GEOMETRY OF MEASURES: PARTITIONS AND CONVEX BODIES

ROMAN KARASEV

1. Introduction

In this course we start with several topological techniques that allow to partition several measures in a Euclidean space into equal parts, or partition the space into parts of prescribed measure. These are classical results in discrete geometry and measure theory, and they find applications in a variety of problems.

After that, we give applications to point-line incidences and spanning trees with low crossing number, following the excellent review of H. Kaplan, J. Matoušek, and M. Sharir [KMS12]. The reader is also encouraged to read the review of L. Guth on a similar topic [Guth13].

Then we discuss the monotone transportation and the Brunn–Minkowski inequality, following the nice course of K. Ball [Ball04]. We also consider the Prékopa–Leindler inequality for log-concave measures, the Minkowski theorem on facet areas, the needle decomposition, and the isoperimetric inequality for the Gaussian measure, following, in particular, the blog post of T. Tao [Tao11] and the nice paper of F. Nazarov, M. Sodin, and A. Vol'berg [NSV02].

We touch the topic of the isoperimetric inequality and concentration on the round sphere, as well as another result about the Gaussian measure, known as the Šidák lemma. Then we give some simple facts about volumes of sections of a cube, facet and vertex numbers of centrally symmetric polytopes, and sketch a proof of the Dvoretzky theorem, following another brief course of K. Ball [Ball97]. We also discuss the topological approach to the Dvoretzky theorem and recent positive and negative results in this direction.

2. The Borsuk–Ulam theorem

One common tool to prove results about partitions of measures is the classical Borsuk–Ulam theorem [Bor33]:

Theorem 2.1. *For any continuous map $f : S^n \to \mathbb{R}^n$ there exists a pair of antipodal points $x, -x \in S^n$ such that $f(x) = f(-x)$.*

Proof. By putting $g(x) = f(x) - f(-x)$, we reduce this theorem to the following: For an odd map $g : S^n \to \mathbb{R}^n$ there exists a point $x \in S^n$ such that $g(x) = 0$. A map g is called *odd* if $g(-x) = -g(x)$ for any x.

Then we consider a simple map g_0 defined as follows: If S^n is the unit sphere in $\mathbb{R}^{n+1}$, then g_0 is the projection to a coordinate subspace $\mathbb{R}^n \subset \mathbb{R}^{n+1}$. It is easy to observe that for g_0 there is a unique antipodal pair $x_0, -x_0 \in S^n$ that is mapped to zero. Moreover, at this x_0 (and $-x_0$) the Jacobian matrix Dg_0 is nondegenerate.

Assume that g does not map any point to zero. Now we connect g_0 and g by the homotopy $h_t(x) = (1-t)g_0(x)+tg(x)$. For any t the map $h_t(x)$ remains an odd continuous

2000 *Mathematics Subject Classification.* 52C35, 60D05.

Key words and phrases. ham sandwich theorem, monotone maps, log-concavity, isoperimetry, polytopes.

Supported by the Dynasty Foundation, the President's of Russian Federation grant MD-352.2012.1, the Federal Program "Scientific and scientific-pedagogical staff of innovative Russia" 2009–2013, and the Russian government project 11.G34.31.0053.

map. From standard facts of differential geometry we may perturb the homotopy h_t slightly to obtain another homotopy $\tilde{h}_t(x)$ with the following properties:

1) $h_0(x)$ is still equal to $g_0(x)$; 2) zero is a regular value for $h : S^n \times I \to \mathbb{R}^n$ ($I = [0,1]$ is the segment), and $h^{-1}(0)$ is a one-dimensional submanifold $Z \subset S^n \times I$ with boundary in $S^n \times \partial I$. 3) the map $h_1(x)$ may be not equal to $g(x)$, but it still misses zero in $\mathbb{R}^n$.

Now starting from the unique pair $\{(x_0, 0), (-x_0, 0)\} \in \partial Z$ and trace this pair along the one-dimensional set Z. This pair of point must finally arrive at some other pair $\{(x_1, t_1), (-x_1, t_1)\} \subset S^n \times I$, but there is nowhere to arrive: $t_1 = 1$ is impossible because of the assumption (3), $t_1 = 0$ would mean that the pair $(x_1, -x_1)$ is the same as $(x_0, -x_0)$ but with reversed order. The latter is impossible because if $(x_0, 0)$ and $(-x_0, 0)$ are connected by a component of Z then the antipodal action $(x, t) \mapsto (-x, t)$ would have a fixed point in this component, which is wrong.

Thus the assumption was wrong and we conclude that $g^{-1}(0)$ is nonempty. □

Let us state another similar theorem:

Theorem 2.2. *Any odd map $g : S^n \to S^n$ has odd degree.*

Proof. The proof follows from taking quotient by the antipodal action $\mathbb{R}P^n = S^n/\mathbb{Z}_2$ and considering the induced map $g' : \mathbb{R}P^n \to \mathbb{R}P^n$. One may check that the map $g'_* : H_1(\mathbb{R}P^n) \to \mathbb{H}_1(\mathbb{R}P^n)$ is an isomorphism. Then from the explicit description of the cohomology $H^*(\mathbb{R}P^n; \mathbb{F}_2) = \mathbb{F}_2[w]/(w^{n+1})$ it follows that g' induces an isomorphism in modulo 2 cohomology and therefore its degree is odd. □

The above theorem has the following corollary due to H. Hopf [Hopf44]:

Theorem 2.3. *Let M be a compact n-dimensional Riemannian manifold and $\delta > 0$ is a positive real number. For any map $f : M \to \mathbb{R}^n$ there exist two points $x, y \in M$ connected by a geodesic of length δ such that $f(x) = f(y)$.*

The proof is left to the reader. Hint: Consider the point $x \in M$ such that $f(x)$ is the extremal point of the image $f(M)$. Then for every direction $\nu \in T_x M$ consider the geodesic $\ell(t, \nu)$ from x in the direction of ν and, assuming the contrary, construct two homotopic maps from the set of directions (identified with S^{n-1}) to S^{n-1}, one of them being odd and the other being non-surjective (and therefore having zero degree).

For more information about the Borsuk–Ulam theorem the reader is referred to the book of Matoušek [Mat03].

3. The ham sandwich theorem and its polynomial version

Now we are ready to prove the classical 'ham sandwich' theorem [ST42, Ste45]:

Theorem 3.1. *Let $\mu_1, \dots, \mu_n$ be probability measures in $\mathbb{R}^n$ that attain zero on every hyperplane. Then some hyperplane H partitions $\mathbb{R}^n$ into a pair of halfspaces H^+ and H^- so that $\mu_i(H^+) = \mu_i(H^-) = 1/2$ for any i.*

Proof. We put $A = \mathbb{R}^n$ to $\mathbb{R}^{n+1}$ as the affine hyperplane defined by $x_{n+1} = 1$. Then for any unit vector $\nu \in S^n$ the inequality $(\nu, y) \ge 0$ defines a halfspace H_ν^+ in A with the complement H_ν^-. For ν equal to $(0, \dots, 0, \pm 1)$ those halfspaces become degenerate, that is coinciding with the empty set of with the whole A.

Now we consider the map $f : S^n \to \mathbb{R}^n$ defined as follows:

$$f(\nu) = (\mu_1(H_\nu^+), \dots, \mu_1(H_\nu^+)).$$

Be Theorem 2.1 there exist a pair $\nu, -\nu \in S^n$ with $\mu_i(H_\nu^+) = \mu_i(H_{-\nu}^+) = \mu_i(H_\nu^-)$ for any i. Since the total measure of A is 1 with respect to each μ_i, we obtain $\mu_i(H_\nu^+) = \mu_i(H_\nu^-) = 1/2$ for any i. □

Now we are going to consider more general partitions of the space. We start from the simplest case of the line $\mathbb{R}$ and consider the space of univariate polynomials of degree at most d, which we denote by $P_d(\mathbb{R})$. For every $f \in P_d(\mathbb{R})$ it is natural to consider the sets

$$H_f^+ = \{x : f(x) \geq 0\} \quad \text{and} \quad H_f^- = \{x : f(x) \leq 0\}.$$

We claim that for any d absolute continuous probability measures $\mu_1, \dots, \mu_d$ in $\mathbb{R}$ there exists a polynomial $f \in P_d(\mathbb{R})$ that splits (with $\mathbb{R} = H_f^+ \cup H_f^-$) every measure into two equal halves. This fact is established by considering the *moment map* $v_1^d : \mathbb{R} \to \mathbb{R}^d$ that takes $t \in \mathbb{R}$ to the vector $(t, t^2, \dots, t^d)$. The images of the measures μ_i are defined and it is important that they attain zero in every halfspace; this follows from the fact that the original μ_i attain zero on every finite set. Now we apply the ham sandwich theorem to these measures in $\mathbb{R}^d$ and obtain an equipartitioning halfspace in $\mathbb{R}^d$ with equation

$$\lambda(x) \geq 0,$$

where λ is a linear function with possible constant term. The function $\lambda(v_1^d)$ then becomes a polynomial of degree at most d in one variable. A nontrivial generalization of this one-dimensional fact for splitting into a given proportion $\alpha : (1-\alpha)$ is given in [SW85], in this case the partitioning set has to be twice more complex than in the simple case $\alpha = 1/2$.

As an exercise the reader may try to prove another result in the line:

Theorem 3.2. *Assume $f_1, \dots, f_n$ are integrable functions on the segment $[0, 1]$. Then there exists another function g, orthogonal to every f_i, that only takes values ± 1 and has at most n discontinuity points.*

The general case of the polynomial ham sandwich theorem follows by considering the Veronese map $v_n^d : \mathbb{R}^n \to \mathbb{R}^{\binom{d+n}{n}} - 1$ that takes an n-tuple $(x_1, \dots, x_n)$ to the set of all possible nonconstant monomials in x_i's of degree at most d. After counting such monomials we obtain:

Theorem 3.3. *Let n and d be positive integers and $r = \binom{d+n}{n} - 1$. Then any r absolutely continuous measures $\mu_1, \dots, \mu_r$ in $\mathbb{R}^n$ may be partitioned into equal halves simultaneously by a partition $\mathbb{R}^n = H_f^+ \cup H_f^-$, where f is a polynomial of degree at most d.*

This theorem has a version for partitioning finite point sets, which is frequently needed in different problems:

Theorem 3.4. *Let n and d be positive integers and $r = \binom{d+n}{n} - 1$. Then for any r finite sets $X_1, \dots, X_r$ in $\mathbb{R}^n$ there exists a partition $\mathbb{R}^n = H_f^+ \cup H_f^-$, where f is a polynomial of degree at most d, such that $|X_i \cap H_f^+|, |X_i \cap H_f^-| \geq 1/2|X_i|$ for any i.*

Proof. Replace every point $x \in X_i$ with a density distributed uniformly over a ball $B_\varepsilon(x)$ and sum those densities over all $x \in X_i$ to obtain the density of the measure μ_i.

Then apply Theorem 3.3 to μ_i and pass to the limit $\varepsilon \to +0$. It is easy to see that all possible partitioning polynomials f_ε may be chosen to be contained in a bounded subset of $P_d(\mathbb{R}^n)$ and therefore it is possible to select a limit polynomial f that will satisfy the requirements. □

4. Partitioning a single point set with successive polynomials cuts

In the review of Kaplan, Matoušek, and Sharir [KMS12] the importance of the following corollary of the polynomial ham sandwich theorem is emphasized:

Lemma 4.1. *Let X be a finite set in $\mathbb{R}^n$ and r be a positive integer. It is possible to find a polynomial of degree at most $C_n r^{1/n}$ with the following property: The set $Z = \{x : f(x) = 0\}$ partitions $\mathbb{R}^n$ into connected components $V_1, \ldots, V_N$ so that $|X \cap V_i| \le 1/r|X|$ for every i.*

Proof. We first use Theorem 3.4 to partition X into almost equal halves using the zero set Z_{f_1} of a linear function f_1. Then we partition every part into equal halves with another zero set Z_{f_2} of a function f_2, which may be still chosen to be linear if $n \ge 2$. Then we do the same j times. After that, we have a collections of polynomials $f_1, \ldots, f_j$ and consider their product $f = f_1 f_2 \ldots f_j$. The zero set Z_f partitions $\mathbb{R}^n$ into at least $r = 2^j$ connected components, each containing at most $1/r$ fraction of the set X.

It remains to bound from above the degree of f. On the i-th step we partitioned 2^{i-1} sets and the required degree of the polynomials was at most $(n!2^{i-1})^{1/n}$. The summation over i of this geometric progression gives the estimate

$$\deg f \le \frac{(n!r)^{1/n}}{1 - 2^{-1/n}} = C_n r^{1/n}.$$

We proved the result for r powers of two, for other r we could choose 2^j to be the least power of two not less than r. □

Following [KMS12], we make several comments on this lemma. Seemingly we partitioned the space into 2^j parts, but some parts could actually split into several connected component in that process. So we actually do not control the number of parts. The other issue is that some points of X (and actually many of them) can lie on the set Z_f and need a separate treatment in most applications.

One may consider a simpler approach that gives partition into convex parts with larger intersection with lines. We may partition a measure into equal halves with a line in arbitrary direction. Then we can partition both parts simultaneously into equal quarters by the ham sandwich theorem, on this step the partitioning line is unique. Therefore we obtain a partition into 4 equal parts such that any line intersects (essentially intersects in the interior) at most 3 of them. Iterating this procedure hierarchically in k steps we partition a measure into $N = 4^k$ parts, and it is easy to see that any line intersects at most $3^k = N^{\log 3/\log 4}$ of them. This estimate is asymptotically worse than the one obtained with polynomial cuts, but is has an advantage that the parts are convex.

When trying to generalize the above example to higher dimensions and intersections with hyperplanes, we see that it is not trivial to find a convex equipartition of a single measure so that every hyperplane does not intersect at least one of them in the interior. The corresponding result is known as the Yao–Yao theorem [YY85]:

Theorem 4.2. *It is possible to partition an absolutely continuous finite measure in $\mathbb{R}^n$ into 2^n equal convex parts so that any hyperplane does not intersect the interior of at least one of the parts.*

Sketch of the proof. We are not going to make the full proof because it is quite technical and hard to visualize, we only sketch the main ideas instead.

The two-dimensional case is already proved. Then we make induction on the dimension and try to find a partition which is a twisted (in some sense) partition into coordinate orthants. We select the basis $e_1, \ldots, e_n$ and partition the measure μ into equal halves with a hyperplane H perpendicular to e_1. Then we consider all possible unit vectors v such that $(v, e_1) > 0$ and project both halves of the measure onto H along v. For every one of the halves we obtain an $(n-1)$-dimensional Yao–Yao partition, and it is possible to prove by induction that it is unique once the basis $e_2, \dots, e_n$ in H is selected.

Then a version of the Brouwer fixed point theorem (a similar fact is Lemma 9.2 below) helps to prove that for some v the *centers* of the Yao–Yao partitions for the two halves coincide and give the new center c.

This gives the required partition, because every hyperplane H' is either parallel to H, in this case everything is clear, or intersects H in an affine $(n-2)$-subspace. One of the rays $\{c+tv\}_{t>0}$ and $\{c-tv\}_{t>0}$ is not touched by H' and we select the corresponding half H^+ or H^-, let it be H^+ without loss of generality. Applying the inductive assumption to the projection of $\mu|_{H^+}$ along v and the intersection $H \cap H'$ we find a part in H^+, whose interior is not intersected by H'.

To finish the proof one has to prove that the Yao–Yao center c is defined uniquely by μ and $e_1, \dots, e_n$. We omit these details. ☐

Then it is easy to iterate and obtain a partition into N equal convex parts so that any hyperplane intersects at most $N^{\frac{\log(2^n-1)}{\log 2^n}}$ of them in interiors. The polynomial partition can give a better result: it is possible to partition a measure into N equal parts so that any hyperplane intersects at most $O(N^{\frac{n-1}{n}})$ of them, see [KMS12] for the details. But for the polynomial partition, the parts may be non-convex and even not connected, so convex partitions are still useful in some cases. The reader is referred to the paper of Bukh and Hubard [BH11] for an interesting application of the Yao–Yao theorem.

5. The Szemerédi–Trotter theorem

We are going to apply Lemma 4.1 and deduce the Szemerédi–Trotter theorem about the number of incidences between points and lines. We start from the definition:

Definition 5.1. Let P be a set of points and L be a set of lines in the plane. Denote by $I(P, L)$ their *incidence number*, that is the number of pairs $(p, \ell) \in P \times L$ such that $p \in \ell$.

Theorem 5.2. *In the plane* $I(P, L) \le C(|P|^{2/3}|L|^{2/3} + |P| + |L|)$ *for a suitable absolute constant* C.

Remark 5.3. This theorem is also valid for *pseudolines*, that is subsets of the place that behave like lines in terms of their intersection. Such a generalization so far seems to be out of reach of algebraic methods.

Proof. We start from a much weaker estimate, which we are going to apply to different sets of points and lines:

Lemma 5.4. $I(P, L) \le |L| + |P|^2$.

This lemma is proved by splitting L into two families: one for the lines intersecting at most one point of P and all other lines. The details are left to the reader.

Now we put $m = |P|$ and $n = |L|$. Take some parameter r, whose value we will define later. We choose an algebraic set Z of degree $O(\sqrt{r})$ (from now on we use the notation $O(\cdot)$ to avoid different constants) that partitions $\mathbb{R}^2$ into connected sets $V_1, \dots, V_N$.

Let $P_0 = P \cap Z$ and $P_i = P \cap V_i$ for any i. Also denote by L_0 the lines in L that lie entirely on Z and denote by L_i the lines in L that intersect V_i. Note that L_i's are not disjoint, and $|L_0| \le \deg Z = O(\sqrt{r})$. The crucial fact is that every line from $L \setminus L_0$ intersects Z in at most $O(\sqrt{r})$ points and intersects at most $O(\sqrt{r})$ of the regions V_i.

First, we obviously estimate:

$$I(P_0, L_0) \leq m|L_0| = O(m\sqrt{r}), \quad \sum_i I(P_0, L_i) = nO(\sqrt{r}), \quad I(P_i, L_0) = 0.$$

Summing up those obvious estimates we obtain $O((m+n)\sqrt{r})$ in total. It remains to use Lemma 5.4 and bound

$$\sum_i I(P_i, L_i) \leq \sum_i |L_i| + |P_i|^2 \leq nO(\sqrt{r}) + m^2/r.$$

Now we make several observations. The projective duality allows us to interchange points and lines and assume $m \leq n$. Then Lemma 5.4 allows us to concentrate on the case $\sqrt{n} \leq m \leq n$. After that putting $r = \frac{m^{4/3}}{n^{2/3}}$ we make all the estimates made so far to be of the form $O(n^{2/3}m^{2/3})$. □

An interesting application of the Szemerédi–Trotter theorem is the sum-product estimates. We quote the simplest of them, due to G. Elekes. For a subset A of a ring put

$$A + A = \{a_1 + a_2 : a_1, a_2 \in A\}, \quad A \cdot A = \{a_1 a_2 : a_1, a_2 \in A\}.$$

Theorem 5.5. *For any finite subset A of $\mathbb{R}$ we have:*

$$|A + A| \cdot |A \cdot A| \geq C|A|^{5/2}$$

for an absolute constant C.

Proof. Consider the set of points in $\mathbb{R}^2$

$$P = \{(b + c, ac) : a, b, c \in A\}$$

and the set of lines

$$L = \{y = a(x - b) : a, b \in A\}.$$

Obviously, $|L| = |A|^2$ and every $\ell \in L$ contains at least $|A|$ points of P with given a and b and variable c. Hence

$$|I(P, L)| \geq |A|^3.$$

Now by the Szemerédi–Trotter theorem

$$|P| \cdot |L| \geq C|I(P, L)|^{3/2}, \Rightarrow |P| \cdot |A|^2 \geq C|A|^{9/2},$$

and therefore $|P| \geq C|A|^{5/2}$. But $P \subseteq (A + A) \times A \cdot A$ and we obtain the required inequality. □

Theorem 5.5 implies that at least one of the cardinalities $|A + A|$ or $|A \cdot A|$ is at least $\sqrt{C}|A|^{5/4}$. The exponent $5/4$ can be slightly improved through a more careful counting, see the details in [TaoVu10, Section 8.3]. Actually, P. Erdős and E. Szemerédi conjectured that it is possible to replace $5/4$ with $2 - \varepsilon$ with arbitrarily small positive ε, this remains an open problem.

6. Spanning trees with low crossing number

Sometimes it is important to estimate the number of parts for the partition in Lemma 4.1. In order to do this we pass to the projective space and prove:

Lemma 6.1. *A hypersurface $Z \subset \mathbb{R}P^n$ of degree d partitions $\mathbb{R}P^n$ into at most d^n connected components.*

Proof. Select a hyperplane $H \subset \mathbb{R}P^n$ that intersects Z generically. Without loss of generality assume H to be the hyperplane at infinity, which is naturally $\mathbb{R}P^{n-1}$.

By the inductive assumption H intersects at most d^{n-1} components of $\mathbb{R}P^n \setminus Z$. Other components $C_1, \dots, C_N$ are bounded. When considering the affine space $\mathbb{R}^n = \mathbb{R}P^n \setminus H$, we choose a degree d polynomial f with the set of zeros Z. Every bounded component C_i has the property that on its boundary f vanishes and keeps sign in its interior. Hence every C_i has a maximum or a minimum of f in its interior. Every critical point is a root of the system of equations:

$$\begin{cases} \frac{\partial f}{\partial x_1}(x) &= 0 \\ & \dots \\ \frac{\partial f}{\partial x_n}(x) &= 0 \end{cases}$$

These are algebraic equations of degree at most $d-1$ each and therefore the system has at most $(d-1)^n$ solutions. Of course, to conclude this we need the solution set to be discrete. The general case is handled by bounding not the number of solutions, but the number of connected components of solutions. By an appropriate perturbation of the equations we may split the components of its solution set into several isolated points each, thus proving what we need.

Now we have at most $(d-1)^n$ bounded components and at most d^{n-1} unbounded components. The total number of components is therefore at most d^n. □

For the affine case we note that every infinite component in projective setting may give two components in affine setting, therefore in $\mathbb{R}^n \setminus Z$ we have at most $(d-1)^n + d^{n-1}$ components (from the proof above), which is at most $d^n + 1$ always. Hence, in Lemma 4.1 the number of components is actually of order r.

For the number of connected components of the set Z itself, there is a more precise result in the plane. This number is bounded from above by the number $\binom{\deg Z - 1}{2} + 1$ by the Harnack theorem [Har76]. The reader may try to prove the Harnack theorem considering the real algebraic curve as a set of cycles on the corresponding complex algebraic curve of genus $g = \binom{\deg Z - 1}{2}$.

Now we give another application of Lemma 4.1 is the following theorem of Chazelle and Welzl [Wel88, Cha89, Wel92]:

Theorem 6.2. *Any finite set $P \subset \mathbb{R}^2$ has a spanning tree T with the following property: Any line ℓ (apart from a finite number of exceptions) intersects T in at most $C\sqrt{|P|}$ points.*

Proof. The first observation is that it is sufficient to find an arcwise connected subset X containing P and having small crossings with almost all lines. Then it is easy to select a tree T inside X that will still connect P and has crossings at most twice of the crossings of X. Then the edges of T may be replaced with straight line segments without increasing the number of crossings. The details of this reduction are left to the reader.

Now we prove the following:

Lemma 6.3. *It is possible to find a set $Y \supset P$ with at most $|P|/2$ connected components and line crossing number at most $C\sqrt{|P|}$.*

The lemma is proved as follows: Taking $r = |P|/C_1$ we obtain by Lemma 4.1 an algebraic set Z of degree at most $C_2\sqrt{|P|/C_1}$ that splits P into parts of size at most C_1. By Lemma 6.1 for sufficiently large C_1 (but still an absolute constant) the number of connected components of $\mathbb{R}^2 \setminus Z$, and therefore Z, will be at most $|P|/2$. Then in every component V_i of $\mathbb{R}^2 \setminus Z$ we have at most C_1 points of P, which we span by a tree T_i and attach this tree to the set Z. Put $Y = Z \cup T_1 \cup \dots \cup T_N$. Any line ℓ (apart from a finite number of exceptions) will intersect Z at most $\sqrt{|P|}$ times and will intersect at most $\sqrt{|P|}$ trees of T_i. Hence this line will have at most $(1 + C_1)\sqrt{|P|}$ points of intersection with Y. The lemma is proved.

Now we apply the lemma once, then select a point in every component of Y thus obtaining the set P_2 with $|P_2| \le 1/2|P|$. Then apply the lemma again to P_2, pass to another point set P_3 with $|P_3| \le 1/2|P_2|$ and so on. As it was in the proof of Lemma 4.1, in $\log |P|$ number of steps we arrive at a connected set $X = Y_1 \cup Y_2 \cup \dots$ spanning P. The number of crossings of X with a line ℓ is bounded from above by the sum of a geometric progression with denominator $2^{-1/2}$ and the leading term $(1 + C_1)\sqrt{|P|}$, so it is bounded by $C\sqrt{|P|}$, where C is another absolute constant.

□

The reader is now referred to the review [KMS12] and a more advanced paper of Solymosi and Tao [ST12] for other interesting applications of Lemma 4.1.

7. Counting point arrangements and polytopes in $\mathbb{R}^d$

Lemma 6.1 of the previous section has interesting applications to estimating the number of configurations of n points in $\mathbb{R}^d$ up to a certain equivalence of relation.

Definition 7.1. Let $x_1, \dots, x_n \in \mathbb{R}^d$ be an ordered set of points. We define its *order type* to be the assignment of signs

$$\operatorname{sgn} \det(x_{i_1} - x_{i_0}, \dots, x_{i_d} - x_{i_0})$$

to all $(d+1)$-tuples $1 \le i_0 < \dots < i_d \le n$. The configuration $x_1, \dots, x_n$ is in *general position* if all those determinants are nonzero.

Now we can prove:

Theorem 7.2. *The number of distinct order types for ordered sets of n points in general position in $\mathbb{R}^d$ is at most $n^{d(d+1)n}$.*

Proof. A configuration in general is characterized by nd coordinates of all its points. Its order type depends on the signs of some $\binom{n}{d+1}$ polynomials of degree d each, we denote the product of these polynomials by $P(x_1, \dots, x_n)$, this polynomial has degree $d\binom{n}{d+1}$ and nd variables.

It is obvious that distinct order types of sets in general position must correspond to distinct connected components of the zero set of P. Hence, by Lemma 6.1 and the remark about its affine case, we have at most

$$\left(d\binom{n}{d+1}\right)^{nd} + 1 \le \frac{n^{d(d+1)n}}{d^{nd^2/2}}$$

such connected components and order types. □

The above argument comes from the paper [GP86] of Jacob Eli Goodman and Richard Pollack. After that they easily observe that the number of combinatorially distinct simplicial polytopes on n vertices in dimension d is also at most $n^{d(d+1)n}$. The generalization of this result to possibly not simplicial polytopes and some other improvements can be found in the paper [Alon86] of Noga Alon.

In the case of $d = 4$ these results bound the number of polytopal triangulations of the 3-sphere on n vertices. Curiously (see [NW13] and the references therein) it is possible to construct much more triangulations of the 3-sphere of n vertices and conclude that most of them are not coming from any 4-dimensional polytope.

8. Chromatic number of graphs from hyperplane transversals

A wide range of applications of the appropriately generalized Borsuk–Ulam theorem started when László Lovász proved [Lov78] an estimate on the chromatic number of a certain graph, known as the Kneser conjecture. Let us give some definitions, we denote the segment of integers $\{1, \dots, n\}$ by $[n]$:

Definition 8.1. Let $G = (V, E)$ be a graph with vertices V and edges E. Its *chromatic number* $\chi(G)$ is the minimum number χ such that there exists a map $V \to [\chi]$ sending every edge $e \in E$ to two distinct numbers (colors).

Definition 8.2. The Kneser graph $K(n, k)$ has all k-element subsets of $[n]$ as the sets of vertices V and two vertices $X_1, X_2 \in V$ form an edge if they are disjoint as subsets of $[n]$.

A simple argument, left as an exercise to the reader, shows that it is possible to color $K(n, k)$ in $n - 2k + 2$ colors in a regular way. The opposite bound $\chi(K(n, k)) \geq n - 2k + 2$ was much harder to establish. In order to prove it, we follow the approach of Dol'nikov [Dol94] and first prove the hyperplane transversal theorem:

Theorem 8.3. *Let $\mathcal{F}_1, \dots, \mathcal{F}_n$ be families of convex compacta in $\mathbb{R}^n$. Assume that every two sets C, C' from the same family $\mathcal{F}_i$ have a common point. Then there exists a hyperplane H intersecting all the sets of $\bigcup_i \mathcal{F}$.*

Proof. For every normal direction ν every set $C \in \bigcup_i \mathcal{F}$ gives a segment of values of the product $\nu \cdot x$ for $x \in C$. For a given family $\mathcal{F}_i$ all these segments intersect pairwise, and therefore have a common point of intersection. Let this point be d_i. In other words, all sets of $\mathcal{F}_i$ intersect the hyperplane

$$H_i(n) = \{x : \nu \cdot x = d_i\}.$$

Actually, the values d_i can be chosen to depend continuously on ν (if we choose the middle of all candidates, for example) and by definition they are also odd functions of n. The combinations $d_2 - d_1, \dots, d_n - d_1$ make an odd map $S^{n-1} \to \mathbb{R}^{n-1}$, which must take the zero value according to the Borsuk–Ulam theorem.

Hence, for some direction ν we can put $d_1 = d_2 = \dots = d_n$ and the corresponding hyperplane will intersect all members of the family $\bigcup_i \mathcal{F}$. □

Now we prove:

Theorem 8.4. *Let $n \geq 2k$. For the Kneser graph we have: $\chi(K(n, k)) = n - 2k + 2$.*

Proof. The upper bounds is already mentioned, so we prove the lower bound. Assume the contrary and consider a coloring of the graph in $n - 2k + 1$ or less colors.

Put the n vertices of the underlying set (from Definition 8.2) as a general position finite point set X in $\mathbb{R}^{n-2k+1}$. For any color $i = 1, \dots, n-2k+1$, let $\mathcal{F}_i$ consist of all convex hulls of the k-element subsets of X that correspond to the ith color of the coloring of $K(n, k)$. Since the coloring is proper, any two k-element subsets of X with the same color have a common point and therefore these families $\mathcal{F}_i$ satisfy the assumptions of Theorem 8.3.

Hence there is a hyperplane H touching every convex hull of every k-element subset of X. But this is impossible: H can contain itself at most $n - 2k + 1$ points of X from the general position assumption, of $2k - 1$ remaining points some k must lie on one side of H; and therefore the convex hull of this k-tuple is not intersected by H. This is a contradiction. □

It is in fact possible to find a much smaller induced subgraph of $K(n, k)$ having the same chromatic number.

Definition 8.5. Assume the ground set of n elements is arranged in a circle. Its subset is called a *Schrijver subset* if it contains no two consecutive elements in the circular order. The Schrijver graph $S(n,k)$ is a the subgraph of $K(n,k)$ induced on Schrijver sets.

Theorem 8.6. *Let $n \geq 2k$. For the Schrijver graph we have: $\chi(S(n,k)) = n - 2k + 2$.*

Proof. We still need to use Dolnikov's hyperplane transversal theorem, but in a more sophisticated manner.

Let the ground set $X = [n]$ be identified with a subset of the circle $S^1 = \mathbb{R}/2\pi\mathbb{Z}$. Let V be the space of trigonometric polynomials of order $\leq k-1$,

$$V = \left\{ a_0 + \sum_{j=1}^{k-1} (a_j \cos jx + b_j \sin jx) \right\}.$$

We will consider the restriction of functions on the circle S^1 to the set X, the set of all functions $f : X \to \mathbb{R}$ is then naturally identified with $\mathbb{R}^n$. A nonzero function in V cannot have more than $2k-2$ zeros in the circle and therefore its restriction to X is also nonzero. Hence we may assume that $V \subset \mathbb{R}^n$ and $\dim V = 2k-1$. Let W be the orthogonal complement of V in $\mathbb{R}^n$, thus $\dim W = n - 2k + 1$.

We have a natural map

$$I : P \to W^*, \quad p \mapsto (f \mapsto f(p)).$$

As in the proof of Theorem 8.4 we assume a coloring of the Schrijver k-tuples in $n-2k+1$ colors so that every two k-tuples of the same color intersect. The images of the k-tuples under I have the property that every two k-tuples of the same color intersect and the number of colors equals the dimension of W^*. By Theorem 8.3 there exists a hyperplane in W^* that intersects all the convex hulls of images of Schrijver k-tuples. It is given by the equation $f = a$, where $f \in (W^*)^* = W$.

Without loss of generality assume $a \geq 0$. Look at the elements of X whose image under I lies in the halfspace $\{f < a\}$, we want to find a Schrijver k-tuple of such elements contradicting the fact that $\{f = a\}$ intersects all the convex hulls of Schrijver k-tuples. In fact, we will be done if we find a Schrijver k-tuple of the elements of X where $f < 0$.

Consider the subset $N = \{p \in X : f(p) < 0\}$. If this subset consists of at least k segments of consecutive points (in the circular order) then we can take a point from each segment and they will make a Schrijver k-tuple. Otherwise it is possible to have precisely $k-1$ segments $[u_1, v_1], \dots, [u_{k-1}, v_{k-1}]$ of the circle with endpoints not in X such that the points of X in these segments are precisely the points of N (some segments may contain no point from X and are added just to have $k-1$ segments in total).

The system of $2k-2$ equations

$$g(u_1) = g(v_1) = \cdots = g(u_{k-1}) = g(v_{k-1}) = 0$$

has a nonzero solution in V, since $\dim V > 2k-2$. Since g cannot have more than $2k-2$ zeros, it only changes sign at the points u_i, v_i. Hence we may choose g positive outside the union of the segments $[u_1, v_1], \dots, [u_{k-1}, v_{k-1}]$ and negative inside the segments. By the definition of N, the product $g(p)f(p)$ is non-negative on X and is positive on N. In case $N = \emptyset$ the product must also be positive on some point $p \in X$ since f represents a nonzero element of W. In any case,

$$\sum_{p \in P} g(p)f(p) > 0$$

that contradicts the orthogonality of $g \in V$ and $f \in W$. $\square$

Theorem 8.3 can also be generalized the following way:

Theorem 8.7. *Let $\mathcal{F}_0, \dots, \mathcal{F}_k$ be families of convex compacta in $\mathbb{R}^n$. Assume that every $n-k+1$ or less number of sets from the same family $\mathcal{F}_i$ have a common point. Then there exists a k-dimensional affine subspace L intersecting all the sets of $\bigcup_i \mathcal{F}$.*

The proof of this result [Dol94] uses a more advanced topological fact: Any $n-k$ sections of the canonical k-dimensional bundle over the real Grassmannian $G_{n,k}$ have a common zero. We sketch the proof as follows: For a generic set of $n-k$ sections, there is an odd number of such common zeros. This is established by considering a certain "linear" set of sections with precisely one nondegenerate common zero, and then showing that the parity of the number of common zeros does not depend on the choice of the generic set of $n-k$ sections of the bundle. This is essentially the same argument that shows that the degree (modulo 2) of a proper map between manifolds of the same dimension is well-defined.

Arguing like in the proof of Theorem 8.4 it is possible to establish some results about chromatic numbers of certain hypergraphs, see [ABMR11] for example. Though the results for hypergraphs are not so precise as in the case of graphs.

More systematic topological approach of [Lov78] (see also the book [Mat03]) to the chromatic number of graphs works as follows. For any two graphs G, H consider the *homomorphisms* $G \to H$, that is maps between the vertex sets $V(G)$ and $V(H)$ sending every edge of G to an edge of H. There is a natural way to consider such homomorphisms as the vertex set of a simplicial (or cellular) complex $\mathrm{Hom}(G, H)$.

Now we check that the definition of the chromatic number of G reads as the smallest χ such that there exists a homomorphism from G to K_χ, where K_χ is the full graph of χ vertices. Such a homomorphism induces a morphism of simplicial complexes $c : \mathrm{Hom}(I_2, G) \to \mathrm{Hom}(I_2, K_\chi)$, where I_2 is the *segment graph* with two vertices and an edge between them. Now the crucial facts are:

- We can interchange the vertices of I_2 thus obtaining an involution on both the simplicial complexes $\mathrm{Hom}(I_2, G)$ and $\mathrm{Hom}(I_2, K_\chi)$.
- The complex $\mathrm{Hom}(I_2, K_\chi)$ has as faces pairs F_1, F_2 of disjoint subsets of $[\chi]$, and therefore is combinatorially the $(\chi-1)$-dimensional boundary of the crosspolytope (the higher dimensional octahedron) in $\mathbb{R}^\chi$. This is the same as the $(\chi-1)$-dimensional sphere with the standard involution.
- In some cases it is possible to map a sphere S^n of dimension larger than $\chi - 1$ to $\mathrm{Hom}(I_2, G)$, so that this map respects the involution on the sphere S^n and on $\mathrm{Hom}(I_2, G)$. In particular, it is possible when the simplicial complex $\mathrm{Hom}(I_2, G)$ is $(\chi - 1)$-connected.
- Then the existence of the map $c : \mathrm{Hom}(I_2, G) \to \mathrm{Hom}(I_2, K_\chi)$, commuting with involution, contradicts the Borsuk–Ulam theorem.

For more information the reader is referred to the book [Mat03].

9. Partition into prescribed parts

Here we stop using the ham sandwich theorem and introduce another technique of measure partitions that has some useful consequences.

Let us introduce the notion of a *generalized Voronoi partition*. The standard Voronoi partition is defined as follows: Start from a finite point set $\{x_1, \dots, x_m\} \subset \mathbb{R}^n$ and put

$$R_i = \{x \in \mathbb{R}^n : \forall j \neq i\ |x - x_i| \leq |x - x_j|\}.$$

The sets $\{R_i\}$ give a partition of $\mathbb{R}^n$ into convex parts. A generalization of a Voronoi partition is obtained when we assign weights w_i to every point x_i and put:

$$R_i = \{x \in \mathbb{R}^n : \forall j \neq i\ |x - x_i|^2 - w_i \leq |x - x_j|^2 - w_j\}.$$

In this case the inequalities in the definition are actually linear, because the both parts have the same quadratic in x term $|x|^2$.

Having noticed the linear nature of the generalized Voronoi partition we may redefine it as follows. Let us work with a finite dimensional linear space V and its dual V^*. For a finite set of linear forms $\lambda_1, \dots, \lambda_m \in V^*$ and weights $w_1, \dots, w_m \in \mathbb{R}$ put

$$R_i = \{x \in \mathbb{R}^n : \forall j \neq i \ \lambda_i(x) + w_i \geq \lambda_j(x) + w_j\}.$$

Definitely, every generalized Voronoi partition has this kind of representation, and the reader may check that the converse is also true. With such a definition it is clear that the regions R_i are projections of facets of the convex polyhedron $G^+ \subset V \times \mathbb{R}$ defined by the system of linear inequalities:

$$\begin{array}{rcl} y & \geq & \lambda_1(x) + w_1 \\ & \dots & \\ y & \geq & \lambda_m(x) + w_m. \end{array}$$

We are going to establish the following fact:

Theorem 9.1. *Let μ be a probability measure on V that attains zero on every hyperplane, $\lambda_1, \dots, \lambda_m \in V^*$ be a system of linear forms, and $\alpha_1, \dots, \alpha_m$ be positive integers with unit sum. Then there exists weights $w_1, \dots, w_m \in \mathbb{R}$ such that the generalized Voronoi partition $\{R_i\}$ corresponding to $\{\lambda_i\}$ and $\{w_i\}$ has the following property:*

$$\mu(R_i) = \alpha_i.$$

Before proving it we exhibit an appropriate topological tool:

Lemma 9.2. *Assume that a continuous map $f : \Delta \to \Delta$ of the n-dimensional simplex to itself maps every face of Δ to itself. Then the map f is surjective.*

Proof. We prove a stronger assertion, the map f of the pair $(\Delta, \partial\Delta)$ to itself has degree 1, by induction.

Since every facet $\partial_i\Delta$ (for $i = 0, \dots, n$) is mapped to itself with degree 1, then the restriction of f to $\partial\Delta$ has degree 1. This means that the map $f_* : H_{n-1}(\partial\Delta) \to H_{n-1}(\partial\Delta)$ is the identity. From the long exact sequence of reduced homology groups

$$0 = \tilde{H}_n(\Delta) \longrightarrow H_n(\Delta, \partial\Delta) \xrightarrow{\delta} \tilde{H}_{n-1}(\partial\Delta) \longrightarrow \tilde{H}_{n-1}(\Delta) = 0$$

we obtain that $f_* : H_n(\Delta, \partial\Delta) \to H_n(\Delta, \partial\Delta)$ also must be an isomorphism. The lemma is proved. ☐

Proof of Theorem 9.1. Consider the barycentric coordinates $t_1, \dots, t_m$ in an $(m-1)$-dimensional simplex Δ. Define the map $f : \Delta \to \Delta$ as follows: For a point $(t_1, \dots, t_m)$ consider the set of weights $(-1/t_1, \dots, -1/t_m)$ and let $f(t_1, \dots, t_m)$ be the set $(\mu(R_1), \dots, \mu(R_m))$ that corresponds to the partition $\{R_i\}$ with given weights.

It is easy to check that when some t_i's vanish and we substitute $w_i = -\infty$ the definition remains valid and the map f is continuous up to the boundary of Δ. Moreover, when t_i vanishes and w_i turns to $-\infty$ the corresponding region R_i becomes empty, and therefore f maps faces of Δ to faces. Now we apply Lemma 9.2 and conclude that the point $(\alpha_1, \dots, \alpha_m)$ must be in the image of f. ☐

Lemma 9.2 also allows to prove several classical results. Here is the Brouwer fixed point theorem:

Theorem 9.3. *Every continuous map f from a convex compactum to itself has a* fixed point*, that is a point x such that $f(x) = x$.*

Proof. Topologically every convex compactum is homeomorphic to a simplex Δ of appropriate dimension. So we assume $f : \Delta \to \Delta$. We also assume that the center of Δ is the origin and its diameter is 1.

If $f(x)$ is never equal to x then from compactness $|f(x) - x| \geq \varepsilon$ for some positive constant ε. Now we replace f with another map

$$\tilde{f}(x) = (1 - \varepsilon/2)f(x),$$

which still has no fixed points and maps Δ to its interior. Next, we construct the continuous map $g : \Delta \to \Delta$ as follows: take the ray $\rho(x)$ from $\tilde{f}(x)$ towards x and mark the first time it touches the boundary $\partial\Delta$, this is the point $g(x)$. The assumptions that $\tilde{f}$ has no fixed points and maps the simplex to its interior guarantee that $g(x)$ is continuous. From the construction it is clear that $g(x) = x$ for any $x \in \partial\Delta$. Therefore Lemma 9.2 is applicable and g must be surjective. But the construction also implies that its image is $\partial\Delta$ and therefore we have a contradiction. □

Another application is the classical Knaster–Kuratowski–Mazurkievicz theorem:

Theorem 9.4. *Consider the n-dimensional simplex Δ with facets $\partial_0\Delta, \dots, \partial_n\Delta$. Assume $\{X_i\}_{i=0}^n$ is a closed covering of Δ such that any X_i does not intersect its respective $\partial_i\Delta$. Then the intersection $\bigcap_{i=0}^n X_i$ is not empty.*

Hint. Replace the covering with the corresponding continuous partition of unity. □

10. Monotone maps

Theorem 9.1 has the following interpretation. Let μ be a probability measure on V, zero on hyperplanes, and ν be a discrete measure on V^*, assigning to every λ_i from the finite set $\{\lambda_i\} \subset V^*$ the measure α_i. Then the convex piecewise linear function

$$u(x) = \sup_{1\leq i\leq m} (\lambda_i(x) + w_i)$$

has the following property: The (discontinuous) map $f : V \to V^*$, defined by $f : x \mapsto du_x$, transports the measure μ to the measure ν.

Using the approximation of arbitrary measures by discrete measures, we may replace ν with any "reasonably good" measure on V^* and obtain the following theorem:

Theorem 10.1. *For two absolutely continuous probability measures μ on V and ν on V^* there exits a convex function $u : V \to \mathbb{R}$ such that the map $f : x \mapsto du_x$ sends μ to ν.*

The maps given by $x \mapsto du_x$ with a convex u are called *monotone maps*. More precise claims about the assumptions on μ and ν and continuity of the resulting map f in Theorem 10.1 can be found in the beautiful review [Ball04]. If the map f is continuously differentiable, then its differential Df turns out to be a positive semidefinite quadratic form, and the conclusion that μ is sent to ν just means that $\rho_\mu = \det Df \cdot \rho_\nu$ for the corresponding densities of the measures.

From the general properties of convex functions one easily deduces that

$$\langle x - y, f(x) - f(y)\rangle \geq 0 \tag{10.1}$$

for a monotone map, and the inequality becomes strict if $x \neq y$, ρ_μ is everywhere positive and ρ_ν is defined.

Another description of a monotone map (see [Ball04] for details) for μ and ν is a map that sends μ to ν and maximizes the following "cost function":

$$C(f) = \int_{\mathbb{R}^n} \langle x, f(x)\rangle\, d\mu.$$

Also, it is possible to describe the function u and its polar w as the solution to the linear optimization problem:

$$\int_V u\,d\mu + \int_{V^*} w\,d\nu \to \min \quad \text{under constraints} \quad \forall x\in V, y\in V^*\ u(x)+w(y)\ge \langle x,y\rangle.$$

In the paper of M. Gromov [Grom90] we find an example of an explicit construction of a monotone map. Let a measure μ be supported in a convex compactum K so that $\operatorname{conv}\operatorname{supp}\mu = K \subset V$ and put for any $k\in V^*$:

$$u(k) = \log \int_K e^{\langle k,x\rangle}\,d\mu(x).$$

It can be checked by hand that the differential map $f(k) = du_k : V^* \to V$ maps V^* precisely to the relative interior of K and its differential Df is a symmetric positive semidefinite matrix, which is positive definite if K has nonempty interior. Therefore f is injective and the volume of K is found as

$$\operatorname{vol} K = \int_{V^*} \det Df\,dk.$$

This formula may be used as a starting point in studying mixed volumes, see Section 13. By straightforward differentiation we observe a curious fact: The values $f(k)$ and $Df(k)$ are the first and the second moment of the measure $e^{\langle k,x\rangle}\mu$ after its normalization.

11. The Brunn–Minkowski inequality and isoperimetry

An interesting application of monotone maps (following [Ball04]) is:

Theorem 11.1 (The Brunn–Minkowski inequality)**.** *Let A and B be open subsets of $\mathbb{R}^n$ of finite volume each, then*

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

where $A+B$ denotes the Minkowski sum $\{a+b : a\in A,\ b\in B\}$ of A and B.

Proof. Consider the probability measures μ and ν distributed uniformly on A and B respectively. Let f be the monotone map between them, this means that $\det Df_x = V_B/V_A$ at any point $x\in A$, where we put $V_A = \operatorname{vol} A$ and $V_B = \operatorname{vol} B$ for brevity.

Note that by (10.1) and the remark after it the map $g(x) = x + f(x)$ is injective on A and therefore

$$\operatorname{vol}(A+B) \ge \int_A \det Dg\,dx = \int_A \det(\mathrm{id} + Df(x))\,dx.$$

Now we are going to use the inequality $\det(X+Y)^{1/n} \ge \det X^{1/n} + \det Y^{1/n}$ for positive semidefinite $n\times n$ matrices (the Brunn–Minkowski inequality for matrices), its proof is simply achieved by diagonalization and is left to the reader. We conclude that

$$\det(\mathrm{id} + Df(x)) \ge \left(1 + \left(\frac{V_B}{V_A}\right)^{1/n}\right)^n$$

and therefore

$$\operatorname{vol}(A+B) \ge \left(1 + \left(\frac{V_B}{V_A}\right)^{1/n}\right)^n V_A = \left(V_A^{1/n} + V_B^{1/n}\right)^n,$$

which is what we need. ∎

The most famous consequence of the Brunn–Minkowski theorem is the isoperimetric inequality. For an open subset $A \subset \mathbb{R}^n$ we want to define its *surface area* $\operatorname{vol}_{n-1}(\partial A)$. Several definitions are possible, but for A with piecewise smooth boundary the following formula (the Minkowski surface area) works:

$$\operatorname{vol}_{n-1}(\partial A) = \lim_{h\to+0} \frac{\operatorname{vol}(A+B_h) - \operatorname{vol} A}{h},$$

where B_h is the ball of radius h. Let v_n and s_n be the volume and the surface area of the unit ball B_1 in $\mathbb{R}^n$.

Theorem 11.2. *For reasonable A we have*

$$\frac{\operatorname{vol}_{n-1}(\partial A)}{s_n} \geq \left(\frac{\operatorname{vol} A}{v_n}\right)^{\frac{n-1}{n}}.$$

Proof. From the Brunn–Minkowski inequality we obtain:

$$\operatorname{vol}(A+B_h) \geq (\operatorname{vol} A^{1/n} + v_n^{1/n}h)^n$$

and therefore

$$\operatorname{vol}_{n-1}(\partial A) \geq n(\operatorname{vol} A)^{\frac{n-1}{n}} v_n^{1/n}.$$

Differentiating the equality $\operatorname{vol} B_h = v_n h^n$ we obtain $s_n = nv_n$ and therefore

$$\operatorname{vol}_{n-1}(\partial A) \geq \frac{s_n}{v_n}(\operatorname{vol} A)^{\frac{n-1}{n}} v_n^{1/n},$$

which is equivalent to the required inequality. □

Remark 11.3. A more careful argument can show that the equality is attained only if A is a ball.

For another approach to the Brunn–Minkowski inequality the reader is referred to the paper of M. Gromov [Grom90] and the post of T. Tao [Tao11]. That approach uses upper triangular Jacobi matrices in place of positive semidefinite matrices and turns out to be useful in several generalizations of the Brunn–Minkowski inequality.

Let us mention other consequences of the Brunn–Minkowski inequality:

Corollary 11.4. *Let A and B be centrally symmetric convex bodies in $\mathbb{R}^n$. Then the volume* $\operatorname{vol}(A+x)\cap B$ *is maximal if $x = 0$.*

Proof. Observe that

$$A \cap B \supseteq \frac{1}{2}((A+x)\cap B + (A-x)\cap B),$$

then apply the Brunn–Minkowsky inequality. □

Theorem 11.5 (The Rogers–Shepard inequality)**.** *For any convex A*

$$\operatorname{vol}(A-A) \leq \binom{2n}{n} \operatorname{vol} A.$$

Proof. Definitely, $A - A = \{a_1 - a_2 : a_1, a_2 \in A\}$ is a projection of $A \times A \subseteq \mathbb{R}^{2n}$ onto the $\mathbb{R}^n$, under the map $\pi : (a,b) \mapsto a - b$. The fibers of this projection over $x \in \mathbb{R}^n$ are the sets $(a_1, a_2) \in A \times A$ such that $a_1 - a_2 = x$, that is $a_1 = a_2 + x$. Hence up to translation the fiber over x is $A \cap (A + x)$.

When $A \cap (A + x_1)$ and $A \cap (A + x_2)$ are both nonempty then

$$A \cap \left(A + \frac{1}{2}(x_1 + x_2)\right) \supseteq \frac{1}{2}\left(A \cap (A + x_1) + A \cap (A + x_2)\right).$$

Hence $\operatorname{vol}(A\cap(A+x))^{1/n}$ is a concave function on $A-A$. If we introduce the norm $\|x\|$ with the unit ball $A-A$, then from concavity

$$\operatorname{vol}(A\cap(A+x)) \ge (1-\|x\|)^n \operatorname{vol} A.$$

Now integrate this over x to obtain:

$$\operatorname{vol}_{2n} A\times A = (\operatorname{vol} A)^2 = \int_{A-A} \operatorname{vol}(A\cap(A+x))\, dx \ge$$
$$\ge \operatorname{vol} A\cdot \int_{A-A} (1-\|x\|)^n\, dx = \operatorname{vol} A\cdot \int_{A-A}\int_0^{(1-\|x\|)^n} 1\, dy\, dx =$$
$$= \operatorname{vol} A\cdot \int_0^1 \int_{\|x\|\le 1-y^{1/n}} 1\, dx\, dy = \operatorname{vol} A\cdot \int_0^1 (1-y^{1/n})^n \operatorname{vol}(A-A)\, dy =$$
$$= \operatorname{vol} A\cdot \operatorname{vol}(A-A)\cdot \int_0^1 (1-y^{1/n})^n\, dy.$$

Let us calculate the last integral by substitution $y=t^n$ and using the beta-function:

$$\int_0^1 (1-y^{1/n})^n\, dy = \int_0^1 (1-t)^n n t^{n-1}\, dt = nB(n+1,n) = \frac{n\Gamma(n+1)\Gamma(n)}{\Gamma(2n+1)} =$$
$$= \frac{nn!(n-1)!}{(2n)!} = \binom{2n}{n}^{-1}.$$

Substituting this into the previous inequality we complete the proof. ☐

12. Log-concavity

Let us discuss logarithmically concave measures (abbreviated "log-concave") and their properties, see also [Ball04, Tao11]. A log-concave measure on $\mathbb{R}^n$ is a measure having density ρ such that $\log \rho(x)$ is a concave function, though there are more general definitions for measures having no density. Following the usual convention, we allow values $-\infty$ for concave functions, corresponding to 0 for log-concave functions.

The log-concavity is expressed by the inequality:

$$\rho((1-t)x_1 + tx_2) \ge \rho(x_1)^{1-t}\rho(x_2)^t. \tag{12.1}$$

Moreover, for continuous density ρ it is sufficient to check the case $t=1/2$, that is

$$\rho\left(\frac{x_1+x_2}{2}\right) \ge \sqrt{\rho(x_1)\rho(x_2)}.$$

The main result about log-concave measures is the Prékopa–Leindler inequality:

Theorem 12.1. *If a measure μ is log-concave and $\pi : \mathbb{R}^m \to \mathbb{R}^n$ is a linear surjection then $\pi_*\mu$ is log-concave.*

After the trivial observation that a Cartesian product of log-concave measures is log-concave we immediately obtain:

Corollary 12.2. *The convolution of two log-concave measures is log-concave.*

Proof of Theorem 12.1. It sufficient to consider the case, when π drops the dimension by 1 and use induction. Moreover, since the concavity property is essentially one-dimensional it suffices to consider the case of $n=1$ and $\pi : (x,y)\mapsto x$, where $y\in\mathbb{R}$.

Now we have to prove that

$$\int_{\mathbb{R}} \rho((1-t)x_1 + tx_2, y)\, dy \geq \left(\int_{\mathbb{R}} \rho(x_1, y)\, dy\right)^{1-t} \cdot \left(\int_{\mathbb{R}} \rho(x_2, y)\, dy\right)^{t}$$

under the assumption

$$\rho((1-t)x_1 + tx_2, (1-t)y_1 + ty_2) \geq \rho(x_1, y_1)^{1-t} \rho(x_2, y_2)^t.$$

Put for brevity

$$f(y) = \rho(x_1, y), \quad g(y) = \rho(x_2, y), \quad h(y) = \rho((1-t)x_1 + tx_2, y),$$

and also

$$F = \int_{\mathbb{R}} f(y)\, dy, \quad G = \int_{\mathbb{R}} g(y)\, dy, \quad H = \int_{\mathbb{R}} h(y)\, dy.$$

Consider the monotone map $\varphi : \mathbb{R} \to \mathbb{R}$ such that is transports $f(y)/F dy$ to $g(y)/G dy$, that is for all y

$$\frac{1}{F} \int_{-\infty}^{y} f(y)\, dy = \frac{1}{G} \int_{-\infty}^{\varphi(y)} g(y)\, dy.$$

It follows that $\varphi'(y) = \frac{f(y)G}{g(\varphi(y))F}$. As y runs from $-\infty$ to $+\infty$ the value $\varphi(y)$ does the same, therefore $\psi(y) = (1-t)y + t\varphi(y)$ also runs monotonically from $-\infty$ to $+\infty$. So we write

$$H = \int_{\mathbb{R}} h(\psi(y))\, d\psi(y) = \int_{\mathbb{R}} h(\psi(y)) \left(1 - t + \frac{tf(y)G}{g(\varphi(y))F}\right) dy.$$

Using the assumption $h(\psi(y)) \geq f(y)^{1-t} g(\varphi(y))^t$ we obtain:

$$H \geq F^{1-t} G^t \int_{\mathbb{R}} \left(\frac{f(y)G}{Fg(\varphi(y))}\right)^{1-t} \left((1-t)\frac{g(\varphi(y))}{G} + t\frac{f(y)}{F}\right) dy$$

and using the mean inequality $\left((1-t)\frac{g(\varphi(y))}{G} + t\frac{f(y)}{F}\right) \geq \left(\frac{f(y)}{F}\right)^t \cdot \left(\frac{g(\varphi(y))}{G}\right)^{1-t}$ we conclude

$$H \geq F^{1-t} G^t \int_{\mathbb{R}} \frac{f(y)}{F}\, dy = F^{1-t} G^t.$$

□

Now we make several observations. Every measure with constant density inside a convex body and zero density outside is log-concave by definition. So any its projection is also log-concave.

Then we start from two convex bodies $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^n$ and put the into $\mathbb{R}^{n+1}$ as $A' = A \times \{0\}$ and $B' = B \times \{1\}$. Note that the convex hull $\operatorname{conv} A' \cup B'$ contains the set $((1-t)A + tB) \times \{t\}$. Therefore by projection to the last coordinate and log-concavity we obtain:

$$\operatorname{vol}((1-t)A + tB) \geq \operatorname{vol} A^{1-t} \operatorname{vol} B^t,$$

this inequality is the dimension-independent version of the Brunn–Minkowski inequality. Indeed, it actually implies the standard Brunn–Minkowski inequality as follows: Replacing A and B with their homothetic copies $\frac{1}{1-t}A$ and $\frac{1}{t}B$ and using the homogeneity of the volume we rewrite:

$$\operatorname{vol}(A + B) \geq \frac{1}{(1-t)^{(1-t)n} t^{tn}} \operatorname{vol} A^{1-t} \operatorname{vol} B^t.$$

The reader is invited to check that after finding the maximum of the right hand side in t we arrive again at the Brunn–Minkowski inequality.

If fact, the inequality

$$\mu((1-t)A + tB) \geq \mu A^{1-t}\mu B^t \tag{12.2}$$

holds for arbitrary log-concave measure μ and convex bodies A and B, with almost the same proof as for $\mathrm{vol} = \mu$. This inequality (12.2) is sometimes considered as the definition of log-concavity, which is also applicable to measures with no density. The reader may check that for a measure with density the two definitions are equivalent.

Actually, the case of possibly non-convex A and B in the dimension-independent Brunn–Minkowski inequality is also valid, and is contained in the following version of the Prékopa–Leindler inequality (see also [Tao11] for generalizations to nilpotent Lie groups), also known as the functional Brunn–Minkowski inequality:

Theorem 12.3. *Assume f, g, h are nonnegative densities in $\mathbb{R}^n$ such that for some $t \in [0,1]$ and $x, y \in \mathbb{R}^n$*

$$h((1-t)x + ty) \geq f(x)^{1-t}g(y)^t.$$

Then

$$\int_{\mathbb{R}^n} h(x) \geq \left(\int_{\mathbb{R}^n} f(x)\,dx\right)^{1-t} \cdot \left(\int_{\mathbb{R}^n} g(y)\,dy\right)^t.$$

Sketch of the proof. One way of proving this theorem is to show that the property $h((1-t)x+ty) \geq f(x)^{1-t}g(y)^t$ is preserved under projections dropping dimension by 1, like we did in the proof of Theorem 12.1.

Another approach is to consider a monotone map φ, sending the measure $f(x)dx$ to $g(x)dx$, and then consider the monotone map $\psi : x \mapsto (1-t)x + t\varphi(x)$ and write inequalities for $h((1-t)x + t\varphi(x))$, similarly to the proof of Theorem 12.1. □

Another result that benefits from log-concavity is the Minkowski theorem on facet areas. We split it into the simple direct theorem and the harder converse theorem:

Theorem 12.4. *If $P \subset \mathbb{R}^n$ is a polytope, $\nu_1, \dots, \nu_m$ are its facet normals, and $A_1, \dots, A_m$ are its respective facet surface areas, then*

$$A_1\nu_1 + A_2\nu_2 + \dots + A_m\nu_m = 0. \tag{12.3}$$

Proof. Consider a constant vector field v, it has zero divergence and by the Gauss theorem its flux through ∂P is zero. That is,

$$A_1(\nu_1, v) + A_2(\nu_2, v) + \dots + A_m(\nu_m, v) = 0.$$

Since v is arbitrary, the result follows. □

Theorem 12.5. *For any prescribed set of normals $\nu_1, \dots, \nu_m$ spanning $\mathbb{R}^n$ and positive facet surface areas $A_1, \dots, A_m$ satisfying (12.3), there exists a unique (up to translations) polytope P corresponding to these data.*

Sketch of the proof. Introduce real variables $t_1, \dots, t_m$ and consider the polytope

$$P(t) = \{x \in \mathbb{R}^n : \forall i = 1, \dots, m\ (x, \nu_i) \leq t_i\}. \tag{12.4}$$

Since a positive combination of ν_i's equals zero, then $P(t)$ is always bounded. The equation (12.4) may treat x_j's and t_i's as variables and define an $(n+m)$-dimensional polyhedron $\widetilde{P}$, this way $P(t)$ is a parameterized n-dimensional section of $\widetilde{P}$.

By the Prékopa–Leindler inequality the function $f(t) = \mathrm{vol}\, P(t)$ is log-concave in t. The standard geometric differentiation reasoning shows that its logarithmic derivative equals

$$d \log f(t) = \frac{1}{f(t)} \left(A_1(t)dt_1 + \dots + A_m(t)dt_m\right),$$

where $A_i(t)$ is the surface area of the corresponding facet. The map $\varphi(t) = d\log f(t)$ is therefore "minus" monotone and its image must be convex.

Definitely, the image of $\varphi(t)$ satisfies (12.3) (divided by $f(t) = \operatorname{vol} P$). We have to check that for every positive vector y satisfying

$$y_1\nu_1 + y_2\nu_2 + \cdots + y_n\nu_n = 0 \tag{12.5}$$

there exists a proportional vector in the image of φ. Denote the set of nonnegative y satisfying (12.5) by Q. The polyhedron Q is a cone centered in 0 and from convexity of the image of φ and the linear programming considerations to prove what we need it is sufficient to check that every extremal ray of Q intersects the image of φ. But such an extremal ray corresponds to a vector y with at most $n+1$ positive coordinates, put $I = \{i : y_i > 0\}$.

It may happen that I contains less than n indexes. We avoid such degenerate cases by perturbing slightly the set of normals $\nu_1, \dots, \nu_m$, the general case then can be deduced by going to the limit. So it is sufficient to consider the case $|I| = n+1$. It is easy to check that such vectors y correspond to simplices $\{x : \forall i \in I\ (x, \nu_i) \le t_i\}$ that are obviously contained in the image of φ. Hence all nonnegative vectors y satisfying (12.5) are in the image of φ up to scaling.

The uniqueness of P follows from the fact that the image of φ is essentially $(m-n)$-dimensional and φ is monotone, hence the preimage $\varphi^{-1}(y)$ is n dimensional and consists of a set of polytopes taken to each other by translations. ∎

In this section we presented the analytical approach to log-concavity, and the reader is invited to read the review of R.P. Stanley [Stan89] about algebraic point of view on log-concavity. One simple fact, left as an exercise for the reader, is: If a univariate monic polynomial $P(x)$ has all roots real and negative, then its coefficients from a log-concave sequence.

In a wonderful way, the algebraic approach returns with a proof of the Alexandrov–Fenchel inequality for mixed volumes of convex bodies, which, in turn, gives another proof of the Brunn–Minkowski inequality. A typical example of an algebraic log-concavity result is:

Theorem 12.6. *Let X be an n-dimensional normal projective algebraic variety and L and M be ample divisor classes on X. Then the sequence $v_0, \dots, v_n$ of intersection numbers*

$$v_k = \#(\underbrace{L, \dots, L}_{k}, \underbrace{M, \dots, M}_{n-k})$$

is log-concave.

Sketch of the proof. For corresponding facts from algebraic geometry please consult the textbook [GH78].

We have to prove the inequality:

$$\#(\underbrace{L, \dots, L}_{k}, \underbrace{M, \dots, M}_{n-k})^2 \ge \#(\underbrace{L, \dots, L}_{k-1}, \underbrace{M, \dots, M}_{n-k+1}) \cdot \#(\underbrace{L, \dots, L}_{k+1}, \underbrace{M, \dots, M}_{n-k-1}). \tag{12.6}$$

After multiplying L and M by some positive integers we assume that L and M correspond to projective embeddings of X. By an appropriate version of the Bertini theorem we choose the hyperplane sections (in L) for the first $k-1$ of L in the intersection formula and the hyperplane sections (in M) for the last $n-k-1$ of M, so that X reduces to a normal surface S and over this surface we have to prove:

$$\#(L, M)^2 \ge \#(L, L) \cdot \#(M, M).$$

By the Hodge theorem, the intersection form of an algebraic surface has positive index 1, then this inequality is just the inverse Cauchy–Schwarz inequality. □

The general form of (12.6) is

$$\#(L_1, L_2, L_3, \dots, L_n)^2 \geq \#(L_1, L_1, L_3, \dots, L_n) \cdot \#(L_2, L_2, L_3, \dots, L_n), \tag{12.7}$$

which is the algebraic form of the Alexandrov–Fenchel inequality.

13. Mixed volumes

Let us make a brief discussion of mixed volumes, following [Grom90]. The crucial fact is:

Theorem 13.1. *Let $K_1, \dots, K_n$ be convex bodies in $\mathbb{R}^n$, the expression*

$$\operatorname{vol}(t_1K_1 + \dots + t_nK_n)$$

for nonnegative t_i is a polynomial in t_i's of degree n and the coefficient at $t_1 \dots t_n$, divided by $n!$, is called the mixed volume *of $K_1, \dots, K_n$, and is denoted by* $\mathrm{MV}(K_1, \dots, K_n)$.

Proof. Consider the monotone maps f_i from $\mathbb{R}^n$ to the respective K_i with potentials

$$u_i(k) = \log \int_K e^{\langle k,x\rangle}\, d\mu_i(x),$$

where μ_i are some measures with convex hulls of support equal to the respective K_i. The map

$$f(k) = t_1f_1(k) + \dots t_nf_n(k)$$

is also monotone, with potential $u(k) = t_1u_1(k) + \dots + t_nu_n(k)$. It is a simple fact about convex functions that for continuously differentiable $u(k)$ that image of the differential map $du(k) = f(k)$ is convex. Therefore the image $\widetilde{K}$ of $f(k)$ is convex, obviously $\widetilde{K} \subseteq K = t_1K_1 + \dots + t_nK_n$.

We claim that (the open convex set) $\widetilde{K}$ actually equals K up to boundary. To prove this, it is sufficient to compare their support functions:

$$h(p, K) = \sup_{x\in K}\langle p, x\rangle, \quad h(p, \widetilde{K}) = \sup_{x\in \widetilde{K}}\langle p, x\rangle.$$

Take $k = \alpha p$, it is easy to see that for $\alpha \to +\infty$ the measure $e^{\langle k,x\rangle}\mu_i(x)$ on K_i gets concentrated near the points of K_i, where the linear form $\langle p, x\rangle$ attains its maximum. Hence $\langle p, f_i(\alpha p)\rangle \to h(p, K_i)$ for $\alpha \to +\infty$. Since the support functions are obviously additive with respect to the Minkowski sum, we obtain

$$\langle p, f(\alpha p)\rangle \to h(p, K),$$

when $\alpha \to +\infty$. Since $h(p, \widetilde{K}) \geq \langle p, f(\alpha p)\rangle$ by definition, there must be an equality $h(p, \widetilde{K}) = h(p, K)$.

Now we can calculate $\operatorname{vol} \widetilde{K} = \operatorname{vol} K$:

$$\operatorname{vol}(t_1K_1 + \dots + t_nK_n) = \int_{\mathbb{R}^n} \det(t_1Df_1(k) + \dots + t_nDf_n(k))\, dk. \tag{13.1}$$

Obviously, the determinant under the integral is a polynomial of degree n, and the result follows. □

Actually, the mixed volume $\mathrm{MV}(K_1, \dots, K_n)$ is positive, the reader may deduce it from (13.1) noting that Df_i's are positive definite. In fact, a stronger fact, the Alexandrov–Fenchel inequality, holds:

Theorem 13.2. *For any convex bodies $K_1, \dots, K_n \subset \mathbb{R}^n$ we have*

$$\operatorname{MV}(K_1, K_2, \dots, K_n)^2 \ge \operatorname{MV}(K_1, K_1, K_3, \dots, K_n) \cdot \operatorname{MV}(K_2, K_2, K_3, \dots, K_n).$$

It seems that there is no easy way to prove this inequality. One way is to relate the mixed volumes to intersection numbers of ample divisors over a *toric* variety, and then prove (12.7) arguing similarly to the proof of Theorem 12.6, see [Ful93]. The other way is to prove an analogue of this inequality for positive definite matrices and their *mixed discriminants*, and then use (13.1) along with other nontrivial observations, see [Grom90].

It makes sense to mention the Bernstein theorem [Bern76], connecting the mixed volumes and intersection numbers of divisors. First, for every Laurent polynomial in n variables

$$P(\bar{x}) = \sum_{\bar{k}} c_{\bar{k}} \bar{x}^{\bar{k}},$$

where we use the notation $\bar{x}^{\bar{k}} = x_1^{k_1} \dots x_n^{k_n}$, we define the Newton polytope:

$$N(P) = \operatorname{conv}\{\bar{k} \in \mathbb{Z}^n : c_{\bar{k}} \ne 0\}.$$

Now the theorem reads:

Theorem 13.3. *The system of equations:*

$$\begin{aligned} P_1(\bar{x}) &= 0 \\ P_2(\bar{x}) &= 0 \\ &\dots \\ P_n(\bar{x}) &= 0, \end{aligned}$$

for $\bar{x} \in (\mathbb{C}^*)^n$, has either an infinite number of solutions, or a finite number of solutions not exceeding $n! \operatorname{MV}(N(P_1), \dots, N(P_n))$. For a generic choice of the coefficients $c_{\bar{k}}$ of the polynomials keeping the Newton polytopes the same, the number of solutions is precisely $n! \operatorname{MV}(N(P_1), \dots, N(P_n))$.

In [Bern76], besides this fact, a certain sufficient condition, in terms of faces of the Newton polytopes, was given that guarantees that the set of solution consists of precisely $n! \operatorname{MV}(N(P_1), \dots, N(P_n))$ points, without using the term *generic*.

We do not give a proof of this theorem here. An elementary, but technical, reasoning can be found in the original paper [Bern76]. A more conceptual approach using symplectic/Kähler geometry can be found in [Ati83].

14. The Blaschke–Santaló inequality

We give an application of the Prékopa–Leindler inequality (in form of Theorem 12.3) to the Blaschke–Santaló inequality. Following the paper of Lehec [Leh09A], we start from proving the functional version of this inequality:

Theorem 14.1. *Assume f and g are nonnegative measure densities such that*

$$f(x)g(y) \le e^{-(x,y)}$$

for any $x, y \in \mathbb{R}^n$. Also assume that $\int_{\mathbb{R}^n} y g(y)\, dy = 0$ and $\int_{\mathbb{R}^n} x f(x)\, dx$ converges. Then

$$\int_{\mathbb{R}^n} f(x)\, dx \cdot \int_{\mathbb{R}^n} g(y)\, dy \le (2\pi)^n.$$

We are going to make the proof in several steps. First, we observe that it is sufficient to prove the following:

Lemma 14.2. *For any nonnegative measure density $f(x)$ with converging $\int_{\mathbb{R}^n} xf(x)\,dx$ there exists $z \in \mathbb{R}^n$ such that for any other nonegative density $g(x)$ with converging $\int_{\mathbb{R}^n} xg(x)\,dx$ the inequality*

$$f(x+z)g(y) \le e^{-(x,y)}$$

for any $x, y \in \mathbb{R}^n$ implies

$$\int_{\mathbb{R}^n} f(x)\,dx \cdot \int_{\mathbb{R}^n} g(y)\,dy \le (2\pi)^n.$$

Informally, the lemma replaces the mass centering condition by an arbitrary centering condition for one of the functions, independent of the other function.

Proof of Theorem 14.1 assuming Lemma 14.2. Suppose we have an appropriate z from Lemma 14.2. From the assumption of Theorem 14.1 it follows that:

$$f(z+x)g(y)e^{(z,y)} \le e^{-(z,y)-(x,y)+(y,z)} = e^{-(x,y)}.$$

Hence

$$\int_{\mathbb{R}^n} f(x)\,dx \cdot \int_{\mathbb{R}^n} g(y)e^{(z,y)}\,dy \le (2\pi)^n.$$

Let us estimate $e^{(y,z)} \ge 1 + (y,z)$ and integrate this inequality after multiplying by $g(y)$:

$$\int_{\mathbb{R}^n} g(y) + (y,z)g(y)\,dy \le \int_{\mathbb{R}^n} g(y)e^{(z,y)}\,dy.$$

Taking into account that $\int_{\mathbb{R}^n} yg(y)\,dy = 0$ we obtain

$$\int_{\mathbb{R}^n} g(y)\,dy \le \int_{\mathbb{R}^n} g(y)e^{(z,y)}\,dy,$$

which implies the required inequality. □

In order to prove Lemma 14.2 we start with proving the corresponding one-dimensional fact:

Lemma 14.3. *Let $f(x)$ and $g(y)$ be nonnegative densities on the half-line $\mathbb{R}_+$. If $f(x)g(y) \le e^{-xy}$ for any $x, y \in \mathbb{R}_+$ then*

$$\int_0^{+\infty} f(x)\,dx \cdot \int_0^{+\infty} g(y)\,dy \le \frac{\pi}{2}.$$

Proof. Put $u(s) = f(e^s)e^s$, $v(t) = g(e^t)e^t$, and $w(r) = e^{-e^{2r}/2}e^r$. From the assumptions it follows that for any $s, t \in \mathbb{R}$

$$w\left(\frac{s+t}{2}\right) = e^{-e^{s+t}/2}e^{(s+t)/2} \ge \sqrt{f(e^s)g(e^t)}e^{(s+t)/2} = \sqrt{u(s)v(t)}.$$

Now the one-dimensional case of Theorem 12.3 implies:

$$\int_0^{+\infty} f(x)\,dx \cdot \int_0^{+\infty} g(y)\,dy = \int_{-\infty}^{+\infty} u(s)\,ds \cdot \int_{-\infty}^{+\infty} v(t)\,dt \le$$
$$\le \left(\int_{-\infty}^{+\infty} w(r)\,dr\right)^2 = \left(\int_0^{+\infty} e^{-z^2/2}\,dz\right)^2 = \frac{\pi}{2}.$$

□

Proof of Lemma 14.2. The proof is by induction. Assume the normalization $\int_{\mathbb{R}^n} f(x)\,dx = 1$, which can be achieved by multiplying $f(x)$ by a constant and dividing $g(y)$ by the same constant.

The one-dimensional case follows from Lemma 14.3 as follows: choose z to be the median of f. After a shift of z to 0 we have to prove that the inequality

$$f(x)g(y) \le e^{-xy}$$

implies

$$\int_{-\infty}^{+\infty} g(y)\,dy \le 2\pi.$$

But Lemma 14.3 is applicable to $f(x)$ and $g(y)$ in the range $x, y > 0$ and to $f(-x)$ and $g(-y)$ in the same range. Summing up the results and noting that

$$\int_{-\infty}^{0} f(x)\,dx = \int_{0}^{+\infty} f(x)\,dx = 1/2$$

we obtain the required inequality.

Now we use induction as follows. Let μ be the measure with density $f(x)$. We can partition μ into equal halves H^+ and H^- with a hyperplane H orthogonal to the last basis vector e_n. Since we may translate f, without loss of generality we assume that H^+ and H^- are defined by $x_n \ge 0$ and $x_n \le 0$ respectively.

Let c^+ and c^- be mass centers of $\mu|_{H^+}$ and $\mu|_{H^-}$. After another translation of f we also assume that the segment $[c^-, c^+]$ intersects H at the origin. Now let v be the vector parallel to $[c^-, c^+]$ and normalized so that $(v, e_n) = 1$. Then we consider the linear operators A and B defined by:

$$Ae_1 = e_1, \dots, Ae_{n-1} = e_{n-1}, Ae_n = v, \quad B = (A^{-1})^T,$$

so that $(Ax, By) = (x, y)$ for any $x, y \in \mathbb{R}^n$. Note that the determinants of these maps equal 1 and A maps H to itself, while B may not map H to itself. Consider the functions of $x', y' \in H$:

$$F(x') = \int_0^{+\infty} f(x' + sv)\,ds, \quad G(y') = \int_0^{+\infty} g(Bx' + te_n)\,dt.$$

From the normalization, $\int_H F(x')\,dx = 1/2$. Also, the mass center of $F(x')$ is zero by the selection of vector v. The assumption can be rewritten:

$$f(x' + sv)g(By' + te_n) \le e^{-(x'+sv, By'+te_n)} = e^{-(Ax'+sAe_n, By'+te_n)} = \\ = e^{-(x',y')-(x',te_n)-(sAe_n,By')-st(v,e_n)} = e^{-(x',y')-st}.$$

We now fix x' and y' and consider $f(x'+sv)$ and $g(By'+te_n)e^{(x',y')}$ as functions of positive variables s and t. By Lemma 14.3 we obtain:

$$F(x')G(y') \le \frac{\pi}{2} e^{-(x',y')}.$$

Now we invoke the inductive assumptions interchanging $F(x')$ and $G(y')$, taking into account Lemma 14.2, the fact $\int_H x'F(x')\,dx' = 0$ implies that

$$\int_H F(x')\,dx' \cdot \int_H G(y')\,dy' \le \frac{\pi}{2}(2\pi)^{n-1}.$$

Since $\int_H F(x')\,dx = 1/2$ we obtain

$$\int_{BH^+} g(y)\,dy = \int_H G(y')\,dy' \le \pi(2\pi)^{n-1}.$$

Similarly, inverting e_n and v, we obtain

$$\int_{BH^-} g(y)\,dy \le \pi(2\pi)^{n-1}$$

and it remains to sum these inequalities to obtain

$$\int_{\mathbb{R}^n} g(y)\,dy \le (2\pi)^n.$$

□

Remark 14.4. A finer argument shows that it is possible to satisfy the assumption $[c_-, c_+] \parallel e_n$, after a careful choice of the coordinate system. In this case $v = e_n$, the operators A, B equal the identity operator, and the formulas get considerably simpler. This assumption is satisfied if we choose the hyperplane $H = \{x_n = 0\}$ so that to minimize the integral

$$\int_{\mathbb{R}^n} \operatorname{dist}(x, H) f(x)\,dx.$$

By varying the normal of H and its constant term we obtain that

$$\int_{H^+} x f(x)\,dx - \int_{H^-} x f(x)\,dx \perp H, \quad \text{and} \quad \int_{H^+} f(x)\,dx - \int_{H^-} f(x)\,dx = 0,$$

which is exactly what we need.

Another approach (see [Leh09B]) to Lemma 14.2 invokes the Yao–Yao theorem (Theorem 4.2) to partition $\mathbb{R}^n$ into 2^n convex cones $A_1, \dots, A_{2^n}$ with center at z so that $\int_{A_i} f(x)\,dx = 2^{-n}$ for every i. After translating z to the origin one observes that the characterizing property of the Yao–Yao partition means that the space $\mathbb{R}^n$ is covered by the family $\{A_i^\circ\}_{i=1}^{2^n}$ of polar cones.

Then one invokes the logarithmic form of the Prékopa–Leindler inequality:

Lemma 14.5. *Let f, g, h be nonnegative absolute integrable function on the positive cone $\mathbb{R}^n_+$. Assume that*

$$h(\sqrt{x_1 y_1}, \dots, \sqrt{x_n y_n}) \ge \sqrt{f(x) g(y)}$$

for any $x, y \in \mathbb{R}^n_+$. Then

$$\int_{\mathbb{R}^n_+} f(x)\,dx \cdot \int_{\mathbb{R}^n_+} g(y)\,dy \le \left(\int_{\mathbb{R}^n_+} h(z)\,dz \right)^2.$$

Proof. Substitute

$$\begin{aligned}
\bar f(s_1, \dots, s_n) &= f(e^{s_1}, \dots, e^{s_n}) e^{s_1 + \dots + s_n},\\
\bar g(t_1, \dots, t_n) &= g(e^{t_1}, \dots, e^{t_n}) e^{t_1 + \dots + t_n},\\
\bar h(r_1, \dots, r_n) &= h(e^{r_1}, \dots, e^{r_n}) e^{r_1 + \dots + r_n}.
\end{aligned}$$

It is easy to check that for any $s, t \in \mathbb{R}^n$

$$\left(\bar h \left(\frac{s+t}{2} \right) \right)^2 \ge \bar f(s) \bar g(t).$$

Then Theorem 12.3 implies that

$$\int_{\mathbb{R}^n} \bar f(s)\,ds \cdot \int_{\mathbb{R}^n} \bar g(t)\,dt \le \left(\int_{\mathbb{R}^n} \bar h(r)\,dr \right)^2,$$

that is equivalent to what we need. □

Now Lemma 14.5 applied to $h(z) = e^{-|z|^2/2}$ gives for any Yao–Yao cone A_i the inequality:

$$\int_{-A_i^\circ} g(y)\ dy \le \pi^n.$$

It remains to sum over $i = 1, \dots, 2^n$ to prove Lemma 14.2.

Now we deduce the classical form of the Blaschke–Santaló inequality for convex bodies:

Corollary 14.6. *Let K be a convex body in $\mathbb{R}^n$ with mass center at the origin, at let*

$$K^\circ = \{y \in \mathbb{R}^n : \forall x \in K\ (x, y) \le 1\}$$

be its polar body. Then

$$\operatorname{vol} K \cdot \operatorname{vol} K^\circ \le v_n^2,$$

where v_n is the volume of the unit ball in $\mathbb{R}^n$.

Proof. Consider the corresponding "norms" (note that $\|x\|$ may not coincide with $\|-x\|$):

$$\|x\| = \min\{r \ge 0 : x \in rK\}, \quad \|x\|_\circ = \min\{r \ge 0 : x \in rK^\circ\}.$$

The definition of the polar body means that for any $x, y \in \mathbb{R}^n$

$$(x, y) \le \|x\| \cdot \|y\|_\circ.$$

Now we introduce two functions

$$f(x) = e^{-\|x\|^2/2}, \quad g(y) = e^{-\|y\|_\circ^2/2}$$

and check that

$$f(x)g(y) = e^{-\|x\|^2/2 - \|y\|_\circ^2/2} \le e^{-(x,y)}.$$

Also, $f(x)$ has mass center at the origin. Hence, by Theorem 14.1 the product of their integrals is at most $(2\pi)^n$. Now we calculate by changing the integration order:

$$\int_{\mathbb{R}^n} f(x)\ dx = \int_{\mathbb{R}^n} \int_0^{f(x)} 1\ dydx = \int_0^1 \operatorname{vol} K(-2\log y)^{n/2}\ dy = c_n \operatorname{vol} K$$

for the constant $c_n = \int_0^1 (-2\log y)^{n/2}\ dy$. The same holds for $g(y)$:

$$\int_{\mathbb{R}^n} g(y)\ dy = c_n \operatorname{vol} K^\circ.$$

It remains to calculate c_n, this is simple if we take the unit ball in place of K and the Euclidean norm $|x|$ in place of $\|x\|$:

$$(2\pi)^{n/2} = \int_{\mathbb{R}^n} e^{-|x|^2/2}\ dx = c_n v_n.$$

Hence

$$\operatorname{vol} K \cdot \operatorname{vol} K^\circ \le \frac{(2\pi)^n}{c_n^2} = v_n^2.$$

□

It is a famous problem (the Mahler conjecture) to establish the lower bound for centrally symmetric convex bodies:

$$\operatorname{vol} K \cdot \operatorname{vol} K^\circ \ge \frac{4^n}{n!},$$

which is an equality for the cube and the crosspolytope. Some known partial results on this conjecture are summarized in the blog post by Tao [Tao07]. The best known result is established by Greg Kuperberg [Ku08]:

$$\operatorname{vol} K \cdot \operatorname{vol} K^\circ \ge \frac{\pi^n}{n!}. \tag{14.1}$$

15. Needle decomposition

Now we are going to consider an interesting tool in studying inequalities for measures, the needle decomposition. The reader is referred to [NSV02] for a deeper review of this subject.

The main result is the following theorem:

Theorem 15.1. *For any two absolutely continuous finite measures μ and ν on $\mathbb{R}^n$ supported in a bounded convex set S and a prescribed $\varepsilon > 0$ it is possible to partition S into some number of convex pieces $P_1, \dots, P_N$ so that for every P_i*

$$\frac{\mu(P_i)}{\mu(\mathbb{R}^n)} = \frac{\nu(P_i)}{\nu(\mathbb{R}^n)}$$

and one of the alternatives hold: Either P_i can be included into a ε-neighborhood of a line, or not. But the total measure $\mu + \nu$ of the parts satisfying the latter alternative is less than ε.

A piece P_i in this theorem has small deviation from a line, so it looks like an almost 1-dimensional "needle". This justifies the name "needle decomposition".

Sketch of the proof. We start from the case $n = 2$. We can partition the two measures with a line into equal halves. Then we can partition every part with another line so that both measures are partitioned into equal fourths. Then we proceed this way many times. Is we impose an additional technical assumption, say, that the sum of densities of μ and ν is separated from zero on S, then it is easy to check that after an appropriate number of steps all the parts P_i become ε-needles (ε-close to lines).

The technical assumption on positive density can be avoided as follows: Take some positive δ and break the parts P_i into two categories: Those for which $\mu(P_i) + \nu(P_i) > \delta \operatorname{vol} P_i$, the "essential" parts and all the other "inessential" parts. For the essential parts after a definite number of steps, depending on ε and δ, we conclude that they are ε-needles. For the inessential parts we observe that their total measure $\mu + \nu$ is at most $\delta \operatorname{vol} S$, which can be chosen to be less than ε.

Now consider the case $n > 2$. For an $(n-2)$-dimensional linear subspace $L \subset \mathbb{R}^n$ we make the 2-dimensional partition of the 2-plane $\mathbb{R}^n/L$ into parts so that the essential parts are ε/n-needles and the total measure of the inessential parts is small. The corresponding partition in $\mathbb{R}^n$ is a partition with hyperplanes parallel to L.

Then the trick is to repeat this construction for different L's passing sufficiently close to any given $(n-2)$-dimensional direction, also requiring that the total measure of the inessential parts on the j'th step is at most $\varepsilon/2^j$. The measures μ and ν get partitioned into many equal pieces and it is possible to check that most of the parts cannot be essentially two-dimensional (contain a two-dimensional disk of a prescribed size), while the inessential parts together have total measure $\mu+\nu$ less than ε. Indeed, if some essential part P_i has a 2-dimensional disk D of radius ε/n inside, then we take the orthogonal to the disk linear subspace $L_D \subset \mathbb{R}^n$ and observe the following: Since we made hyperplane cuts almost parallel to L_D on some stage, this disk could not survive for the essential parts. □

The useful observation is as follows. For every ε-needle part P_i we choose the line ℓ_i, to which it is ε-close. If the measures μ were log-concave then their restrictions μ_i and ν_i to the convex body P_i remain log-concave. Since P_i is almost one-dimensional then it makes sense to consider the projections of μ_i and ν_i to ℓ_i, which become log-concave measures in the line, by Theorem 12.1. And if the densities of μ and ν were continuous, then this measure on the line approximates the original measures on P_i with an appropriate precision. Of course, by approximating measures the assumption of continuity for the density can be dropped in most practical applications. To put it short, the needle decomposition can sometimes reduce questions about pairs of log-concave measures on $\mathbb{R}^n$ to pairs of log-concave measures in the line.

There is a more sophisticated version of the needle decompositions, called *pancake* decomposition used, for example, in [Grom03]. Informally, when the number of measures to partition increases then the "essential dimension" of the parts increases accordingly.

16. Isoperimetry for the Gaussian measure

Let us apply the needle decomposition to establish the isoperimetric inequality for a Gaussian measure on $\mathbb{R}^n$. After some rescaling any such measure obtains the density $e^{-\pi|x|^2}$, which we like for its simplicity and normalization $\int_{\mathbb{R}^n} e^{-\pi|x|^2}\,dx = 1$. The crucial property of the Gaussian measure is that the density of its orthogonal projection is proportional to restriction of its density to any line.

Since the notion of the surface area does not make much sense for Gaussian measures we formulate the "integrated" form of the isoperimetric inequality:

Theorem 16.1. *For a Gaussian measure μ on $\mathbb{R}^n$ consider an open subset U and a halfspace H of the same measure $\mu(H) = \mu(U)$. Then for their ε-neighborhoods we have:*

$$\mu(U_\varepsilon) \geq \mu(H_\varepsilon).$$

Remark 16.2. Note that the value $\mu(H_\varepsilon)$ depends on $\mu(H)$ and ε and *does not depend* on the dimension n.

So we see that for the standard measure in $\mathbb{R}^n$ the "most round" body is the ball, and for the Gaussian measure the "most round" body is the halfspace.

Sketch of the proof. Let $\bar{U} = \mathbb{R}^n \setminus U$ be the complement of U. After making the needle decomposition for two restrictions $\mu|_U$ and $\mu|_{\bar{U}}$ we reduce the problem to a one-dimensional problem, preserving the equality

$$\frac{\mu(U \cap P_i)}{\mu(P_i)} = \mu(U) = \mu(H).$$

The measure μ on P_i, after projection to ℓ_i, becomes a log-concave measure with density ρ. Moreover, it is *strongly log-concave* in the following sense:

$$\rho((1-t)x_1 + tx_2) \geq \rho(x_1)^{1-t}\rho(x_2)^t e^{\pi t(1-t)|x_1-x_2|^2}.$$

It is easy to check that the proof of Theorem 12.1 can be modified to prove that the strong log-concavity is preserved under projections.

Now everything reduces to the following:

Lemma 16.3. *Let μ be the probability measure on the line with density $e^{-\pi x^2}$ and ν be a strongly log-concave probability measure on the line. Consider an open subset U and a halfline H such that $\mu(H) = \nu(U)$. Then for their ε-neighborhoods we have:*

$$\nu(U_\varepsilon) \geq \mu(H_\varepsilon).$$

The proof of this lemma is left as a technical exercise for the reader. It makes sense to approximate U with a union of intervals and then show that these intervals can be moved to the left or to the right decreasing $(\nu(U_\varepsilon))'_\varepsilon$. Finally all the intervals can be merged into a single one either $(-\infty, t)$ or $(t, +\infty)$. In this case it is easy to verify that the value $(\nu(U_\varepsilon))'_\varepsilon = \rho(t)$ is at least $e^{-\pi x^2}$, where

$$\nu(U) = \int_{-\infty}^{x} e^{-\pi\xi^2}\,d\xi.$$

□

17. Isoperimetry and concentration on the sphere

It is interesting and important that on the round sphere $S^n \subseteq \mathbb{R}^{n+1}$ the corresponding version of Theorem 16.1 also holds:

Theorem 17.1. *For a the uniform probability measure σ on S^n consider an open subset U and a halfspace $H \subset \mathbb{R}^{n+1}$ such that $\sigma(H \cap S^n) = \sigma(U)$. Then for their ε-neighborhoods (in the geodesic path metric of S^n) we have:*

$$\sigma(U_\varepsilon) \geq \sigma((H \cap S^n)_\varepsilon).$$

This means, in particular, that the minimal $(n-1)$-dimensional volume of ∂U for given $\sigma(U)$ is attained at spherical caps like $H \cap S^n$. We do not give a proof here because it depends on some deep results in Riemannian geometry, see the appendix to [MS86] by M. Gromov, for example. The crucial notion here is the Ricci curvature of the Riemannian metric, which equals $n-1$ for the unit n-dimensional round sphere.

Another heuristic evidence for Theorem 17.1 is that, for large dimensions n, the Gaussian measure with density $e^{-\pi|x|^2}$ on $\mathbb{R}^n$ gets concentrated near the round sphere of radius $\sqrt{n/\pi}$. After rescaling by $\sqrt{\pi/n}$ we see that it "approaches" the measure σ and therefore the isoperimetric inequality for the sphere is "asymptotically" correct.

The isoperimetric inequality has the following famous consequence, known as the *concentration of measure phenomenon* on the sphere:

Theorem 17.2. *Let $U \subset S^n$ have measure $\sigma(U) \geq 1/2$. Then the measure of the neighborhood $\sigma(U_\varepsilon)$ becomes almost 1 for ε of order $\sqrt{n}$, in particular the following estimate holds:*

$$\sigma(U_\varepsilon) \geq 1 - e^{-\frac{(n-1)\varepsilon^2}{2}}.$$

In [Led01] it is shown that this theorem can be deduced from the isoperimetry of the Gaussian measure using some analytical technicalities. Following [GM2001], we give a simple proof of a weaker fact:

Proof of a weaker assertion. Consider the complement $V = S^n \setminus U_\varepsilon$, note that the spherical distance between U and V is at least ε. Now we pass to subsets of the unit ball B^{n+1} defined as follows:

$$U_0 = \{x \in B^{n+1} : x/|x| \in U\}, \quad V_0 = \{x \in B^{n+1} : x/|x| \in V\},$$

for their volumes we have ($v = v_{n+1}$ is the volume of the unit ball here):

$$\operatorname{vol} U_0 = v\sigma(U), \quad \operatorname{vol} V_0 = v\sigma(V).$$

Consider the set $X = \frac{1}{2}(U_0 + V_0)$, simple trigonometry shows that X consists of vectors with lengths at most $\cos\frac{\varepsilon}{2}$, and therefore

$$\operatorname{vol} X \leq v\cos^{n+1}\frac{\varepsilon}{2} = v\left(1 - \sin^2\frac{\varepsilon}{2}\right)^{\frac{n+1}{2}} \leq ve^{-\frac{(n+1)\sin^2\frac{\varepsilon}{2}}{2}}.$$

The Brunn–Minkowski inequality, in particular, gives $\operatorname{vol} V_0 \leq \operatorname{vol} X$, and therefore

$$v\sigma(V) \leq ve^{-\frac{(n+1)\sin^2\frac{\varepsilon}{2}}{2}},$$

which implies

$$\sigma(U_\varepsilon) \geq 1 - e^{-\frac{(n+1)\sin^2\frac{\varepsilon}{2}}{2}}.$$

□

A usual way to use the concentration on the sphere is the following (Lévy's lemma):

Corollary 17.3. *Let f be a 1-Lipschitz ($|f(x) - f(y)| \le \operatorname{dist}(x, y)$) function on S^n, then the measure of the set where $f(x)$ differs from its median value M_f is estimated by*

$$\sigma\{x : |f(x) - M_f| \ge \varepsilon\} \le 2e^{-\frac{(n-1)\varepsilon^2}{2}}.$$

Informally, on most of the sphere the function f differs from M_f by at most $O(\frac{1}{\sqrt{n}})$.

Much more information about the concentration phenomenon on the sphere and many other metric-measure spaces can be found in the book [Led01].

18. More remarks on isoperimetry

There are many other instances of the isoperimetric inequality. Here we just give a brief sketch of them, for more details see the book [Lub94].

First, the isoperimetric inequality on Riemannian manifolds is connected through the Cheeger–Buser inequalities to the smallest positive eigenvalue of the Laplace operator (the Laplacian). Let us give some explanations without proofs. The Laplace operator for differential forms on a smooth closed manifold M is defined as

$$\Delta = dd^* + d^*d,$$

where d^* is the formal adjoint to the d operator on differential forms. It is easy to check the characteristic property of the Laplace operator:

$$\int_M (\omega, \Delta\omega)\nu = \int_M |d\omega|^2\nu + \int_M |d^*\omega|^2\nu,$$

where ν is the volume form associated with the Riemannian structure. For functions this reduces to $\Delta f = d^*df$ and

$$\int_M f\Delta f\nu = \int_M |df|^2\nu.$$

From this formula (more precisely, for its version with $f\Delta g$) it is clear that Δ is self-adjoint and non-negative. Moreover, since we assume M to be compact and connected, the only *harmonic* functions (i.e. satisfying $\Delta f = 0$) are constant functions. Therefore on the subspace

$$L_0^2(M) = \left\{f \in L^2(M) : \int_M f\nu = 0\right\}$$

the Laplace operator is strictly positive. By some standard tools of functional analysis it can be shown that Δ^{-1} is a compact operator on $L_0^2(M)$, and the smallest eigenvalue of $\Delta|_{L_0^2(M)}$ makes sense and is denoted by $\lambda_1(M)$. Therefore $\lambda_1(M)$ is the largest positive number satisfying the following

$$\int_M |df|^2\nu \ge \lambda_1(M) \cdot \int_M |f|^2\nu, \quad \text{for all} \quad f \quad \text{such that} \quad \int_M f\nu = 0.$$

The connection to the isoperimetric inequalities now becomes more clear, since for a partition $M = A \cup B$ into two subsets we can consider a functions almost constant on A, almost constant on B, and changing in a reasonable way near the boundary $\partial A = \partial B$. After normalizing it to have a zero mean, we may apply the definition of $\lambda_1(M)$ and deduce some kind of isoperimetric inequality, showing that $\partial A = \partial B$ is sufficiently large.

It is also possible to show the converse: By considering the sublevel sets $M_c = \{x \in M : f(x) \le c\}$ and applying a certain kind of isoperimetric inequality to them, a lower bound for $\lambda_1(M)$ follows by careful integration over c.

One particular case when the above construction simplifies greatly is the case of the isoperimetry on graphs, which is the main topic of the book [Lub94]. First, the isoperimetric constant of a graph $G = G(E, V)$ is the number c such that

$$|E(A, B)| \geq c \min\{|A|, |B|\}$$

for every partition of the set of vertices $V = A \cup B$, where $E(A, B)$ denotes the set of edges between A and B. It can be easily proved that $c(G)$ can be related to the smallest positive eigenvalue of the graph Laplacian, this is left as an exercise for the reader. The graph Laplacian is defined as follows: For a function f on vertices we put

$$\Delta f(y) = \deg y f(y) - \sum_{(x,y) \in E} f(x).$$

The only functions corresponding to the zero eigenvalue are constants and it is easy to show that Δ is self-adjoint and positive on the functions with zero mean. The corresponding minimal positive eigenvalue is denoted by $\lambda_1(G)$.

Now we make the following definition:

Definition 18.1. A family of graphs G_n is called a family of *expanders* is $|V(G_n)| \to \infty$ as $n \to \infty$, the degrees of their vertices are uniformly bounded, $\deg v \leq k$ for any $v \in V(G_n)$, and their isoperimetric constants $c(G)$ (or the numbers $\lambda_1(G)$) are uniformly bounded by a positive number c.

Informally, the expander graphs are quantitatively more than connected, while having the bounded degree of any vertex. The theory of expander graphs is large and interesting, so here we only mention the main facts without proofs.

First, if we try to construct a bipartite graph on two sets of size n and connect any vertex on the left hand side to some k vertices on the right hand side randomly, then, as n tends to infinity, the probability to have of the inequality $c(G) \geq c > 0$ tends to 1 if $c < k/2$. This means that appropriately constructed "random graphs" are expanders, and this fact has serious practical applications.

Second, an explicit construction of expanders is rather difficult. The most usual way is to consider an infinite discrete group Γ with a set of generators S and build its Cayley graph $G(\Gamma, S)$. Then take the set of quotients Γ/N_n by a family of normal subgroups such that $|\Gamma/N_n| \to \infty$ and consider the induced graph on Γ/N_n. It turns out that some properties of the group Γ and its representations on the Hilbert space (the "Kazhdan property T" or other similar properties) allow to prove that the constructed family of graphs is a family of expanders. For the details (and many other interesting stuff, like the Banach–Tarski paradox) the reader is referred to the already mentioned book [Lub94], or other sources.

19. Šidák's lemma

Here we are going to prove a useful fact about Gaussian measures, known as Šidák's lemma. We define a *centrally symmetric strip* in $\mathbb{R}^n$ as the set $S = \{x : |\lambda(x)| \leq w\}$, for some $\lambda \in \mathbb{R}^{\ltimes *}$ and $w \in \mathbb{R}^+$.

Theorem 19.1. *Let A be a centrally symmetric convex body in $\mathbb{R}^n$ and S be a centrally symmetric strip in $\mathbb{R}^n$. Then for a Gaussian probability measure μ we have:*

$$\mu(A) \cdot \mu(S) \leq \mu(A \cap S).$$

The key idea of the proof is to introduce a new probability measure $\nu(X) = \frac{\mu(X \cap S)}{\mu(S)}$. The inequality that we want to obtain is formalized in the following:

Definition 19.2. Suppose μ and ν are two probability measures on $\mathbb{R}^n$. We say that ν is *more peaked* than μ if for any centrally symmetric convex body A

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

Now the proof of Theorem 19.1 consist of two lemmas:

Lemma 19.3. *If ν is more peaked than μ then for any centrally symmetric log-concave density f we have*

$$\int_{\mathbb{R}^n} f\ d\nu \geq \int_{\mathbb{R}^n} f\ d\mu.$$

Proof. Rewrite the integral:

$$\int_{\mathbb{R}^n} f\ d\nu = \int_0^{f(x)} \int_{\mathbb{R}^n} 1\ d\nu dy = \int_0^{f(x)} \nu(C_y) dy,$$

where $C_y = \{x : f(x) \leq y\}$. For log-concave and centrally symmetric f the sets C_y are convex and centrally symmetric, so the inequality follows by integration in y. □

Lemma 19.4. *If the measure ν is more peaked than μ, as measures on $\mathbb{R}^n$, and τ is a finite centrally symmetric log-concave measure on $\mathbb{R}^m$, then $\nu \times \tau$ is more peaked than $\mu \times \tau$ on $\mathbb{R}^{n+m}$.*

Proof. Denote by x and y the points in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Consider a centrally symmetric convex body $A \subset \mathbb{R}^{n+m}$. The measure $1 \times \tau$ is log-concave and its restriction τ' to A is also log-concave and centrally symmetric. Hence by Theorem 12.1 the density

$$f(x) = \int_{\{y:(x,y)\in A\}} 1\ d\tau$$

is log-concave and centrally symmetric, since it is the density of the projection of τ' to $\mathbb{R}^n$.

Now we observe that

$$\nu \times \tau(A) = \int_{\mathbb{R}^n} f(x)\ d\nu \geq \int_{\mathbb{R}^n} f(x)\ d\mu = \mu \times \tau(A)$$

by Lemma 19.3. □

And the proof of Theorem 19.1 is complete by the following obvious:

Lemma 19.5. *If μ is the Gaussian measure on the line and S is a centrally symmetric segment then the measure defined by*

$$\nu(X) = \frac{\mu(X \cap S)}{\mu(S)}$$

is more peaked than μ.

Then it suffices to take the Cartesian product of the result of this last lemma with the Gaussian measure on $\mathbb{R}^{n-1}$ to obtain Theorem 19.1.

Remark 19.6. Note that it is conjectured (the Gaussian correlation conjecture) that Theorem 19.1 can be generalized to the case of two arbitrary centrally symmetric convex bodies A and B with the same inequality: $\mu(A) \cdot \mu(B) \leq \mu(A \cap B)$.

Another result related to the notion of "more peaked" is the lower bound for the section volume of the unit cube $Q^n = [-1/2, 1/2]^n$:

Theorem 19.7. *Let L be a k-dimensional linear subspace of $\mathbb{R}^n$. Then the k-dimensional volume of the section $L \cap Q^n$ is at least 1.*

Remark 19.8. The upper bounds on the section volume are harder to prove. It is known that any hyperplane section of Q^n has area at most $\sqrt{2}$. This result was extended to lower values of k, but for some pairs (k, n) the precise upper bound is still not known. See the review of K. Ball [Ball01] for the details of this and many other interesting facts.

Proof. Consider the uniform measure ν on Q^n, and compare it with the Gaussian measure μ with density $e^{-\pi|x|^2}$. Actually, ν is more peaked than μ, the proof is reduced to the one-dimensional case by applying Lemma 19.4 and noting that any product of two such measures (in possibly different dimensions) is log-concave.

Now take an ε-neighborhood L_ε of L. By the definition of "more peaked" we obtain:

$$\nu(L_\varepsilon) \geq \mu(L_\varepsilon).$$

This can be decoded as

$$\operatorname{vol}(L_\varepsilon \cap Q^n) \geq \int_{B_\varepsilon^{n-k}} e^{-\pi|x|^2}\, dx,$$

where B_ε^{n-k} is the $(n-k)$-dimensional ball of radius ε. The right hand side for $\varepsilon \to +0$ is asymptotically $v_{n-k}\varepsilon^{n-k}$. And the trick is that the k-dimensional volume of $L \cap Q^n$ may be defined (following Minkowski) to be

$$\operatorname{vol}_k L \cap Q^n = \lim_{\varepsilon \to +0} \frac{\operatorname{vol}(L \cap Q^n)_\varepsilon}{v_{n-k}\varepsilon^{n-k}}.$$

It remains to note that the difference between $\operatorname{vol}(L \cap Q^n)_\varepsilon$ and $\operatorname{vol} L_\varepsilon \cap Q^n$ is $o(\varepsilon^{n-k})$ (the proof is left to the reader) and the result then follows. □

20. Centrally symmetric polytopes

A usual way to apply the Šidák lemma (Theorem 19.1) is to analyze the behavior of a centrally symmetric convex polytope in terms of its facets. Any such polytope K is defined by a system of inequalities

$$|(n_i, x)| \leq w_i, \quad i = 1, \dots, N,$$

where n_i are unit normal vectors to facets of K and w_i are positive reals. Each such inequality defines a strip and therefore Theorem 19.1 is applicable. For example, the following lemma holds:

Lemma 20.1. *Assume a centrally symmetric convex polytope $K \subset \mathbb{R}^n$ contains the unit ball B and has $2N$ facets. Then it intersects more than a half of the sphere S of radius $c\sqrt{\frac{n}{\log N}}$ (for sufficiently large N), where $c > 0$ is some absolute constant.*

Sketch of the proof. Choose a Gaussian measure with density $\left(\frac{\alpha}{\pi}\right)^{n/2} e^{-\alpha|x|^2}$. An easy estimate using integration by parts shows that any strip P_i defined by $|(n_i, x)| \leq w_i$ has measure (here we use $w_i \geq 1$)

$$\mu(P_i) \geq \int_{-1}^{1} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha x^2}\, dx \geq 1 - \frac{1}{\sqrt{\pi\alpha}} e^{-\alpha}.$$

It is known (for example, from the direct calculation of moments with gamma function and the Chebyshev inequality) that this Gaussian measure is mostly concentrated around the radius $|x| = \sqrt{\frac{n}{\alpha}}$. By Theorem 19.1

$$\mu(K) = \mu(P_1 \cap \dots \cap P_N) \geq \mu(P_1) \cdot \dots \cdot \mu(P_N) \geq \left(1 - \frac{1}{\sqrt{\pi\alpha}} e^{-\alpha}\right)^N,$$

so, in order to prove the lemma (that K intersects more than a half of a sphere of radius $c\sqrt{\frac{n}{\log N}}$), we have to take α of order $\log N$ and check that $\mu(K)$ is greater than some absolute positive constant, that is

$$\left(1 - \frac{1}{\sqrt{\pi\alpha}} e^{-\alpha}\right)^N \geq c_2,$$

or

$$N \log\left(1 - \frac{1}{\sqrt{\pi\alpha}} e^{-\alpha}\right) \geq c_3 \tag{20.1}$$

for a negative constant c_3. Then we observe the value $\alpha = c\log N$ with some $c < 1$ satisfies this inequality for sufficiently large N. ☐

Using Lemma 21.5 from the next section we conclude (the Figiel–Lindenstrauss–Milman theorem):

Theorem 20.2. *Let K be a centrally symmetric convex polytope with $2N$ facets and $2M$ vertices, then*

$$\log N \cdot \log M \geq \gamma n$$

for some absolute constant $\gamma > 0$.

Proof. By Lemma 21.5 we assume that K contains the unit ball B and is contained in the ball $\sqrt{n}B$. The dual body K^*, defined by

$$K^* = \{x \in \mathbb{R}^n : \forall y \in K\ (x, y) \leq 1\},$$

is therefore contained in B and contains $\frac{1}{\sqrt{n}}B$. By Lemma 20.1 K intersects more than a half of the sphere of radius $r = c\sqrt{\frac{n}{\log N}}$, and K^* intersects more than a half of the sphere of radius $r^* = c\sqrt{\frac{1}{\log M}}$.

Note that the inequality $r^*r \leq 1$ is what we need to prove. Assume the contrary: $r^*r > 1$. Then for any x with $|x| = r$ consider its renormalized $x^* = r^*\frac{x}{|x|}$. Note that it cannot happen that $x \in K$ and $x^* \in K$ simultaneously, because $(x, x^*) > 1$. Hence either K intersects at most half of the sphere S_r or K^* intersects at most half of the sphere S_{r^*}, which is a contradiction.

Small values of N and M, not suitable for Lemma 20.1, may be considered separately in a similar fashion. ☐

In fact, the above lemma and theorem can be proved without the Šidák lemma by estimating the surface areas of spherical caps. The idea is that a spherical cap may be very close to a halfsphere in terms of distance and still have very small surface area. Then we observe that the complement to a strip on a sphere is a pair of caps of small area, and if the sum of all areas is at most half of the area of the sphere then Lemma 20.1 is established. This cap area estimate corresponds to the value $\alpha = c\log N$, while the stronger estimate (20.1) was not fully used in the above proof.

However, the approach with caps requires writing down and estimating some integrals of trigonometric functions, while the Šidák lemma allows us to work with simple exponential expressions. Which is more important, a careful use of the Šidák lemma allowed A. Barvinok [Barv11] to prove the following strengthening of Theorem 20.2: Under the assumption of the above theorem the inequality $N \leq \alpha n$ implies $M \geq e^{\beta n}$, for some positive $\beta = \beta(\alpha)$.

Another result estimates from below the volume of a centrally symmetric polytope containing the unit ball:

Lemma 20.3. *Assume a centrally symmetric convex polytope $K \subset \mathbb{R}^n$ contains the unit ball B and has $2N$ facets. Then its volume is at least*

$$\left(\frac{cn}{\log N}\right)^{n/2} v_n,$$

(for sufficiently large N), where $c > 0$ is some absolute constant.

Proof. The intersection of K with the ball, bounded by the sphere from 20.1, is at least half of this ball by volume. □

Corollary 20.4. *Let $K \subset \mathbb{R}^n$ be a polytope with N vertices on the unit ball B. Then*

$$\frac{\operatorname{vol} K^\circ}{\operatorname{vol} K} \geq \left(\frac{cn}{\log N}\right)^n.$$

Proof. We can replace K with $\operatorname{conv}\{K \cup -K\}$ to make it centrally symmetric. At this step the volume of K increases and the volume of K° decreases. So assume K centrally symmetric with $2N$ vertices. Note that the polar K° has $2N$ facets and then

$$\frac{\operatorname{vol} K^\circ}{\operatorname{vol} K} = \frac{(\operatorname{vol} K^\circ)^2}{\operatorname{vol} K \cdot \operatorname{vol} K^\circ} \geq \frac{(\operatorname{vol} K^\circ)^2}{v_n^2} \geq \left(\frac{cn}{\log N}\right)^n,$$

where we used Corollary 14.6 and Lemma 20.3. □

The same approach with more care allows to prove that $\log N$ can be replaced with $\log(N/n)+1$, as was shown by Gluskin in [Glu1988]. Bárány and Füredi in [BF87] proved a similar result using different technique and deduced the following: Suppose we have a black box, that for any point $x \in \mathbb{R}^n$ either asserts that x is inside the unknown convex body K, or provides a separating hyperlane for K and x. From the above results we conclude that an algorithm estimating the volume of K with such a black box needs a huge number of queries to the black box. But it turns out that practically the volume can be efficiently estimated with a randomized algorithm.

Another famous problem about centrally symmetric polytopes is to prove that the total number of faces of all dimensions (including the polytope itself) is at least 3^n. The reader may easily check that this bound is attained for the cube or its dual, the crosspolytope. In the case of simple or simplicial polytopes Stanley proved [Stan87] this conjecture using the correspondence between the linear space spanned by the faces of a polytope up to a certain equivalence relation and the cohomology ring of the corresponding algebraic (*toric*) variety. Let us give a sketch for those knowing or willing to learn some toric geometry: The symmetry of a polytope makes an action of the involution ι on the polytope P and its corresponding toric variety X_P. Let $T = (\mathbb{C}^*)^n$ be the torus acting on X_P. Then it remains to compare the two descriptions of the T-equivariant cohomology of X_P: $H_T^*(X_P) = H^*(X_P) \otimes H^*(BT)$ from the collapsing spectral sequence, on the one hand, and $H_T^*(X_P)$ is the Stanley–Reisner ring of P, that is something defined combinatorially by P. From this it follows that the Hilbert series of the trace of ι on those spaces satisfies:

$$h(\iota, H_T^*(X_P)) = h(\iota, H_T^*(X_P)) \cdot h(\iota, H^*(BT)) = \frac{h(\iota, H_T^*(X_P))}{(1+t)^n} = 1.$$

Hence $h(\iota, H^*(X_P)) = (1+t)^n$ and from this it is possible to deduce the lower bounds on the numbers of faces, calculated from $h_P(t) = h(\iota, H^*(X_P))$, see the details in [Stan87] and [Cox11].

21. Dvoretzky's theorem

We are going to discuss one of the most famous applications of the concentration phenomenon on the sphere, the Dvoretzky theorem:

Theorem 21.1. *For a positive real ε and a positive integer k there exists another positive integer $n(k,\varepsilon)$ with the following property. If $\|\cdot\|$ is any norm on $\mathbb{R}^n$ then there exists a Euclidean norm $|\cdot|$ on $\mathbb{R}^n$ and a k-dimensional linear subspace $L \subseteq \mathbb{R}^n$ such that*

$$|x| \le \|x\| \le (1+\varepsilon)|x|.$$

The proof needs several lemmas, of which we prove all but one. First, we will need δ-nets on the sphere S^{k-1}.

Definition 21.2. A finite set $X \subset S^{k-1}$ is a δ-net, if for every $x \in S^{k-1}$

$$\operatorname{dist}(x, X) \le \delta.$$

In what follows we assume for simplicity that $\delta \le \pi/4$ and assume k to be sufficiently large.

Lemma 21.3. *There exists a δ-net in S^{k-1} of size at most $\frac{k4^{k-1}}{\delta^{k-1}}$.*

Proof. Find an inclusion maximal set of disjoint spherical caps of radius $\delta/2$ (balls in geodesic metric) in S^{k-1}. Let X be the set of their centers. Since we cannot add any other spherical cap of radius $\delta/2$ to the set, every point of S^{k-1} is at distance at most δ from some $x \in X$ and therefore X is a δ-net.

Now we estimate $|X|$ comparing the total surface area of the caps, which is at least $|X|v_{k-1}(\delta/4)^{k-1}$ (here we use that δ is not too big), to the surface area of the sphere $s_k = kv_k$. Hence

$$|X| \le \frac{kv_k 4^{k-1}}{v_{k-1}\delta^{k-1}} \le \frac{k4^{k-1}}{\delta^{k-1}}$$

here we use that $v_k \le v_{k-1}$ for $k \ge 6$, which can be seen from the explicit formula $v_k = \frac{\pi^{k/2}}{\Gamma(k/2+1)}$. ☐

Lemma 21.4. *If X is a δ-net in S^{k-1} and $\delta < \pi/4$ then every $x' \in S^{k-1}$ can be expressed as a positive linear combination of vectors from X with sum of coefficients at most $\sqrt{2}$.*

Proof. Let us prove that the ball B' of radius $1/\sqrt{2}$ is contained in $\operatorname{conv} X$. Assuming the contrary by the Hahn–Banach theorem we obtain a vector y with $|y| = 1/\sqrt{2}$ such that the set

$$H_y = \{x : (x,y) \ge (y,y)\}$$

has no common interior point with $\operatorname{conv} X$. Then it is easy to see that the normalized $y/|y|$ is at geodesic distance at least $\pi/4 > \delta$ from X.

So we conclude that S^{k-1} is inside $\sqrt{2}\operatorname{conv} X$, which is equivalent to what we need. ☐

Lemma 21.5. *Let K be a centrally symmetric convex body in $\mathbb{R}^n$. Then there exists a centrally symmetric ellipsoid E such that*

$$E \subseteq K \subseteq \sqrt{n}E.$$

Proof. Take as E the ellipsoid of maximal volume contained in K, the John ellipsoid. After a linear transformation assume that E is a unit ball. If there exist a point $x \in K$ with $|x| > \sqrt{n}$ than it can be shown by a straightforward calculation that after stretching E in the direction of x and shrinking it in the orthogonal to x directions E can increase its volume while remaining inside the set $\operatorname{conv}(E \cup \{x\} \cup \{-x\}) \subseteq B$. ☐

Now, in Theorem 21.1 we consider some norm on $\mathbb{R}^n$. By Lemma 21.5 we choose the ellipsoid E and define $|\cdot|$ to be the Euclidean norm with E as the unit ball. Then we consider $f(x) = \|x\|$ as a function of the sphere S^{n-1}, the conclusion of Lemma 21.5 reads as

$$\frac{1}{\sqrt{n}} \leq f(x) \leq 1$$

on the sphere. The condition $f(x) \leq 1$ (using the triangle inequality for the norm) means that f is 1-Lipschitz on S^{n-1}. Let M be the median of f. For the following lemma we only give a sketch of a proof:

Lemma 21.6. *Under the above assumptions on $\|\cdot\|$ and $|\cdot|$, generated by the John ellipsoid, the median M of $\|\cdot\|$ has the lower bound*

$$c\sqrt{\frac{\log n}{n}}$$

with some absolute constant c.

See the discussion of this lemma in [Ball97]. A different approach to this difficulty is to select an orthonormal (relative to $|\cdot|$) base e_i such that $\|e_i\| \geq \frac{n-i}{n}$ (this is the Dvoretzky–Rogers lemma), then choose the linear span of the first $n/2$ of these vectors as the new space to work in and establish an analogue of Lemma 21.6 for this subspace. This is established using averaging of the expression $\|x_1e_1 + \dots + x_ne_n\|$ over all possible changes of signs of the coordinates x_i, along with the triangle inequality for the norm. The details are given, for example, in [Led01, Section 3.5].

Proof of Theorem 21.1 assuming Lemma 21.6. By Corollary 17.3 (here σ is the probability measure on S^{n-1}) for the set

$$C = \{x \in S^{n-1} : |\|x\| - M| \geq M\varepsilon/8\}$$

we have

$$\sigma(C) \leq 2e^{-\frac{(n-2)M^2\varepsilon^2}{128}} \leq 2e^{-\frac{c^2\varepsilon^2\log n}{128}}.$$

Now we take $\delta = \varepsilon/4$ and choose a δ-net X on S^{k-1}, the latter is considered as the unit sphere of the coordinate subspace $\mathbb{R}^k \subset \mathbb{R}^n$. Applying a random rotation ρ to X we see the following: For any $x_i \in X$ the probability of the event $x_i \in C$ is at most $2e^{-\frac{c^2\varepsilon^2\log n}{128}}$.

If in total

$$2e^{-\frac{c^2\varepsilon^2\log n}{128}}|X| < 1$$

then for some random rotation the whole X gets inside $S^{n-1} \setminus C$. Then we choose L to be the image $\rho(\mathbb{R}^k)$ and denote by $S(L)$ the unit sphere of L. From the bound on $|X|$ of Lemma 21.3, the inequality condition is satisfied when

$$\frac{k16^{k-1}}{\varepsilon^{k-1}}e^{-\frac{c^2\varepsilon^2\log n}{128}} < 1/2,$$

which is indeed true for sufficiently large n.

By Lemma 21.4 we conclude that for any $x \in S(L)$, after a rescaling making the median M equal 1,

$$\|x\| \leq \|\sum_{x_i\in X} c_ix_i\| \leq \sqrt{2}\max_{x_i\in X}\|x_i\| \leq \sqrt{2}(1+\varepsilon/8).$$

Hence, again using the triangle inequality for the norm, $\|\cdot\|$ on $S(L)$ has Lipschitz constant at most $\sqrt{2}(1+\varepsilon/8)$. It follows that the values of $\|\cdot\|$ on $S(L)$ are between

$$(1-\varepsilon/8) - \varepsilon/4 \cdot \sqrt{2}(1+\varepsilon/8)$$

and

$$(1 + \varepsilon/8) + \varepsilon/4 \cdot \sqrt{2}(1 + \varepsilon/8).$$

For sufficiently small ε, after a slight rescaling of $\|\cdot\|$ we obtain the inequality

$$|x| \leq \|x\| \leq (1 + \varepsilon)|x|$$

for any $x \in L$. □

22. Topological and algebraic Dvoretzky type results

It was noted in [Mil88] that a simple proof for the Dvoretzky theorem would follow from the following topological conjecture of Knaster [Kna47]:

Conjecture 22.1. *For any finite set $X \subset S^{n-1}$ with $|X| = n$ and any continuous function $f : S^{n-1} \to \mathbb{R}$ it is possible to find a rotation ρ such that*

$$f(\rho x_1) = \dots = f(\rho x_n),$$

where $\{x_1, \dots, x_n\}$ are the points of X.

Some cases of this conjecture were confirmed: when X is an orthonormal basis, when $n = 3$, when $|X|$ is prime and the points of X form a two-dimensional regular polygon, see the discussion in [DK11] for more details and references. But unexpectedly, in the paper of Kashin and Szarek [KS03] a counterexample was constructed using some knowledge about sections of a cube and similar things. A greatly simplified exposition of this counterexample is given in [Mat10, Miniature 32].

Of course, if the conjecture were true we could take a δ-net on S^{k-1} as X of size at most $\left(\frac{4}{\delta}\right)^k$ by Lemma 21.3. Then, assuming Conjecture 22.1, for $n \geq \left(\frac{4}{\delta}\right)^k$ we could rotate X and rescale the Euclidean norm to have the equality $\|x\| = |x|$ for any $x \in X$. The rest would follow from Lemma 21.4.

This approach could still pass if we establish the *weak* Knaster conjecture with a reasonable estimate for the function $n(m)$:

Conjecture 22.2. *There exists a function $n = n(m)$ with the following property. For any finite set $\{x_1, \dots, x_m\} \subset S^{n-1}$ of size k and any continuous function $f : S^{n-1} \to \mathbb{R}$ it is possible to find a rotation ρ such that*

$$f(\rho x_1) = \dots = f(\rho x_m).$$

About this generalized conjecture, almost as little is known as about the original one. In fact, already for $m = 4$ the existence of $n(m)$ is only known for some very particular types of sets $\{x_1, \dots, x_m\}$.

Another direction of Dvoretzky-type results is the "algebraic Dvoretzky theorem" from [DK11]:

Theorem 22.3. *For an even positive integer d and a positive integer k there exists $n(d, k)$ such that for any homogeneous polynomial f of degree d on $\mathbb{R}^n$, where $n \geq n(d, k)$, there exists a linear k-subspace $L \subseteq \mathbb{R}^n$ such that $f|_L$ is proportional to the $d/2$-th power of the standard quadratic form*

$$Q = x_1^2 + x_2^2 + \dots + x_n^2.$$

This theorem is very similar to the Dvoretzky theorem, but unlike the latter, it gives a precisely "round" section for a polynomial. Unfortunately, the topological tools used in its proof do not give any explicit bound on $n(d, k)$. Moreover, it is not clear how to deduce the Dvoretzky theorem from this result, even if the bound were reasonable.

On the other hand, for odd degree d it is proved with elementary topology [DK11] that any homogeneous polynomial of degree d in n variables vanishes on some linear k-subspace with $n = k + \binom{d+k-1}{d}$. This fact is originally due to B.J. Birch [Bir57], who established it by algebraic tools with a worse estimate for $n(d,k)$.

Let us sketch the proof of the Birch theorem for a homogeneous polynomial P of degree d in n variables. The space of all orthonormal k-frames in $\mathbb{R}^n$ is parameterized by the Stiefel manifold $V_{n,k}$. For any frame $(e_1,\dots,e_k)$ the polynomial

$$P_e(t_1,\dots,t_k) = P(t_1e_1 + t_2e_2 + \dots + t_ke_k)$$

is a degree d homogeneous polynomial in k variables. The space W of all such polynomials has dimension $\binom{d+k-1}{d}$. The correspondence $e \mapsto P_e$ makes an odd map from $V_{n,k}$ to W. Here *odd* is understood for $V_{n,k}$ so that along with a frame $(e_1,\dots,e_k)$ we can consider the frame

$$-(e_1,\dots,e_k) = (-e_1,\dots,-e_k).$$

Then one form of the generalized Borsuk–Ulam theorem (see [Mat03]) asserts that some frame is mapped to zero (that is exactly what we need), provided the space $V_{n,k}$ is $(\dim W - 1)$-connected. This is indeed the case when $n \ge k + \binom{d+k-1}{d}$.

A similar proof works for the following result about "making a convex function symmetric":

Theorem 22.4. *Let $X \subset S^{k-1}$ be a centrally symmetric subset with $|X| = 2m$. If $n \ge m + k$ and $f : S^{n-1} \to \mathbb{R}$ is a continuous function, then it is possible to find an isometric copy $\rho(X) \subset S^{n-1}$ such that for any $x \in \rho(X)$*

$$f(x) = f(-x).$$

References

[Alon86] N. Alon. The number of polytopes, configurations and real matroids. *Mathematika*, 33:62–71, 1986.

[ABMR11] J.L. Arocha, J. Bracho, L. Montejano, J.L. Ramírez Alfonsín. Transversals to the convex hulls of all k-sets of discrete subsets of $\mathbb{R}^n$. *Journal of Combinatorial Theory, Series A* 118(1):197–207, 2011.

[Ati83] M.F. Atiyah. Angular momentum, convex polyhedra and algebraic geometry. *Proceedings of the Edinburgh Mathematical Society (2)*, 26(2):121–133, 1983.

[Ball97] K. Ball. An elementary introduction to modern convex geometry. *Flavors of Geometry. MSRI Publications, 31*, 1997.

[Ball01] K. Ball. Convex geometry and functional analysis. *Handbook of the Geometry of Banach Spaces*, 161–194, 2001.

[Ball04] K. Ball. An elementary introduction to monotone transportation. *Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics 1850*, 99–106, 2004.

[BF87] I. Bárány, Z. Füredi. Computing the volume is difficult. *Discrete and Computational Geometry*, 2(1):319–326, 1987.

[Barv11] A. Barvinok. A bound for the number of vertices of a polytope with applications. *Arxiv preprint arXiv:1108.2871*, 2011.

[Bern76] D.N. Bernstein. The number of roots of a system of equations. *Functional Anal. Appl.*, 9(3):183–185, 1976.

[Bir57] B.J. Birch. Homogeneous forms of odd degree in a large number of variables. *Mathematika* 4:102–105, 1957.

[Bor33] K. Borsuk. Drei Sätze über die n-dimensionale euklidische Sphäre. *Fundam. Math*, 20(1):177–190, 1933.

[BH11] B. Bukh, A. Hubard. Space crossing numbers. *Arxiv preprint arXiv:1102.1275*, 2011.

[Cha89] B. Chazelle, E. Welzl. Quasi-optimal range searching in spaces of finite VC-dimension. *Discrete Comput. Geom.*, 4:467–489, 1989.

[Cox11] D. Cox, J. Little, H. Schenck. *Toric Varieties.* Graduate Studies in Mathematics 124, American Mathematical Society, 2011.

[Dol94] V.L. Dol'nikov. Transversals of families of sets in $\mathbb{R}^n$ and a connection between the Helly and Borsuk theorems. *Sb. Math.* 79(1):93–107, 1994.

[DK11] V.L. Dol'nikov, R.N. Karasev. Dvoretzky type theorems for multivariate polynomials and sections of convex bodies. *Geometric And Functional Analysis*, 21(2):301–318, 2011.

[Ful93] W. Fulton. *Introduction to toric varieties.* Annals of Mathematics Studies, 131. Princeton University Press, Princeton, NJ, 1993.

[GM2001] A. A. Giannopoulos, V. D. Milman. Euclidean structure in finite dimensional normed spaces. *Handbook of the Geometry of Banach Spaces*, 1:707–779, 2001.

[Glu1988] E.D. Gluskin. Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces. (in Russian) *Mat. Sbornik* 136(178):85–96, 1988.

[GH78] P. Griffiths, J. Harris. *Principles of algebraic geometry.* Wiley-Interscience, New York–Chichester–Brisbane–Toronto, 1978.

[GP86] J. E. Goodman, R. Pollack. Upper bounds for configurations and polytopes in $\mathbb{R}^d$. *Discrete Comput. Geom.*, 1:219–227, 1986.

[Grom90] M. Gromov. Convex sets and Kähler manifolds. *IHES*, www.ihes.fr/ gromov/PDF/[68].pdf, 1990.

[Grom03] M. Gromov. Isoperimetry of waists and concentration of maps. *Geometric and Functional Analysis*, 13:178–215, 2003.

[Guth13] L. Guth. Unexpected applications of polynomials in combinatorics. math.mit.edu/ lguth/Exposition/erdossurvey.pdf, 2013.

[Har76] C. G. A. Harnack. Über Vieltheiligkeit der ebenen algebraischen Curven. *Math. Ann.*, 10:189–199, 1876.

[Hopf44] H. Hopf. Eine Verallgemeinerung bekannter Abbildungs und Überdeckungssätze. *Portugaliae Math.*, 4:129–139, 1944.

[KMS12] H. Kaplan, J. Matoušek, M. Sharir. Simple proofs of classical theorems in discrete geometry via the Guth–Katz polynomial partitioning technique. *Discrete Comput. Geom.*, 48(3):499–517, 2012.

[KS03] B. S. Kashin, S. J. Szarek. The Knaster problem and the geometry of high-dimensional cubes. *Comptes Rendus Mathematique*, 336(11):931–936, 2003.

[Kna47] B. Knaster. Problem 4. *Colloq. Math.*, 30:30–31, 1947.

[Ku08] G. Kuperberg. From the Mahler conjecture to Gauss linking integrals. *Geom. Funct. Anal.* 18(3):870–892, 2008.

[Led01] M. Ledoux. *The concentration of measure phenomenon.* Amer. Math. Soc., 2001.

[Leh09A] J. Lehec. A direct proof of the functional Santaló inequality. *Comptes Rendus Mathematique*, 347(1):55–58, 2009.

[Leh09B] J. Lehec. Partitions and functional Santaló inequalities. *Archiv der Mathematik*, 92(1):89–94, 2009.

[Lov78] L. Lovász. Kneser's conjecture, chromatic number and homotopy. *Journal of Combinatorial Theory, Series A* 25:319–324, 1978.

[Lub94] A. Lubotzky. *Discrete groups, expanding graphs and invariant measures.* Birkhäuser, 1994.

[Mat03] J. Matoušek. *Using the Borsuk-Ulam theorem: Lectures on topological methods in combinatorics and geometry.* Springer Verlag, 2003.

[Mat10] J. Matoušek. Thirty-three miniatures. Mathematical and algorithmic applications of linear algebra. *Student Mathematical Library 53*, American Mathematical Society, Providence, RI, 2010.

[MS86] V. D. Milman, G. Schechtman. *Asymptotic theory of finite dimensional normed spaces. Lecture Notes in Mathematics 1200.* Springer Verlag, 1986.

[Mil88] V. D. Milman. A few observations on the connections between local theory and some other fields. *Geometric Aspects of Functional Analysis. Lecture Notes in Mathematics 1317*, 283–289, 1988.

[NSV02] F. Nazarov, M. Sodin, A. Vol'berg. The geometric Kannan-Lovász-Simonovits lemma, dimension-free estimates for the distribution of the values of polynomials, and the distribution of the zeros of random analytic functions. *St. Petersbg. Math J.*, 14(2):351–366, 2002.

[NW13] E. Nevo, S. Wilson. How many n-vertex triangulations does the 3-sphere have? *Arxiv preprint* arXiv:1311.1641, (2013).

[Sid67] Z. Šidák. Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, 62(318):626–633, 1967.

[ST12] J. Solymosi and T. Tao. An incidence theorem in higher dimensions. *Discrete Comput. Geom.*, 48(2):255–280, 2012.

[Stan87] R.P. Stanley. On the number of faces of centrally-symmetric simplicial polytopes. *Graphs and Combinatorics*, 3(1):55-66, 1987.

[Stan89] R.P. Stanley. Log-concave and unimodal sequences in algebra, combinatorics, and geometry. *Annals of the New York Academy of Sciences*, 576:500–535, 1989.

[Ste45] H. Steinhaus. Sur la division des ensembles de éspace par les plans et des ensembles plans par les cercles. *Fundam. Math*, 33:245–263, 1945.

[ST42] A. H. Stone and J. W. Tukey. Generalized "sandwich" theorems. *Duke Mathematical Journal*, 9(2):356–359, 1942.

[SW85] W. Stromquist and D. Woodall. Sets on which several measures agree. *J. Math. Anal. Appl.*, 108:241–248, 1985.

[ST83] E. Szemerédi and W. T. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3:381–392, 1983.

[Tao07] T. Tao. Open question: the Mahler conjecture on convex bodies. terrytao.wordpress.com/2007/03/08/open-problem-the-mahler-conjecture-on-convex-bodies/, 2007.

[TaoVu10] T. Tao, V. H. Vu. Additive combinatorics. Cambridge studies in advanced mathematics 105, 2010.

[Tao11] T. Tao. The Brunn–Minkowski inequality for nilpotent groups. terrytao.wordpress.com/2011/09/16/the-brunn-minkowski-inequality-for-nilpotent-groups/, 2011.

[Wel88] E. Welzl. Partition trees for triangle counting and other range searching problems. *Proc. 4th Annu. ACM Sympos. Comput. Geom.*, 23–33, 1988.

[Wel92] E. Welzl. On spanning trees with low crossing numbers. *Data structures and efficient algorithms, Lecture Notes in Computer Science, vol. 594.* Springer Verlag, 1992, 233–249.

[YY85] A. C. Yao, F. F. Yao. A general approach to d-dimensional geometric queries. *Proceedings of the seventeenth annual ACM symposium on Theory of computing, STOC 85*, New York, NY, USA, ACM, 1985, 163–168.

Roman Karasev, Dept. of Mathematics, Moscow Institute of Physics and Technology, Institutskiy per. 9, Dolgoprudny, Russia 141700

Roman Karasev, Institute for Information Transmission Problems RAS, Bolshoy Karetny per. 19, Moscow, Russia 127994

Roman Karasev, Laboratory of Discrete and Computational Geometry, Yaroslavl' State University, Sovetskaya st. 14, Yaroslavl', Russia 150000

E-mail address: r_n_karasev@mail.ru

URL: http://www.rkarasev.ru/en/