Surprising Geometric Phenomena in High-Dimensional Convexity Theory

V. Milman

Sackler Faculty of Exact Sciences
Tel Aviv University
Tel Aviv 69978, Israel

ABSTRACT. We shall describe asymptotic phenomena for high-dimensional objects, which violate intuition based on traditional studies of objects of low dimension. We will illustrate these phenomena with examples which involve convex bodies and homogeneous structures (both continuous and discrete) for which other parameters play the role of dimension. Our aim will be to convey some of the new intuition about high-dimensional structures, that underlies such counterintuitive results. We will also discuss the notion of “isomorphic” geometry as the study of common geometric features of families of, say, convex bodies, and analyze their unexpected asymptotic behavior as the dimension increases to infinity. The underlying methods use different mathematical tools and are useful in a variety of apparently unrelated mathematical areas.

1. Introduction

In this lecture we discuss results which stand between Geometry and Functional Analysis. The theory was built during the last two decades. It considers geometric problems via a Functional Analysis point of view. Consequently, the “isometric” problems and point of view typical for geometry are substituted by “isomorphic” ones. This became possible with the asymptotic approach (with respect to dimension increasing to infinity) to the study of high-dimensional spaces. The goal of this lecture is to demonstrate, with some examples and results, a new intuition which corresponds to high dimensional spaces and present some discoveries and highlights of this theory. We avoid the technical part of the theory and will emphasize the geometric insight. In fact, we restrict this presentation to one facet of the theory: isomorphic (asymptotic) geometry and related phenomena. We do not discuss (for example) the very important type-cotype theory with the deep probabilistic connections behind it. To fill these and other omissions I recommend the following books: [MSch86], [P86a], [P86b], [P89a], [TJ88] and surveys [M88a], [M92], [L92], [LM93].

Consider a finite dimensional normed space $X = (\mathbb{R}^n, \| \cdot \|)$. Such a space is defined by its unit ball $K_X = \{ x \in \mathbb{R}^n, \| x \| \leq 1 \}$. Inversely, if $K$ is a convex...
centrally-symmetric compact body in \( \mathbb{R}^n \), then \( X_K = (\mathbb{R}^n, \| \cdot \|_K) \) is the normed space with the unit ball \( K \). Let \( |x| \) be the canonical euclidean norm in \( \mathbb{R}^n \), \( (x,y) \) the standard inner product, and denote \( D \) the standard euclidean ball, i.e. \( D = K(\mathbb{R}^n, |\cdot|) \).

### A few preliminary examples.

We start with three observations:

a) ([GrM84], [MP89]) Fix \( 0 < \delta < \frac{1}{2} \) (say, \( \delta = \frac{1}{4} \)). Define the floating body \( K_\delta \) of \( K \) as the intersection of the half spaces \( \{ x \in \mathbb{R}^n \mid (x, \theta) \leq m_\delta(\theta) \}, \theta \in \mathbb{R}^n \), where \( m_\delta(\theta) \) is defined by

\[
\text{Vol} \left\{ x \in K \mid (x, \theta) > m_\delta(\theta) \right\} = \delta \text{Vol} K.
\]

Then, there is a number \( C(\delta) \), independent of dimension \( n \) or \( K \subset \mathbb{R}^n \), such that for any symmetric convex body \( K \) the floating body \( K_\delta \) is uniformly, up to a factor \( C(\delta) \), isomorphic to an ellipsoid; this means that there is an ellipsoid \( E \), s.t.

\[
E \subset K_\delta \subset C(\delta)E.
\]

[Note that the initial body \( K \) could be very far from any ellipsoid; it could be, say, a cross-polytope (= the unit ball of \( \ell_1^n \)) or a cube, but the “regularization” described above, which is obtained by cutting a fixed portion of volume in any direction, brings us to a \( C(\delta) \)-neighborhood of an ellipsoid.]

Moreover, this ellipsoid is homothetic to the Legendre ellipsoid of inertia \( \mathcal{L}(K) \) of \( K \).

Let us formally introduce a multiplicative geometric distance \( d(\cdot, \cdot) \) between convex bodies \( K \) and \( T \): \( d(K, T) = \inf \{ b/a \mid K \subset bT \text{ and } aT \subset K \} \). The Banach-Mazur distance between two normed spaces \( X_K \) and \( X_T \) is \( d(X_K, X_T) = \inf \{ d(K, uT) \mid u \in GL_n \} \).

b) Centroid body (of Petty [Pet61]) (or “zonoid of inertia” by [MP89]). Consider a body \( Z(K) = -n \int_0^1 K(x) dx \) which is a “Minkowski” integral, meaning that the finite sums approximating the integral are Minkowski sums of sets: intervals \([0, x_i]\) with suitable weights. Geometrically, \( \partial Z(K) \) is the collection of centroid points (centers of gravity) of “halves” of our body \( K \) obtained by cutting \( K \) by hyperplanes passing through the origin.

Again, as in a), there is a universal constant \( C > 0 \), such that \( \forall n, \forall K \subset \mathbb{R}^n \), \( d(Z(K), \mathcal{L}(K)) \leq C \) (and \( C \leq \sqrt{8} \)).

c) Lattice tiling. Let the inner part \( \overset{\circ}{K} \neq \emptyset \). Denote \( K_i = K + x_i \) for \( x_i \in \mathbb{R}^n \). We call \( \{ K_i \} \) a tiling of \( \mathbb{R}^n \) (and say that \( K \) produces a tiling by shifts) if (i) \( \bigcup_i K_i = \mathbb{R}^n \) and (ii) \( \overset{\circ}{K}_i \cap \overset{\circ}{K}_j = \emptyset \) for \( i \neq j \). A lattice tiling is a tiling such that the set \( \{ x_i \} \) is a lattice, i.e. \( \{ x_i \} = AZ^n \), where \( A \) is an invertible linear map and \( Z^n \subset \mathbb{R}^n \) is the set of all integer vectors of \( \mathbb{R}^n \).

Trivially, an affine image of the cube \([-1,1]^n\) produces a lattice tiling and the euclidean ball does not. However, do uniformly isomorphic versions \( K_n \) of the euclidean balls \( D_n \) exist which produce a tiling? Surprisingly, the answer is “Yes”. More precisely, for any integer \( n \) there is a convex symmetric body \( K_n \subset \mathbb{R}^n \), \( D_n \subset K_n \subset 3 \cdot D_n \), such that \( K_n \) produces a lattice tiling of \( \mathbb{R}^n \). (This observation...
of Alon-Milman follows immediately from some known results of [R50], [Bu72], [Bou87], [Ban90], on lattice covering-packing; see [M92] for details and references.)

**Isomorphic Geometry.** The three examples above lead us to a notion of “isomorphic ellipsoid”: a family of convex bodies \( \{K_\alpha\} \) of infinitely increasing dimension represents an “isomorphic ellipsoid” if there is a constant \( C \) and a family of ellipsoids \( \{E_\alpha\} \) such that \( E_\alpha \subset K_\alpha \subset CE_\alpha \) for every \( K_\alpha \) in the family, i.e. \( d(K_\alpha, E_\alpha) \leq C \).

So, in example a) we can state that for a fixed \( \delta, 0 < \delta < \frac{1}{2} \), the family of \( \delta \)-floating bodies \( \{K_\delta \mid \forall n \forall K \subset \mathbb{R}^n\} \) is an isomorphic ellipsoid. (We may add: \( C(\delta) \)-isomorphic ellipsoid).

So, we naturally derive isomorphic geometric objects and isomorphic geometric results. Of course, an “isomorphic” geometric object is, in fact, a family of objects in different spaces of increasing dimension and by “isomorphic” geometric properties of such an “isomorphic” object we mean a common property of this family. Classical Geometry (in a fixed dimension) is an isometric theory (or, at most, “almost” isometric, \( \varepsilon \)-isometric for small \( \varepsilon > 0 \)). “Isomorphic” study means asymptotic behavior determined by some parameter (most often it is dimension \( n \to \infty \)) and exact control of dependence of the constants involved on this parameter (say, dimension in most examples in this lecture).

The appearance of such an isomorphic geometric object is a new feature of asymptotic high-dimensional theory. Geometry and Analysis meet here in a non-trivial way. We will meet later geometric inequalities in isomorphic form. The study of special examples of such inequalities, “isomorphic isoperimetric inequalities” led to the discovery of the “Concentration Phenomenon” – one of the most powerful tools of asymptotic theory, responsible for many counter-intuitive results in the theory.

### 2. Another example – theorem; correction of intuition

Our next example is already a non-trivial theorem originally proved in 1988 (see [M91]).

**THEOREM 2.1.** There is a universal constant \( C \) so that for any \( n \) and any centrally-symmetric, compact, convex body \( K \subset \mathbb{R}^n \), there are two linear operators \( u_1, u_2 \in SL_n \) so that if \( P = K \cap u_1 K \) and \( Q = \text{conv}(P \cup u_2 P) \), then the body \( Q \) is within distance \( C \) of an ellipsoid: that is, there is an ellipsoid \( E \) for which \( E \subset Q \subset CE \).

In the language of Functional Analysis, the same fact can be reformulated in the following form:

**THEOREM 2.2.** For every finite dimensional normed space \( X = (\mathbb{R}^n, \| \cdot \|) \) there are three linear operators \( T_1, T_2, T_3 \subset GL_n \), such that the following relation holds:

Consider \( p(x) = \| T_1 x \| + \| T_2 x \| \) and an inf-convolution

\[
q(x) = p(x) \ast p(T_3 x) \quad [\text{by definition } p(x) \ast p(u x) = \inf_{y+z=x} \{ p(y) + p(u z) \}] ;
\]

then \( q(x) \) is \( C \)-isomorphic to the standard euclidean norm in \( \mathbb{R}^n \):

\[
|x| \leq q(x) \leq C|x|.
\]
Note that $C$ does not depend on the dimension $n$ or on the initial norm $\| \cdot \|$.

In fact, an even stronger statement is correct:

**THEOREM 2.3.** For any four centrally-symmetric convex bodies $K_i \subset \mathbb{R}^n$, $i = 1, 2, 3, 4$, $|K_i| = |\mathcal{D}|$, there are $\{u_i\}_{i=1}^4 \subset SL_n$ such that if $P_1 = u_1K_1 \cap u_2K_2$, $P_2 = u_3K_3 \cap u_4K_4$ and $Q = \text{Conv } P_1 \cup P_2$, then $\frac{1}{C}\mathcal{D} \subset Q \subset C\mathcal{D}$ for some universal constant $C$ independent of the dimension $n$ and $\{K_i\}_{i=1}^4$.

There are many remarkable ellipsoids associated with a given body $K$. They recover different traces of hidden symmetries which exist in any high-dimensional convex body. If $K$ has a “large” group of symmetries then all these ellipsoids coincide. Well-known, for example, are the maximal volume ellipsoid inscribed in $K$ (introduced by Löwner and studied by F. John), the minimal volume ellipsoid circumscribed around $K$ or the Legendre and Binet ellipsoids of inertia. There are many other important constructions of ellipsoids associated with $K$ and an ellipsoid which Theorem 2.1 provides is one of them. We call it the $M$-ellipsoid and know how to identify it a priori through some geometric inequality. Changing coordinates we may consider this ellipsoid $\mathcal{E}$ to be the standard euclidean ball $\mathcal{D} \subset \mathbb{R}^n$; let $w \in GL_n$ be such that $w\mathcal{E} = \mathcal{D}$. We call the family of position of $K$ the family $\{vK\}_{v \in GL_n}$. So the affine image $wK = \hat{K}$ is a position of $K$. Applying Theorem 2.1 to $K$ in such a position, i.e. to $\hat{K}$, we may additionally claim that $(u_1; u_2)$ are orthogonal pairs of operators and “most” (meaning with high probability with respect to the Haar probability measure of orthogonal groups) pairs $(u_1; u_2) \in O(n) \times O(n)$ will regularize $\hat{K}$ to an isomorphic euclidean ball $\mathcal{D}$.

Fig. 1
Clearly, every position of $K$ produces the unit ball of isometrically the same normed space as $X_K$. It is an interesting feature of an (asymptotic) high dimensional theory of convex sets which we will meet throughout the lecture that we are, in fact, forced to consider the family of all positions of a given $K$ (that is all affine images $uK$, $u \in \text{GL}_n$) even when we are aiming at some volume inequalities or other properties of an individual $K$. (See K. Ball [B91] for the use of the family of positions in isoperimetric type problems.)

Let us try to develop an intuition of high-dimensional spaces. To start with, we should understand how to draw “high-dimensional” pictures, how high-dimensional convex bodies “look”.

The first non-intuitive point is that 2 or 3 dimensional “pictures” of a high-dimensional convex body should have a “hyperbolic” form! The reason is that the volume of parallel hypersections decays exponentially after passing the median level (see Fig.1).

This is a fact observed in [GrM84]. Yes, $K$ is a convex set but the rate of volume decay has a crucial influence on the geometry and we should find a way to visualize it in our pictures.

Now, with such a picture in mind, let us try, intuitively, to understand the theorem.

We use the following notation: $|K|$ defines the volume of $K$ and $K^o = \{x \in \mathbb{R}^n \mid (x,y) \leq 1 \text{ for every } y \in K\}$ is the polar body of $K$.

Let $|K| = |D|$; for some $R > 1$ almost no $(n-1)$-dimensional volume belongs to the intersection $R \cdot \partial D \cap K$. Then find a turn $u_1$ s.t. (see Fig. 2) $$ (R\partial D \cap K) \cap (R\partial D \cap u_1 K) = \emptyset . $$

So $P = K \cap u_1 K \subset R \cdot D$. 

![Fig. 2](image-url)
But $K$ is convex (finally, we recall this fact), so consider $P^o \supset \frac{1}{R}D$. Do the same operation with the body $P^o$: $Q^o = P^o \cap u_2P^o \subset RD$ and it looks as though we succeeded:

$$\frac{1}{R}D \subset Q \subset R \cdot D.$$ 

However: When does exponential decay start? What is the value of $R$? For many classical ellipsoids (say, maximal vol. ellipsoids), $R$ may be as large as $\sim \sqrt{n/\log n}$. For some other ellipsoids $R \sim \log n$. It was necessary to construct a special ellipsoid (better to say, to prove its existence) for which $R \sim \text{Const}$. However, this is not yet enough. The result above is a "metric one": we want $(K \cap u_1K \cap \partial(R\cdot D))$ to be empty, not just to have small volume; so more corrections are necessary, but the idea is correct. Just that instead of estimating decay of measure, we should consider decay of coverings.

3. Isomorphic geometric inequalities

Let us return to the ellipsoid which appeared in Theorem 2.1. It was originally constructed in connection with some geometric inequality (but, again, inequality in isomorphic form).

**Theorem 3.1.** There exists a universal constant $C$, such that for any convex symmetric body $K$ in $\mathbb{R}^n$, there is an ellipsoid $M_K$, $|M_K| = |K|$, such that for any (not only convex) set $T$

(i) $\frac{1}{C}|M_K + T|^{1/n} \leq |K + T|^{1/n} \leq C|M_K + T|^{1/n}$

where $K + T = \{x + y \mid x \in K, y \in T\}$ is the Minkowski addition of sets. (We ignore measurability problems; they are not crucial here). And for any convex symmetric set $T$

(ii) $\frac{1}{C}|M_K \cap T|^{1/n} \leq |K \cap T|^{1/n} \leq C|M_K \cap T|^{1/n}$.

Define volume radius of $K$ by $\text{vrad} K = (|K|/|D|)^{1/n}$. Then Theorem 3.1 states that substituting $K$ by some special ellipsoid $M_K$ in (i) and (ii) essentially preserves volume radii.

Of course, it immediately implies reversion of the Brunn-Minkowski inequality (which was the initial result in this direction). Recall that the Brunn-Minkowski inequality states that for any two compact sets (say, convex) $K$ and $T$ in $\mathbb{R}^n$,

$$|K + T|^{1/n} \geq |K|^{1/n} + |T|^{1/n}.$$

The reverse Brunn-Minkowski inequality is the following statement:

**Theorem 3.2 [M86].** $\exists C$ such that for any integer $n$ and any two convex bodies $K$ and $T$ there is a relative position $\hat{K} = uK$, $u \in SL_n$, of say, $K$ such that

$$|\hat{K} + T|^{1/n} \leq C(|K|^{1/n} + |T|^{1/n}).$$

(Also the same inequality holds for the polars $\hat{K}^o, T^o$.)
This is a very easy consequence of Theorem 3.1(i) (but it precedes it).

The original proof of Theorem 3.2 was followed by three other proofs: two by G. Pisier ([P89a], [P89b]) and another one of [M88b]. All four proofs being different in their technical presentations added a number of facets to the understanding of the phenomenon behind Theorems 3.1 and 3.2. One such addition is the following covering property.

Let \( N(K,T) = \min\{N \mid \exists \{x_i\}_1^N \subset \mathbb{R}^n, K \subset \bigcup_1^N (x_i + T)\} \).

An equivalent statement for Theorem 3.2 (or Theorem 3.1) is the following fact: There is a number \( C \) such that for any \( n \in \mathbb{N} \), any \( K \) (convex symmetric) there is an ellipsoid \( M_K \), \( |K| = |M_K| \), such that

\[
N(K, M_K) \leq e^{Cn} \quad \text{and} \quad N(M_K, K) \leq e^{Cn}.
\]

This is, of course, the best possible estimate (up to a value of the universal constant \( C \)). Note that, for any ellipsoid \( E \), \( N(2E, E) \sim C_0^n \) (for some \( C_0 > 1 \)) but \( K \) may be very far from any ellipsoid.

Of course, the ellipsoid described above is not unique and there are many with different additional properties.

Such additional important information is a decay of the functions \( N(K, tM_K) \) and \( N(M_K, tK) \) when \( t \) increases. Pisier [P89a] showed that, for any \( p < 2 \) there is an \( M \)-ellipsoid \( E \) such that \( N(E, tK) + N(K, tE) \leq e^{C_p n/t^p} \) for some constant \( C_p \) depending on \( p < 2 \) only.

Remarkably, in fact, ellipsoids do not play a special role in these facts. We may take any convex symmetric body \( P \) and the family of its positions \( \mathcal{P} = \{uP\}_{u \in GL_n} \). Then the same facts are true for \( \mathcal{P} \) instead of the family of ellipsoids. Say:

\[
\exists C \text{ such that for any convex body } K, \text{ there is a position } P_K \in \mathcal{P} \text{ such that } |K| = |P_K| \text{ and } N(P_K, K) + N(K, P_K) \leq e^{Cn}.
\]

Also \( P_K \) can be put in inequalities (i) and (ii) of Theorem 3.1 instead of \( M_K \).

So, we have some (high-dimensional) Principle:

The affine family of any fixed convex symmetric body (say \( P \)) is rich enough to substitute any other body (\( K \) above) in an essential part of volume computations.

This reflects a new intuition about high dimension: instead of the expected increasing of essentially different bodies in \( \mathbb{R}^n \) with increasing \( n \rightarrow \infty \) (diversity of possibilities) we observe their decreasing to essentially one (any) body with its affine class.

In many of the results we have described till now, convexity is not a crucial assumption (but high dimension is). One may substitute convexity by another, very weak condition, and Theorems 2.1, 2.2 and 3.1, 3.2 still hold. Define a convolution of two bodies \( K \triangleq T = \bigcup_{x \in K} [x, y] \), were \([x, y]\) is the interval joining points \( x \) and \( y \).

Let \( K \) be a centrally-symmetric quasi-convex star body, i.e.

(i) \( tK \subset K, \ 0 \leq t \leq 1 \) (star-body condition); and \( K = -K \)

(ii) \( K \cap K \subset C \cdot K \)

(We say that \( K \) is \( C \)-quasi-convex. Example: \( K_{\ell_p^n} \), the unit ball of \( \ell_p^n \) space for \( 0 < p < 1 \), is a \( C(p) \)-quasi-convex for some constant \( C(p) \) independent of dimension \( n \).) Clearly, \( C = 1 \) iff \( K \) is convex.
Then Theorems 3.1 and 3.2 are true for quasi-convex bodies with constant $C$ in these theorems depending on the quasi-convexity constant only ([BBP95]). Using this fact, Theorems 2.1 and 2.2 are also extended to a quasi-convex setting ([M95]):

Let $f \in C(S^{n-1})$, $f > 0$ and even. Consider its homogeneous extension to $\mathbb{R}^n$ by $\tilde{f}(x) = |x|f(\frac{x}{|x|})$, for $x \neq 0$. Let $K = \{x \in \mathbb{R}^n \mid \tilde{f}(x) \leq 1\}$ be a $C$-quasi-convex body. Apply Theorem 2.2 to the function $\tilde{f}(x) \equiv \|x\|$ (it is not a norm now but a quasi-norm). Then the function $q(x)$ we obtained is $C_1$-isomorphic to the standard euclidean norm in $\mathbb{R}^n$ where $C_1$ depends only on the $C$-quasi-convexity constant of $K$.

As we will see below, these facts reflect the probabilistic nature of high dimension. And I mean by this more than just the fact that we are using probabilistic techniques in many steps of the proofs. In the next section we will discuss the so-called *concentration phenomenon* but here I would like to mention one very recent result which gives a probabilistic description for the norm of an operator

$$A : \ell^n_2 \to X.$$  

Of course, the norm $\|A\| = \sup\{|Ax| \mid |x| \leq 1\}$ is not a probabilistic notion. However, up to a universal factor, i.e. in isomorphic form, it is (!) (see [MSch96]): there is a universal constant $C$ such that

$$\frac{1}{C} \sqrt{n/kM(A)} \leq \|A\| \leq C \sqrt{n/kM(A)}$$

where $M(A) = \int_{S^{n-1}} \|Ax\|d\mu(x)$, $\mu(x)$ is the rotation invariant probability measure on the sphere $S^{n-1} = \{x \mid |x| = 1\}$, and $k = k(A)$ has also a probability meaning:

$$k = k(A) = \max \left\{ k \in \mathbb{N} \mid \text{Prob}\{E \in G_{n,k} \mid \|Ax\| \leq 2M(A)|x| \right. \text{for } x \in E \geq 1 - \frac{k}{n+k} \right\}.$$  

Here $G_{n,k}$ is the Grassmann manifold of $k$-dimensional subspaces of $\mathbb{R}^n$ and the probability measure on $G_{n,k}$ is defined by the euclidean structure $\ell^n_2$ on $\mathbb{R}^n$.

4. Isoperimetric inequalities in isomorphic form; concentration phenomenon

One of the major tools in the development of asymptotic high dimensional theory are isomorphic forms of isoperimetric inequalities. Let us introduce a general framework for such problems and then let us describe a few examples to feel what we mean by *isomorphic* form of inequalities.

Let $(X, \rho, \mu)$ be a metric space ($\rho$ is a metric) with a probability measure $\mu$. For $A \subset X$, let $A_\varepsilon = \{x \in X \mid \rho(x, A) \leq \varepsilon\}$. Define

$$\alpha(X; \varepsilon) = \sup \left\{ \mu(X \setminus A_\varepsilon) \mid \forall A, \mu(A) \geq \frac{1}{3} \right\}.$$
So, the value $\alpha(\varepsilon)$ corresponds to a solution of the “$\varepsilon$-isoperimetric problem” for sets of $\frac{1}{2}$-measure.

Only in very few cases, can solutions be described explicitly and $\alpha(X, \varepsilon)$ calculated exactly. The case of a euclidean sphere $S^n$ — solved by P. Lévy [L51] in 1919 is such a lucky case and the most classical one after $M$. The calculation gives an estimate $\alpha(S^{n+1}, \varepsilon) \leq \sqrt{\frac{\pi}{8}} e^{-\varepsilon^2 n/2}$. We consider here $S^n \subset \mathbb{R}^{n+1}$ equipped with geodesic distance $\rho_r$ and rotation invariant probability measure $\mu$.

This means that for (any) $A_n \subset S^{n+1}$, $\mu(A_n) = \frac{1}{2}$, and $\varepsilon > 0$, $\mu((A_n)_\varepsilon)$ tends to 1 when $n \to \infty$ (and exponentially quickly). This is not a geometric point of view. Classical geometry fixes a space (meaning “n”) and considers the dependence on $\varepsilon > 0$ and then it does not observe the phenomenon! However, we do not consider one set “A” but a family of sets $\{A_n \subset S^{n+1}, \mu(A_n) \geq \frac{1}{2}\}$ and a statement which is an isomorphic geometric inequality is a uniform bound for $\varepsilon$-extensions of sets of this family.

At the same time in many (most) interesting cases, we know how to estimate $\alpha(X; \varepsilon)$ without any idea of the structure of the extremal sets. The above estimate for $S^n$ happens to be typical and implies the so-called “concentration phenomenon”. Typical estimates for many natural families $\{X_n\}$ have a similar form:

$$\alpha(X_n, \varepsilon) \leq c_1 e^{-c_2 \varepsilon^2 n}.$$  

However, the exact structure of extremal sets or exact value of $\alpha(X; \varepsilon)$ are not known even for such “simple” cases as the torus $X = \mathbb{T}^n = \prod_{i=1}^n S^1_i$ equipped with Haar probability measure and distance $d(x, y) = \frac{1}{\sqrt{n}} \sqrt{\sum_{i=1}^n \rho(x_i, y_i)^2}$ where $\rho$ is the geodesic distance on $S^1$ and $x = (x_i)_{i=1}^n$, $y = (y_i)_{i=1}^n$, or the Stiefel manifold $W_{n,2}$ — the unit tangent bundle of $S^{n-1}$. (Both these examples satisfy $\ast$.)

Another important geometric example (where no extremal sets are known, as well as the exact value of $\alpha(\varepsilon)$) ([GrM83]):

$$\alpha(SO_n, \varepsilon) \leq \sqrt{\frac{\pi}{8}} e^{-\varepsilon^2 n/8}.$$  

Such geometric estimates, asymptotic by increasing “dimension” $n$, form one of the major technical tools of Asymptotic Theory, and the isomorphic form of the inequalities is enough for all applications [and also the only one typically available]. So the isomorphic (=asymptotic) view on isoperimetric problems has freed us from the necessity of solving (exactly) the isoperimetric problems.

[It is an extremely well-developed direction which uses different methods and with many people involved; besides those mentioned above there are Alon, Amir, Bobkov, Borell, Ledoux, Marton, Maurey, Schechtman, Sudakov, Talagrand, Tsirelson and others. (See the books [MSch86], [LT91] and surveys [M88a] and [T95] for details.)]

Let us outline why a bound of the form $\ast$ is so crucial. Consider a 1-Lip function $f(x)$ defined on $(X, \rho, \mu)$, i.e. $|f(x) - f(y)| \leq \rho(x, y)$.

Denote by $L_f$ the median of $f(x)$, i.e. $\mu\{x \in X \mid f(x) \geq L_f\} \geq \frac{1}{2}$ and $\mu\{x \in X \mid f(x) \leq L_f\} \geq \frac{1}{2}$. Then

$$\mu\{x \in X \mid |f(x) - L_f| < \varepsilon\} \geq 1 - 2\alpha(X, \varepsilon).$$
So, if the value of $\alpha(X, \varepsilon)$ is very small, then values of the Lipschitz function “concentrate” by a measure around one value. This is the case when $X = S^n$ and dimension $n$ is large, as well as for $T^n$ or $SO_n$ for large $n$. It is, in fact, a general property of high-dimensional metric probability spaces which we call the “concentration phenomenon”. Many different techniques were developed to treat different examples of metric probability spaces and to prove concentration phenomenon for them. Let us describe one discrete example considered by B. Maurey [Mau79].

Let $\Pi_n$ be the group of permutations of $\{1, \ldots, n\}$ equipped with the counting probability measure $\mu(A) = \#A/n!$ (for any $A \subset \Pi_n$) and Hamming distance $\rho(s, t) = \#\{i \in [1, \ldots, n] : s(i) \neq t(i)\}$ for any two permutations $s, t \subset \Pi_n$. Then $\alpha(\Pi_n, \varepsilon) \leq c_1 e^{-c_2 \varepsilon^2 n}$ for absolute constants $c_1$ and $c_2 > 0$. Again, we have an inequality of “isoperimetric type” but in a new, isomorphic version, which is enough for what we call the “concentration phenomenon”.

This phenomenon is responsible for many “unexpected”, “strange” properties of high-dimensional spaces. Behind intuitively “expected” properties of high-dimensional spaces stands the behavior of $\varepsilon$-entropy which increases exponentially with increase in dimension. However, this exponential increase is compensated by the exponential effect of the concentration phenomenon. As a result, we often observe only linear behavior where a priori intuition expects an exponential. Some Examples:

Approximation by Minkowski sums. Let $A + B = \{x + y \mid x \in A, y \in B\}$ be the Minkowski sum of two sets $A$ and $B$ in $\mathbb{R}^n$. Let $I_i = [-x_i, x_i] \subset \mathbb{R}^n$ be intervals of length, say, 1. Consider $T = \sum_{i=1}^N I_i$. We want to approximate a Euclidean ball by such sums, that is, for a given $\varepsilon > 0$ we would like to have $d(T, D) \leq 1 + \varepsilon$. Obviously, if $N = n$ then $d(T, D) \geq \sqrt{n}$ and, by an entropy consideration, it looks as if we need at least an exponential by $n$ number of intervals to achieve a good approximation to $D$. However, an easy geometric interpretation of an old result from [FLM77] (see also [Go85], [Sch88]) shows that there exist intervals $I_i \subset \mathbb{R}^n$, $i = 1, \ldots, N_0$, for $N_0 \leq c_3 n^{3/2}$ ($c$ is a numerical constant) such that $d\left(\sum_{i=1}^{N_0} I_i, D\right) \leq 1 + \varepsilon$.

This direction was later treated intensively as part of a program of estimating the number of intervals needed for approximating zonoids by a sum of intervals (in [Sch87], [BLM89a] and [BLM88]). It is shown there that the above situation is essentially preserved when we substitute intervals by other convex bodies or if we consider approximation by sums of other convex bodies instead of $D$. For example [BLM88], [Schm91].

**Theorem 4.1.** Let a convex compact body $K \subset \mathbb{R}^n$ be given. There exist orthogonal operators $A_i \in SO_n$, $i = 1, \ldots, N_0$ for $N_0 < c_4 n^{3/2}$ such that $T = \frac{1}{N_0} \sum_{i=1}^{N_0} A_i K$ satisfies $d(T, D) \leq 1 + \varepsilon$ (as usual, $c$ is a numerical constant).

**Classical symmetrizations.** Surprisingly very few symmetrizations are necessary to transform any convex body $K$ to $\varepsilon$-neighborhood of a euclidean ball. Just $\frac{1}{2} n \log n + c(\varepsilon)n$ Minkowski symmetrizations will do the job [BLM88].

The situation with Steiner symmetrization is less clear: $\exists c > 0$ s.t. starting with $\forall K \subset \mathbb{R}^n$, $|K| = |D|$, we derive $T$, $cD \leq T \leq (1 + \varepsilon)D$, after properly
chosen $c(\epsilon)n\log n$ Steiner symmetrizations (note it is again an isomorphic result) [BLM89b]. However, there is resistance for deriving a $(1 + \epsilon)$-approximation by a small number of symmetrizations. Some partial results for higher dimensional symmetrizations are due to Tsolomitis [Ts96].

Another application of the concentration estimate $(*)$ is the following: again, let $f(x) \in C(X)$ be a Lipschitz function $|f(x_1) - f(x_2)| \leq \rho(x_1, x_2)$. Then $\mathbb{E}|f| \equiv \|f\|_{L_1(X)} \sim L_{\|f\|}$ (median of $|f|$) if $\alpha(X, \epsilon)$ is very small. (So as we have already observed, $f$ is “almost constant” with high probability.) But also

$$\|f\|_{L_p(X)} \sim L_{\|f\|} \sim \|f\|_{L_1(X)},$$

and we see that $L_p$-norms are equivalent for Lip. functions on $X$ with small $\alpha(X; \epsilon)$.

The Brunn-Minkowski inequality implies similar properties for any convex body $K$ ([Bor75], [GrM84], [MP89]):

$$\forall n, \forall K \subset \mathbb{R}^n, |K| = 1 \quad \text{and} \quad 0 < p, q < \infty$$

a) $\|f\|_{L_p(K)} \equiv \left( \int_K |f(x)|^p d\mu \right)^{1/p} \leq c_{p,q}\|f\|_{L_q(K)}$ for $\forall$ linear function $f(x)$

b) For any $\| \cdot \|$ and $0 < p, q < \infty$: $\|\|x\||\|_{L_p(K)} \leq c_{p,q}\|\|x\||\|_{L_q(K)}$

One side of such inequalities are just Hölder inequalities; so another side is “reverse” Hölder inequalities which we also call Khinchine type inequalities because $K = [-\frac{1}{2}, \frac{1}{2}]^n$ in a) gives Khinchine inequalities and in b) gives Kahane’s inequalities.

These cases have a probability interpretation (they correspond to independent random variables uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$). However, from our point of view, the role of the “independence” condition, as well as other standard probability type conditions such as, say, “martingale”, is to create a “high-dimensional” convex body; but any $n$-dimensional convex body will serve the same purpose. The reason behind applications of these conditions is the high dimension created and not, for example, independence.

So, probability type ideas and approach are combined here with high-dimensional convexity theory. In fact, the starting point for P. Lévy’s study of isoperimetric inequality for $S^n$ (and its consequence – the concentration phenomenon for the sphere $S^n$) was a geometric interpretation by E. Borel (~1914) of the law of large numbers.

Let $C^n = [-1, 1]^n$ be a cube in $\mathbb{R}^n$ with the standard euclidean distance “dist”. Then $\text{diam} C^n = 2\sqrt{n}$. Consider a linear functional $f$, $f(x) = \sum_1^n x_i$ (i.e. $\text{Ker} f = (1, \ldots, 1)^\perp$). Then $(1/2^n) \text{Vol}\{x \in C^n : \text{dist}(x, \text{Ker} f) \geq \epsilon \sqrt{n}\} = P\left\{ \left| \frac{1}{n} \sum_1^n x_i \right| > \epsilon \right\}$ are uniformly distributed on $[-1, 1]$ independent random variables] $\leq c \exp \left( -\frac{\epsilon^2 n}{2} \right)$. Therefore, “most” of the volume of $C^n$ is concentrated near a “small slice” (relative to the diameter). However, the concentration phenomenon for a cube $C^n$ implies a similar estimate for any 1-Lip. function on $C^n$ and not only for a particular linear functional. So, in a sense, the concentration phenomenon is a “non-linear” version of deviation-inequality type results in Probability.
Returning to a) and b) equivalences above, Bourgain [Bou91] extended a) to any polynomials (with a constant $c_{p,q}(d)$ depending also on the degree $d$ of the polynomial). An interesting case of a) and b) is $q \to 0$. It was shown in [MP89] that $c_{p,q}$ are independent of $q$ for linear functionals and recently Latala [La96] showed that the constants $c_{p,q}$ are independent of $q$ also in b).

The last example of family of $V$ spaces $X_n$ with a concentration property I would like to demonstrate is also discrete. It is taken from [AIM85] and is useful in Computer Science (see also [Al86]).

Let $G = (V, E)$ be a connected graph on $|V| = n$ vertices. We equip $V$ with the counting probability measure $\mu(A \subset V) = |A|/|V|$ and with the path metric: $\rho(x, y) = \{\text{the smallest number of edges in a path which joins } x \text{ and } y\}$.

Let $L^2(V)$ denote the space of real-valued functions on $V$ with the usual scalar product $(f, g)$ and the usual norm $\|f\| = \sqrt{(f, f)}$ induced by it. Consider the quadratic form
\[
(Qf, f) = \sum_{e \in E} (f(e^+) - f(e^-))^2
\]
defined on $f \in L^2(V)$ where $e^+$ and $e^-$ denote the vertices joined by the edge $e$.

Let $0 = \lambda_0 < \lambda_1 = \lambda_1(G) \leq \lambda_2 \cdots \leq \lambda_{n-1}$ be the eigenvalues of $Q$. Clearly, $Q$ plays a role of Laplacian for the graph $G = (V, E)$. Let $d = \max\{d(v) | v \in V\}$ be the maximum degree of a vertex of the graph $G$. The concentration function of $(V, \rho, \mu)$ is estimated by
\[
\alpha(V; \varepsilon) \leq \frac{1}{2} \exp \left\{ -\varepsilon \sqrt{\lambda_1/2d\log 2} \right\}.
\]
(Note that $\varepsilon$ is an integer in this case and ranges from 1 to the diameter of $V$.)

The most important examples of the use of this estimate are the so-called Cayley graphs. Let $V$ be a finite group and $S \subset V$ some set of generators. Assume $S = S^{-1}$ and the identity $e \notin S$. We join $v$ and $u$ from $V$ by an edge if $u = s^{-1}v$ for some $s \in S$. Then a path distance in such a graph, $G = (V; E)$, is the word distance in $V$ induced by $S$. The degree $d$ of $G$ is equal to $|S|$. Consider $L_2(V)$ and let $\pi$ be the left regular representation of $V : \pi(t)f(v) = f(t^{-1}v)$. So if $\{e_t\}_{t \in V}$ is the natural basis of $L_2$ then $\pi(t)e_v = e_{tv}$ and the Laplacian $Q = |A|I - \sum_{s \in S} \pi(s)$. This is a self-adjoint operator and $\lambda_1(Q) = \{|S| - \text{the second largest positive eigenvalue of } A(G) = \sum_{s \in S} \pi(s)\} \geq |S| - \|A|L_2\|/2$ where $L_2^2 \oplus \{\text{const.}\} = L_2(V)$.

To use this construction successfully we have to find natural families of groups $V_i$, $|V_i| = n_i \to \infty$, and generators $S_i \subset V_i$, $|S_i| = d_i \leq \text{Const}$, such that $\|A(G_i)\|_{L_2^2} \leq |S_i| - \varepsilon$ for some positive fixed $\varepsilon$. This brings us to the $T$-property of Kazhdan [K67] well-known in Representation Theory, which gives us a number of most interesting examples, say $V_m = \text{SL}_k(Z/mZ)$, $k \geq 2$ fixed and $m \to \infty$.

There is a (right) choice of generators $S_m \subset V_m$, $|S_m| = 4$, such that $\alpha(V_m; t) \leq \frac{1}{2} \exp(-ct)$. Note that the bound does not depend on $p$. It is a very strong concentration property and implies families of expanders – special bipartite graphs which play a central role in the construction of fast algorithms (see, e.g. [Al86], [AlSp92]). I also recommend the books [S90] and [Lu94] which show how to apply delicate results from Number Theory to choose the best possible sets of generators $S_m$. Many other fascinating related results can be found in these books.
5. Duality and connected isomorphic inequalities

Let $K$ be a convex compact body in $\mathbb{R}^n$ and $K = -K$. Then the polar body (or the dual body) of $K$ is $K^\circ = \{ x \in \mathbb{R}^n \mid (x, y) \leq 1 \text{ for } \forall y \in K \}$. Obviously $X_K^* = X_{K^\circ}$. It is well-known how important duality is in the study of convexity and also in Functional Analysis. The notion has its roots, in fact, in mechanics. Legendre and Binet ellipsoids of inertia of $K$ are, up to a factor, the dual of each other (see [J37] or [MP89]). Clearly, a larger body has a smaller polar. It is possible to give (and in many different forms) a precise quantitative form to this relation.

5.1. For example, the product of the volume radii of $K$ and $K^\circ$ remains between two universal constants: there is a universal constant $c > 0$ such that $\forall n$ and $\forall K \subset \mathbb{R}^n$ as above

$$c \leq \left( \frac{|K| \cdot |K^\circ|}{|D|^2} \right)^{1/n} \leq 1.$$

(The right side is the classical Blaschke-Santaló inequality, and the left side is [BM85].)

These inequalities imply, perhaps more surprisingly, a similar relation for covering numbers ([KM87]): there are universal constants $C$ and $c > 0$ such that $\forall n$ and for every two convex symmetric bodies $K$ and $T \subset \mathbb{R}^n$

$$c \leq \left( \frac{N(K, T)}{N(T^\circ, K^\circ)} \right)^{1/n} \leq C.$$

5.2. The next two results are an interpretation of facts from Local Theory in the language of classical convexity. However, I do not know of any other proofs than the very delicate and non-trivial analysis of the corresponding normed spaces by purely Functional Analysis methods.

5.2a. [FLM77] Let $P$ be a centrally symmetric polytope, $v(P) = \# \text{ of its vertices and } f(P) = \# \text{ of its faces}$. Then $\log v(P) \cdot \log f(P) \geq cn$ for some universal constant $c > 0$ (and for some polytopes it is exact). (Note the importance of the symmetry condition $P = -P$; otherwise, for simplex $S$, $v(S) = n + 1$ and $f(S) = n + 1$; so exponents are a result of symmetry!)

5.2b. Let $w(K)$ be the mean width of $K$ (again $K = -K$). (It is called $b(K)$ in Convex Geometry; we use the notation $M^*(K)$ in Local Theory; let $||x||^*$ be the norm with $K^\circ$ being the unit norm; then $\frac{1}{2} w(K) = M^*(K) = \int_{S_{n-1}} ||x||^* d\mu(x)$.)

Then $\exists$ a position $uK = \tilde{K}$, $u \in SL_n$, of $K$, s.t.

$$w(\tilde{K}) \cdot w(\tilde{K}^\circ) \leq C \log n$$

(or, more strongly, $\leq c \log d_K$ where $d_K = d(X_K, \ell_2^n) = \min\{d(K, E) \mid E \text{ is an ellipsoid}\}$).
This is a combination of results of Lewis [Lew79]/Figiel-Tomczak [FT79]/Pisier [P82]. Note that it is a different type of isomorphic geometric inequality (not a volume-type) and one of the most important tools in the proofs of all isomorphic volume type inequalities we discussed before. It uses solutions of some variational problems, non-trivial duality arguments and, most important, the estimate of Pisier for Rademacher projections which is a non-trivial application of very concrete Harmonic analysis methods in abstract Functional Analysis.

5.3. There is an even deeper connection between sections of $K$ and $K^\circ$.

5.3a. Fix $\kappa > 0$. Let $k + m = \lfloor (1 - \kappa)n \rfloor$. For "most" [=high probability] sections $E$ of dimension $k$ and most sections $F$ of $K^\circ$ of dimension $m$,

$$w(K \cap E) \cdot w(K^\circ \cap F) \leq \frac{8}{\kappa}.$$  

Also, in fact, for some universal $C$

$$\int_{G_{n,k}} w(K \cap E) d\mu(E) \cdot \int_{G_{n,m}} w(K^\circ \cap F) d\mu(F) \leq \frac{C}{\kappa}.$$  

This is surprising: polarity is a global construction, so why does it connect so precisely with local structure, i.e. with the structure of sections?

5.3b. In fact, even the maximal width $r(K \cap E) = \text{diam}(K \cap E)$ of a central section $E$ of $K$ of dimension, say, $n/2$, can be estimated in average by the mean-width: for some universal constant $C$,

$$\int_{G_{n,n/2}} \text{diam}(K \cap E) d\mu(E) \leq C w(K).$$  

In Local Theory, we call this “the low $M^*$-estimate” (see [M85], [M90], [PT86], [Go88], or any of the books mentioned earlier). In its analytic form it is used as a bridge between results of a Functional Analytic nature and Geometric type Global results.

5.4. Another type of geometric (isomorphic) inequality connecting a body $K$ with its polar $K^\circ$ is the following:

Let $r_\ell = \max\{r \mid r\mathcal{D} \subset \frac{1}{\ell} \sum_{i=1}^\ell u_i K, \ u_i \in O(n)\}$. Then there exists $c > 0$ such that, for all $n$, all $K \subset \mathbb{R}^n$, convex symmetric compact body

$$r_2(K) \cdot r_3(K^\circ) \geq c.$$  

In fact, in this case we know that $c \geq 0.0388$ (see [M91]).

Applying this fact to, say, the unit cubes $C^n = [-1, 1]^n$, we immediately see that $\exists u \in O(n)$ s.t.

$$c\sqrt{nD} \subset \frac{C^n + uC^n}{2} \subset \sqrt{nD}$$  

for an absolute constant $c > 0$. This particular very interesting case was derived long ago by Kashin [Kas77].
6. Local and global asymptotic theories

At the heart of the global results presented in this lecture stand methods of Functional Analysis. By global properties we refer to properties of the original body in question, while the local properties pertain to the structure of lower dimensional sections and projections of the body, i.e. to the linear structure of a normed space in the spirit of functional analysis.

From the beginning of the 70's the needs of Geometric Functional Analysis led to a deep investigation of the linear structure of finite dimensional normed spaces. Dvoretzky’s theorem was the initial result in this direction. But it had to develop a long way before this structure was understood well enough that it could be used for the study of the global properties of a space.

The culmination of this study was an understanding of the fact that subspaces (and quotient spaces) of proportional dimension behave very predictably. This was the bridge between the problems of Functional Analysis and the Global Asymptotic properties of convex sets. The following theorem was one of the first links between local properties and global structure [M85]:

**Theorem 6.1.** Fix $0 < \lambda < 1$. Every finite dimensional normed space $X = (\mathbb{R}^n, \| \cdot \|)$ contains a subspace $sX$ and a quotient $q(sX) = F$ of a subspace $sX$ such that

1) $k = \dim(F) \geq \lambda n$
2) $d(F, \ell^k_2) \leq C(\lambda) \sim \frac{1}{1 - \lambda} \log \frac{1}{1 - \lambda}$.

We know today how to rewrite almost every fact of Local Theory in a Global form. For example, the above theorem is (for $\lambda = \frac{1}{2}$) a “Local form” of the “Global” Theorems 2.1 and 2.2. ($\lambda$ close to 1 has no global analogue.) Also Theorem 4.1 may be considered a global form of Dvoretzky’s Theorem. However, only recently in one case did we succeed in putting this analogy in an exact quantitative form, and I end this lecture with this example.

Let $\| \cdot \|$ define the standard euclidean norm in $\mathbb{R}^n$, $X$ be $\mathbb{R}^n$ with norm $\| \cdot \|$ and $M = \int_{S^{n-1}} \|x\| d\mu(x)$. Let $k = k(X; \| \cdot \|) \leq n$ be the largest integer such that

$$
\mu_{\sigma_n, k} \left( \left\{ E : \frac{M}{2} |x| \leq \|x\| \leq 2M |x| \text{ for all } x \in E \right\} \right) > 1 - \frac{k}{n + k}.
$$

Let $t$ be the smallest integer such that there are orthogonal transformations $u_1, \ldots, u_t \in O(n)$ with

$$
\frac{M}{2} |x| \leq \frac{1}{t} \sum_{i=1}^{t} \|u_i x\| \leq 2M |x|, \text{ for all } x \in \mathbb{R}^n.
$$

So $k = k(X; \| \cdot \|)$ is a quantitative expression for a Local property of a space $X$ equipped with the given euclidean structure; at the same time, $t$ gives us a quantitative form for the corresponding global property of the space $X$.

**Theorem 6.2 [MSch96].** For some universal constant $C$

$$
C^{-1} n \leq kt \leq C n.
$$
Concluding remarks
There are a number of reasons for and origins of orderal behavior. One may mention “repetition” which creates order, as statistics demonstrates. What I have attempted to show in this lecture is that very high dimensions, or more generally, high parametric families are another source of order. Of course, time limitations have restricted me to very specific objects and structures.

Let me note in conclusion a few remarks about the techniques used in this theory, which stand behind the scenes of my lecture. In addition to the concentration phenomenon which absorbed a huge diversity of methods, the probabilistic nature of a high-dimensional space implies that probability notions and parameters should play an important, perhaps central, role. Indeed, notions of the type and cotype of a space which were developed in the 70s (through works of Maurey, Pisier, Kwapień, Krivine and others) give a “right” classification of normed spaces (convex bodies) by their probabilistic behavior. Also Harmonic Analysis techniques enter the abstract theory of Convexity through these notions (Pisier; Bourgain). Let me also note the use of delicate martingale theory, p-stable probability laws (Pisier, Johnson-Schechtman) and the construction of “random normed spaces” (Gluskin; Szarek) with a very interesting behavior, not observed in concrete examples.

I would like to thank Miriam Hercberg for helping me to edit this text and artfully weaving computer graphics into TeX. I also thank Noga Alon and Leonid Polterovich for their useful remarks on the first draft of this talk.

References

Books and Surveys


Articles


