sidak’s Lemma is closely related to the Gaussian correlation problem. Sidak’s Lemma says that the Gaussian measure of an intersection of symmetric slabs, is at least the product of the Gaussian measures of the slabs. It is clear that this statement can be deduced by induction from the following lemma.

**Lemma 4.** If $K$ is a symmetric convex body in $\mathbb{R}^n$, $S$ is a symmetric slab and $\gamma$ is standard Gaussian measure on $\mathbb{R}^n$, then

$$\gamma(K \cap S) \geq \gamma(K) \gamma(S).$$
It was pointed out by A. Giannopoulos [20] that this lemma can readily be recast so as to fit into the framework described earlier. Define a probability \( \mu \) on \( \mathbb{R}^n \) by

\[
\mu(A) = \frac{\gamma(A \cap S)}{\gamma(S)}.
\]

The lemma states that \( \mu \) is more peaked than \( \gamma \). By writing \( \mathbb{R}^n \) as the direct sum of the 1-dimensional space spanned by the vector perpendicular to the slab \( S \) and the \( (n - 1) \)-dimensional space parallel to \( S \), the measure \( \mu \) can be regarded as the product of an \( (n - 1) \)-dimensional Gaussian measure and a 1-dimensional measure. The latter has a density of the form

\[
s \mapsto (\text{constant}) \, e^{-s^2/2} 1_{[-t, t]}(s),
\]

obtained by restricting the Gaussian density to a symmetric interval, and then renormalising. Plainly, this 1-dimensional probability is more peaked than the standard Gaussian. Therefore, by Kanter’s Lemma, \( \mu \) is more peaked than \( \gamma \).

Sidak’s Lemma gives Gluskin’s Theorem in the following way.

**Proof.** Assume that for each \( k \), \( v_k \) has norm at most \((4 \log(1 + \frac{k}{n}))^{-1/2}\) and let

\[
K = \{x: |\langle x, v_k \rangle| \leq 1, \text{ for all } k\} = \bigcap S_k,
\]

where for each \( k \), \( S_k \) is the slab

\[
\{x \in \mathbb{R}^n: |\langle x, v_k \rangle| \leq 1\}.
\]

Clearly

\[
\text{vol}(K) \geq (\sqrt{2\pi})^n \gamma(K),
\]

and by Sidak’s Lemma, this is at least

\[
(\sqrt{2\pi})^n \prod_k \gamma(S_k).
\]

For each \( k \) the slab \( S_k \) has width \( 2t \) where

\[
t = \frac{1}{\|u_k\|} \geq \sqrt{4 \log \left(1 + \frac{k}{n}\right)}.
\]

Now a slab of width \( 2t \) has Gaussian measure

\[
\frac{1}{\sqrt{2\pi}} \int_{-t}^{t} e^{-s^2/2} \, ds
\]
and it is well-known that this is always at least $1 - e^{-t^2/2}$. Hence

$$\text{vol}(K) \geq (\sqrt{2\pi})^n \prod_k \left(1 - e^{-2 \log(1 + \frac{k}{n})} \right) = (\sqrt{2\pi})^n \prod_k \left(1 - \frac{1}{(1 + \frac{k}{n})^2} \right).$$

The product can easily be estimated since

$$\sum \log \left(1 - \frac{1}{(1 + \frac{k}{n})^2} \right) \geq \int_0^\infty \log \left(1 - \frac{1}{(1 + \frac{x}{n})^2} \right) \, dx = -2n \log 2. \quad \square$$

**Remark.** It is clear that the preceding argument can be adapted to show that if the $v_k$ satisfy the stronger estimate

$$|v_k| \leq \frac{\delta}{\sqrt{\log(1 + k)}},$$

then the Gaussian measure of the intersection of the corresponding slabs is at least $1/2$ (not just $1/2^n$). This implies that if $\| \cdot \|$ is the norm whose unit ball is the intersection of these slabs, then the random variable $\| \cdot \|$ on the probability space $(\mathbb{R}^n, \gamma)$ has a median which is at most 1. In fact, it is not too hard to show that under this assumption on $|v_k|$, even the mean

$$\int_{\mathbb{R}^n} \|x\| \, d\gamma(x)$$

is at most a constant. Using his remarkable majorising measure theorem, [37] Talagrand has shown that the “converse” is true: if the mean of $\| \cdot \|$ is at most 1, then the unit ball contains a convex body which is the intersection of slabs, $(S_k)$, whose widths are at least $\delta \sqrt{\log(1 + k)}$.

### 2.2. Upper bounds for the volumes of sections of cubes

The purpose of this section is to explain how the Brascamp–Lieb inequality, quickly provides upper estimates which are complementary to those of the previous section. In fact, the argument here is really no more than a disguised form of the volume ratio estimate for symmetric bodies; but the statement looks a bit different.

**Theorem 10** (Sections of the cube). Every $k$-dimensional section of an $n$-dimensional unit cube has volume at most

$$\left(\frac{n}{k}\right)^{k/2}.$$
The estimate is sharp whenever $k$ divides $n$; but even in the cases where it is not sharp, the estimate has roughly the right "shape". As mentioned above, the alternative estimate $(\sqrt{2})^{n-k}$ is known, and this is sharp whenever $k \geq n/2$.

**Proof.** Let $Q$ be the unit cube, $[-\frac{1}{2}, \frac{1}{2}]^n$. Let $H$ be a $k$-dimensional subspace of $\mathbb{R}^n$ and let $P$ be the orthogonal projection from $\mathbb{R}^n$ onto $H$. For each $i$, $1 \leq i \leq n$, let

$$v_i = Pe_i,$$

where $e_1, e_2, \ldots, e_n$ is the standard basis of $\mathbb{R}^n$, and let

$$c_i = |v_i|^2.$$

Since $P$ acts as the identity $I_H$ on $H$,

$$I_H = \sum_i v_i \otimes v_i = \sum_i c_i u_i \otimes u_i$$

where the $u_i$ are unit vectors. If $x \in H$, then $x \in Q$ if and only if, $|\langle x, v_i \rangle| \leq \frac{1}{2}$, for every $i$. In terms of the $u_i$, $H \cap Q$ is the set

$$\left\{ x \in H : |\langle x, u_i \rangle| \leq \frac{1}{2\sqrt{c_i}} \right\}.$$

An application of Theorem 2 in the space $H$ gives,

$$\text{vol}(H \cap Q) \leq \prod_i \left(\frac{1}{\sqrt{c_i}}\right)^{c_i}. \quad (5)$$

Since $\sum c_i = k$, the convexity of the function $c \mapsto c \log c$ shows that the right-hand side of (5) is maximised when all $c_i$ are equal: and hence equal to $k/n$. \hfill \Box

**3. Plank problems**

In the early 1930's, Tarski asked the following question (in connection with the Banach–Tarski paradox). Let us call the region between two parallel hyperplanes in $\mathbb{R}^n$, a plank.

QUESTION 3. If a convex body of minimum width 1, is covered by a union of planks, must the widths of the planks add up to at least 1?

Clearly it is possible to cover with planks of total width 1, by aligning them perpendicular to the direction in which the body has minimum width.

Tarski himself showed that the answer to the question is yes, if the body is an Euclidean ball in 2 or 3 dimensions. During the 40s, the problem gained a certain notoriety. It was
finally solved in the affirmative by T. Bang in 1951 [6]. The basic lemma used by Bang, is proved below. Though apparently simple, this lemma turns out to be the starting point for several surprising developments.

At the end of his paper, Bang asked a further question, which is a good deal more natural than Tarski’s original one, since it is affine invariant. Given a convex body $K$, the relative width of a plank, $S$, is the width of $S$, divided by the width of $K$ in the direction perpendicular to $S$. Bang’s affine plank problem asks:

**QUESTION 4.** If a convex body is covered by a union of planks, must the relative widths of the planks add up to at least 1?

Plainly, an affirmative answer to Bang’s question, strengthens Bang’s result. The general case of this affine plank problem is still open, but the most important case, in which the body is assumed to be centrally symmetric, was proved in [3]. If $K$ is such a body, then it may be regarded as the unit ball of some finite-dimensional normed space. A plank in a normed space is a set of the form

$$\{x \in \mathbb{R}^n : |\phi(x) - m| \leq w\}$$

for some continuous linear functional $\phi$ on the space. If the norm of $\phi$ is 1, then the relative width of the plank is $w$. The result in [3] applies to infinite-dimensional spaces as well as finite: in the language of functional-analysis, it states:

**THEOREM 11 (Ball).** If $\{\phi_i\}$ is a sequence of linear functionals, of norm 1, on a Banach space $X$, $m_i$ is a sequence of real numbers and $(w_i)$ is a sequence of positive numbers satisfying

$$\sum_{i=1}^{\infty} w_i < 1,$$

then there is a point $x$ in the unit ball of $X$ for which

$$|\phi_i(x) - m_i| > w_i \quad \text{for every } i.$$

This formulation of the theorem makes it clear that the result simultaneously extends the Hahn–Banach Theorem and sharpens the Uniform Boundedness Principle. It also suggests a similarity between the plank problem and the “Coefficient Problem” in harmonic analysis. For any sequence of numbers $(c_k)$ satisfying

$$\sum |c_k|^2 < \infty,$$

the function

$$t \mapsto \sum c_k e^{ikt}$$
is square-integrable on the circle. For an integrable function $f$ on the circle, let $\hat{f}(k)$ denote
the $k$th Fourier coefficient of $f$. The de Leeuw–Kahane–Katznelson Theorem [15] states
that for any sequence $(c_k)$ of positive numbers satisfying (6), there is a bounded function
$f$ on the circle for which
$$|\hat{f}(k)| > c_k$$
for every $k$. Thus, it is impossible to distinguish bounded functions, from other square-
integrable functions, just by looking at the sizes of their Fourier coefficients.

It is natural to wonder just what conditions on a sequence, $(\psi_k)$, of functions will ensure
that for any square-summable sequence $(c_k)$ of coefficients, there is a bounded function $f$
for which
$$|(f, \psi_k)| > c_k$$
for every $k$. The de Leeuw–Kahane–Katznelson Theorem is proved using modifications of
random functions such as
$$\sum \varepsilon_k c_k \psi_k,$$
where $(\varepsilon_k)$ is a sequence of independent choices of sign. The method requires at least
some control of the functions $\psi_k$ in a norm “bigger than 2”: a uniform estimate on their
$L_{2+\epsilon}$ norms, for example. More crucially, it depends heavily upon the orthogonality of the
characters. The method, therefore, leaves open several natural questions; for example, what
happens when the $\psi_k$ are the trigonometric characters, but it is required that the function
$f$ be supported on some restricted part of the circle?

The situation was resolved in a rather startling way by the following theorem of Nazarov
[34]. (All functions and coefficients are, from now on, assumed to be real, since this results
in no loss of generality.)

**Theorem 12 (Nazarov).** Let $(\psi_k)$ be a sequence of unit functions in $L_1$ (of a measure
space) which satisfy an upper-2-estimate: that is, there is a constant $M$ so that
$$\left\| \sum a_k \psi_k \right\|_1 \leq M \left( \sum a_k^2 \right)^{1/2}$$
for any sequence $(a_k)$. Then, for any square-summable sequence $(c_k)$, there is a bounded
function $f$ satisfying
$$\left| \int f \psi_k \right| > c_k$$
for every $k$.

Nazarov’s Theorem is proved in Section 3.3. Although the proof of this theorem does not
formally depend upon the other results in this chapter, it will be clear that the methods used
are closely related. The basic lemma used by Bang, is proved in Section 3.1. Section 3.2 contains a proof of Theorem 11, which uses Bang's Lemma in its original form. Finally, the chapter closes by mentioning some related results and open problems.

3.1. Bang's lemma

The crucial lemma used by Bang in his solution of the original plank problem is the following. This lemma will be used in Section 3.2.

**Lemma 5.** Let \((u_i)\) be a sequence of unit vectors in an Euclidean space, \((m_i)\) a sequence of reals and \((w_i)\) a sequence of positive numbers. Then there is a choice of signs \((\varepsilon_i)\) for which the vector

\[ x = \sum \varepsilon_i w_i u_i \]

satisfies

\[ |\langle x, u_i \rangle - m_i| \geq w_i \] for every \(i\).

**Proof.** The following argument deals with the special case in which all the \(m_i\) are 0. The general case is only marginally more complicated: see, e.g., [3] for a proof.

Choose the sequence of signs for which the vector \(x\) is longest. Now, suppose \(k\) is between 1 and \(n\) and let

\[ y = \sum_{i \neq k} \varepsilon_i w_i u_i \]

so that \(x = y + \varepsilon_k w_k u_k\). By the choice of \(\varepsilon_k\), \(x\) has norm at least as large as that of \(y - \varepsilon_k w_k u_k\). This means that

\[ \varepsilon_k \langle y, u_k \rangle \geq 0. \]

This in turn guarantees that the number

\[ \langle x, u_k \rangle = \langle y, u_k \rangle + \varepsilon_k w_k = \varepsilon_k (\langle y, u_k \rangle + w_k) \]

is at least \(w_k\) in size. This is what was wanted. \(\square\)

3.2. The affine plank theorem

The aim in this section is to prove Theorem 11, at least in the special case in which all \(m_i\) are zero. The argument in the general case is no different; it merely uses the more general case of Bang's lemma.
Most of the work goes into solving the problem on a finite-dimensional space. In this case it may be assumed that there are only finitely many functionals \((\phi_i)_{i=1}^n\), and it suffices to prove that if

\[
\sum w_i = 1,
\]

then there is a point \(x\) in the unit ball for which

\[
|\phi_i(x)| \geq w_i
\]

for each \(i\). By slicing the planks into thin "sheets" it may also be assumed that all the \(w_i\) are the same, i.e., that each is equal to \(1/n\). For each \(i\), let \(x_i\) be a vector of norm 1 for which

\[
\phi_i(x_i) = 1.
\]

Let \(A = (a_{ij})\) be the matrix given by

\[
a_{ij} = \phi_i(x_j)
\]

so that each diagonal of \(A\) is 1. It suffices to find a sequence \((\lambda_j)\) for which \(\sum |\lambda_j| \leq 1\) and

\[
\left| \sum_j a_{ij} \lambda_j \right| \geq \frac{1}{n}
\]

for each \(i\); for, given such a sequence, the vector

\[
x = \sum_j \lambda_j x_j,
\]

has the required properties.

The main task will therefore be to prove the following combinatorial statement:

**Theorem 13.** Let \(A = (a_{ij})\) be an \(n \times n\) real matrix with 1's on the diagonal. Then there is a sequence \((\lambda_j)\) for which

\[
\sum |\lambda_j| \leq 1,
\]

but

\[
\left| \sum_j a_{ij} \lambda_j \right| \geq \frac{1}{n}
\]

for each \(i\).
The argument below will guarantee rather more; namely, that the sequence \((\lambda_j)\) satisfies
\[
\sum \lambda_j^2 \leq \frac{1}{n}.
\] (8)
This implies (7) by Cauchy–Schwarz. The advantage of aiming for (8) is that one can apply Hilbert space methods to modify the matrix \(A\). From now on, a matrix will be called positive, if it is symmetric and positive semi-definite. The modification of a matrix needed here, is described in the following lemma.

**Lemma 6.** Let \(A\) be an \(n \times n\) matrix, of which each row contains a non-zero entry. Then there is a diagonal matrix \(\Theta\), with non-negative entries, and an orthogonal matrix \(U\), so that the matrix
\[
H = \Theta AU
\]
is positive and has all diagonal entries equal to 1.

Lemma 6 can be proved, either from Brouwer’s Fixed Point Theorem or, more directly, by a variational argument: see [3]. The entries in the diagonal matrix, \(\Theta\), whose existence is guaranteed by this lemma, are not easily expressed in terms of the original matrix \(A\). For the present purpose it is necessary to understand how large these entries can be. The next lemma will provide an estimate for them.

**Lemma 7.** Let \(H\) be a positive matrix with 1’s on the diagonal and \(U\) an orthogonal matrix. Then the diagonal entries of the product \(HU\) satisfy
\[
\sum (HU)_{ii}^2 \leq n.
\]

**Remark.** The fact that the trace of \(HU\) is at most \(n\), is standard. The above, quadratic, estimate looks somewhat unusual. The elegant proof below was shown to me by F. Lust-Piquard.

**Proof.** For each \(i\), let \(\gamma_i = (HU)_{ii}\), and let \(\Gamma\) be the diagonal matrix with diagonal entries, \(\gamma_1, \ldots, \gamma_n\). Let \(T\) be the positive square root of \(H\). Then
\[
\sum_i (HU)_{ii}^2 = \text{trace}(\Gamma HU) = \text{trace}(\Gamma TTU)
\]
\[
\leq \text{trace}(\Gamma TT) \frac{1}{2} \text{trace}(TUU^*T)^{1/2}
\]
\[
= \text{trace}(\Gamma^2 H) \frac{1}{2} \text{trace}(H)^{1/2}
\]
\[
= \left( \sum \gamma_i^2 \right)^{1/2} \sqrt{n} = \left( \sum_i (HU)_{ii}^2 \right)^{1/2} \sqrt{n}
\]
the inequality in the middle following from the “matrix Cauchy–Schwarz” inequality. This proves the lemma. □

Theorem 13 is proved by combining these lemmas with Bang’s lemma.

**Proof.** Given the matrix $A$, choose $\Theta$ and $U$ as guaranteed by Lemma 6 and put $H = \Theta A U$. Let $\theta_1, \ldots, \theta_n$ be the diagonal entries of $\Theta$. Since $H$ is positive, and has 1’s on the diagonal it is the Gram matrix of a sequence $(u_i)$ of unit vectors in Euclidean space. By Bang’s Lemma there is a sequence of signs $(\varepsilon_i)$ for which the vector

$$x = \sum \varepsilon_i \theta_j u_j$$

satisfies

$$|\langle x, u_i \rangle| \geq \theta_i$$

for each $i$. This is equivalent to

$$\left| \sum_j (H)_{ij} \theta_j \varepsilon_j \right| \geq \theta_i$$

for each $i$. In view of the definition of $H$ this is in turn equivalent to

$$\left| \sum_j \theta_i (AU)_{ij} \theta_j \varepsilon_j \right| \geq \theta_i,$$

and hence to

$$\left| \frac{1}{n} \sum_j (AU)_{ij} \theta_j \varepsilon_j \right| \geq \frac{1}{n}$$

for each $i$.

The sequence given by

$$\lambda_k = \frac{1}{n} \sum_j U_{kj} \theta_j \varepsilon_j$$

will therefore solve the problem, provided that

$$\sum \lambda_k^2 \leq \frac{1}{n}.$$

Now, since $U$ is an orthogonal matrix, this is the same as asking that

$$\sum \theta_j^2 \leq n.$$
But $\Theta A = HU^*$ and since $A$ has 1's on its diagonal this implies, in particular, that for each $i$,

$$\theta_i = (HU)_{ii}.$$  

The desired estimate now follows from Lemma 3.2.

This solves the finite-dimensional version of the plank problem. The infinite-dimensional version does not follow formally from the finite-dimensional, except in reflexive spaces where there is weak-compactness. However, the “quadratic” estimate obtained in the special case in which all the plank widths are equal, can be transferred back to the more general situation: the result that has really been proved is this.

**Theorem 14.** Let $A = (a_{ij})$ be an $n \times n$ real matrix with 1's on the diagonal and $(w_i)$ a sequence of positive numbers. Then there is a sequence $(\lambda_j)$ for which

$$\sum_j w_j^{-1} \lambda_j^2 \leq \sum_j w_j,$$  

but

$$\left| \sum_j a_{ij} \lambda_j \right| \geq w_i$$

for each $i$.

This quadratic control is sufficient to handle the infinite-dimensional situation. Given

$$\sum_i w_i < 1,$$

use the theorem, for each $n$, to construct a sequence $(\lambda_j)^n$, which corresponds to a point where the first $n$ functionals are large. The quadratic control on these sequences ensures that they are uniformly summable; so it is possible to find an accumulation point.

### 3.3. Nazarov’s solution to the coefficient problem

The purpose of this section is to prove Theorem 12, stated earlier. Nazarov’s argument begins with slightly stronger hypotheses than those above. It initially gives:

**Theorem 15.** Let $(\psi_j)$ be a sequence of unit vectors in $L_1$ of a probability space, which satisfy an upper-2-estimate in $L_2$: i.e., for any sequence of coefficients $(a_j)$,

$$\left\| \sum_j a_j \psi_j \right\|_2 \leq M \left( \sum_j a_j^2 \right)^{1/2}.$$
Then for any sequence \((c_j)\) satisfying
\[ \sum c_j^2 < \infty \]
there is a bounded function \(f\) satisfying
\[ |Ef\psi_j| \geq c_k \]
for each \(j\).

Theorem 12 reduces to Theorem 15 via the following weak form of Grothendieck’s inequality (see [26, Section 10]).

**Theorem 16.** If \((\psi_k)\) is a sequence of unit vectors in \(L_1\) (of a measure space \((\Omega, \mathcal{F}, \mu)\)) which satisfy an upper-2-estimate (in \(L_1\)) then there is a non-negative function \(\phi\) for which
\[ \int \phi \, d\mu = 1 \]
and for which the sequence \((\psi_k/\phi)\) satisfies an upper-2-estimate in \(L_2\) of the probability space whose measure is \(\phi \, d\mu\).

The details of the reduction are left to the reader. The proof of Theorem 15 follows.

**Proof.** Consider the functions of the form
\[ f_\epsilon = \sum_{j=1}^{\infty} \epsilon_j c_j \psi_j \]
for all possible choices of sign \(\epsilon = (\epsilon_j)\). Each of these sums converges in \(L_2\) and hence in \(L_1\). Moreover, the map
\[ \epsilon \mapsto f_\epsilon \]
is continuous from the compact space \((-1, 1)^\omega\) into \(L_1\).

The functions \(f_\epsilon\) certainly need not be bounded: but the functions \(g_\epsilon\) given by
\[ g_\epsilon = \tan^{-1} f_\epsilon \]
are bounded. The aim will be to show that for an appropriate choice of \(\epsilon\), and every \(k\),
\[ |\langle g_\epsilon, \psi_k \rangle| \geq \alpha c_k, \]
(\(\alpha\) being a constant depending only upon \(M\)).
Define \( s : \mathbb{R} \rightarrow \mathbb{R} \) by
\[
 s(x) = \int_{0}^{x} \tan^{-1} t \, dt.
\]

\( s \) is Lipschitz, with constant \( \pi/2 \). This ensures continuity of the function \( S : L^{1} \rightarrow \mathbb{R} \) given by
\[
 S(f) = E s(f).
\]

There is, therefore, a choice of signs \( \varepsilon = (\varepsilon_{j}) \) for which \( S(f_{\varepsilon}) \) is maximum. By altering the signs of the \( \psi_{j} \), it may be assumed that the maximum occurs for
\[
 \varepsilon = (1, 1, 1, \ldots).
\]

Use \( f \) to denote the function \( f_{(1,1,1,\ldots)} = \sum c_{j} \psi_{j} \) and \( g \) to denote \( \tan^{-1} f \). For a fixed integer \( k \), take
\[
 f_{k} = \sum_{j \neq k} c_{j} \psi_{j} - c_{k} \psi_{k} = f - 2c_{k} \psi_{k}.
\]

Then, by the maximality of the choice of signs,
\[
 0 \leq S(f) - S(f_{k}) = E (s(f) - s(f_{k})). \tag{10}
\]

Apply the second mean value theorem to the function \( s \) (at each point of the probability space) to find a function \( h \), so that at each point
\[
 s(f_{k}) = s(f) + s'(f)(f_{k} - f) + \frac{1}{2} s''(h)(f_{k} - f)^{2}
 = s(f) - 2s'(f)c_{k} \psi_{k} + 2s''(h)c_{k}^{2} \psi_{k}^{2}
\]
and \( h \) lies between \( f \) and \( f_{k} \). Combine this with Equation (10) to get
\[
 0 \leq E (2s'(f)c_{k} \psi_{k} - 2s''(h)c_{k}^{2} \psi_{k}^{2}).
\]

This implies that
\[
 E g \psi_{k} = E s'(f) \psi_{k} \geq c_{k} E s''(h) \psi_{k}^{2}.
\]

To complete the proof it remains to show that
\[
 E s''(h) \psi_{k}^{2}
\]
cannot be too small. By virtue of the fact that \( s' = \tan^{-1} \) this integral is

\[
E \frac{1}{1 + h^2} \psi_k^2.
\]

By the Cauchy–Schwarz inequality

\[
1 = E|\psi_k| \leq \left( E \frac{\psi_k^2}{1 + h^2} \right)^{1/2} (E(1 + h^2))^{1/2}.
\]

So it is enough to check that \( E(1 + h^2) \) cannot be too large. But since \( h \) is pointwise between \( f \) and \( f_k \), the latter is no more than

\[
E(1 + f^2 + (f_k)^2) \leq 1 + 2M^2.
\]

Nazarov’s article [34] includes several extensions of the above result, to spaces other than \( L_\infty \).

### 3.4. Related results and open problems

An old problem on the boundary between convex geometry and number theory is that of estimating the maximum density of lattice packings of spheres, in Euclidean space. The densest packings found to date, were obtained by this author in [5], using Bang’s Lemma.

A delightful analogue of Nazarov’s Theorem, in a non-commutative setting, was found by Lust-Piquard [31]. She proves that if \( (a_{ij}) \) is a matrix of reals, for which there is a uniform bound upon the sums of the squares of the entries in any row or column, then there is a bounded operator \( B \) on Hilbert space, whose matrix \( (b_{ij}) \) satisfies,

\[
|b_{ij}| > a_{ij}
\]

for all \( i \) and \( j \).

Nazarov’s Theorem and Bang’s Lemma show, respectively, that if \( X \) is either \( L_1 \) or \( L_2 \), and \( (x_i) \) is a sequence of unit vectors in \( X \), satisfying an upper-2-estimate, then for any square-summable sequence \( (w_i) \), there is a functional \( \phi \) in \( X^* \) for which

\[
|\phi(x_i)| > w_i
\]

for every \( i \). The obvious question is “For which \( X \) is this true?” A natural guess would be “Spaces of cotype 2”. The adventurous might like to conjecture that a still more general statement is true; a statement that also includes a version of Theorem 11 (up to some constant).
References


