THE TEACHING OF MATHEMATICS
2006, Vol. IX, 2, pp. 1–17

TVERBERG'S THEOREM AND SHADOWS OF SIMPLICES AND CONVEX BODIES

Rade T. Živaljević

Abstract. Helge Tverberg published more than forty years ago his original proof of the theorem which has been widely acclaimed and today bears his name. This beautiful result has been one of the most celebrated results of discrete geometry and, together with its relatives, still remains a central and one of the most intriguing results of geometric combinatorics. Here we give a reasonably non-technical presentation of this result having in mind a larger mathematical audience, particularly school teachers and their talented students, hoping that it may raise their interest for this very attractive area of mathematics. In the remaining part of the article we briefly visit some of other branches of convex geometry and outline how "smashing" and 'slicing" of convex bodies offers a deep insight into their structure and behavior.

ZDM Subject Classification: G95; *AMS Subject Classification*: 00A35.

Key words and phrases: Simplex; join of sets; colorful Carathéodory theorem; Radon's theorem; Tverberg's theorem.

1. Introduction

Here is a fragment from "Ode to Simplex" (in Serbian) which gradually emerged in the announcements of Belgrade seminar "Combinatorics in Geometry, Topology and Algebra" (CGTA for short).

Ја сам симплекс троугласте сенке
Смерни атом васколиког Мира
Конвексности древне проповедник
Плод заноса креде и лењира.
... ...
Симплексна су сад дошла времена
Симплексу се скандира и свира
Број и Симплекс два страшна симбола
Који од њих Миром доминира!?

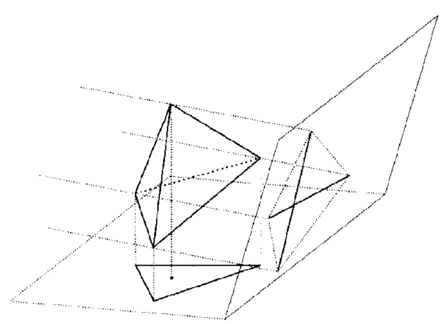


Fig. 1

We wouldn't dare to offer a translation of these verses here, however they are just a reminder about the role the simplex plays (as one of the basic geometric shapes) throughout mathematics and in our perception of reality. Tverberg's theorem itself is a combinatorial statement about the *shadow* of a $(d-1)(q+1)$-dimensional simplex in a d-dimensional Euclidean space.

By a *shadow* of a convex set we mean the projection $\pi(K)$ of K where $\pi\colon \mathbf{R}^n \to \mathbf{R}^d$ is a linear or affine map. We start with a formal definition of the simplex, while a brief reminder of other basic concepts of convex geometry can be found in the Appendix, Section 6.

Definition 1.1. *A d-dimensional simplex $\sigma = \sigma^d$ in $\mathbf{R}^N$ is the convex hull $\sigma := \operatorname{conv}(S)$ of a collection $S = \{a_0, a_1, \dots, a_d\}$ od $d+1$ points in $\mathbf{R}^N$, provided the affine span of S is d-dimensional. A 1-dimensional simplex is a line segment, 2-dimensional simplex is a triangle, and a 3-dimensional simplex is a tetrahedron.*

2. Colorful Carathéodory theorem

Suppose three triangles in the plane have a point in common (Figure 2). Then one can select a vertex from each of these triangles and form a new triangle which also contains this point. This is a 2-dimensional version of “Colorful Carathéodory theorem” proved by Imre Bárány [Bar82].

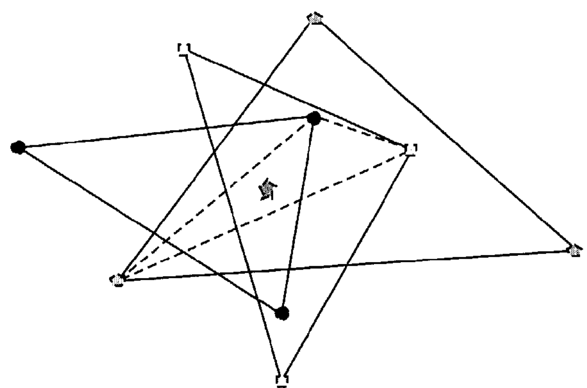


Fig. 2: The colorful Carathéodory theorem

The usual (monochromatic) Carathéodory theorem would under these circumstances guarantee that if a point p is in the convex hull $\operatorname{conv}(S)$ of a planar set S, then $p \in \Delta$ for some (possibly degenerate) triangle with vertices from S. The colored version is more precise in the sense that if three sets S_1, S_2, S_3 of (colored)[1] points in the plane have property

$$p \in \operatorname{conv}(S_1) \cap \operatorname{conv}(S_2) \cap \operatorname{conv}(S_3) \tag{1}$$

then there exists a *rainbow* triangle $\Delta = \operatorname{conv}\{x_1, x_2, x_3\}$, where $x_i \in S_i$, such that $p \in \Delta$.

The idea of the proof of this statement is quite natural (see e.g. [M02] for a more detailed presentation). If a multicolored (rainbow) triangle with the desired

[1] In this black-and-white version, in Fig. 2 red points are represented by circles, yellow points by squares and green ones by small pentagons.

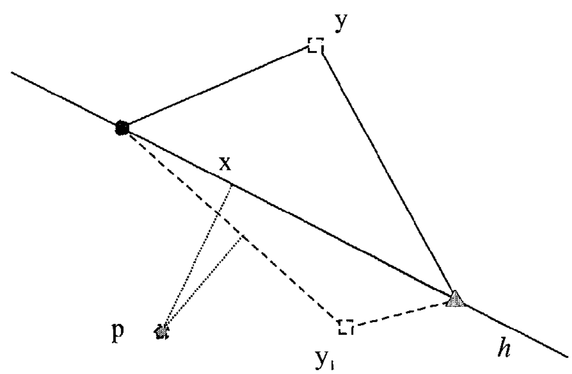


Fig. 3: Improving rainbow triangles

properties does not exist, let us choose one of them that misses the point p by the smallest margin. In other words let $\Delta = \text{conv}\,\{x_1, x_2, x_3\}$ be the rainbow triangle such that the distance $d = d(p, \Delta) > 0$ is the smallest possible.

Suppose that $x \in \Delta$ is a point such that $d(p, x) = d$. Let h be the line containing x and perpendicular to the line segment $\overline{px}$, as in Figure 3. Then p and the triangle Δ are on the opposite sides of this line and let h^- be the open half-space determined by h which contains p. Notice that x belongs to one of the sides of Δ hence one of the colors (yellow = square in the picture) is used for coloring the opposite vertex y. Since p is in the convex hull of yellow points, there must exist a yellow point y_1 in h^-. But if Δ_1 is the rainbow simplex obtained from Δ by replacing vertex y by vertex y_1 then $d(p, \Delta_1) < d(p, \Delta_1)$. Contradiction!

The proof above was carried on in the plane but the ideas are quite general. We leave it as an exercise for the reader to extend the arguments, observing the necessary modifications, to the proof of the general case.

Theorem 2.1. *If (possibly degenerate) simplices $\sigma_0, \sigma_1, \ldots, \sigma_d$ in $\mathbf{R}^d$ have a point in common, then this point is also contained in a simplex of the form $\sigma := \text{conv}\,\{x_j\}_{j=0}^d$, where x_j is a vertex of the simplex σ_j.*

Recall that a simplex $\sigma = \text{conv}\,(S) = \text{conv}\,\{v_j\}_{j=0}^d$ is *degenerate* if one of its vertices v_i is in the affine span of the set $S \setminus \{v_i\}$, i.e. if the condition that the affine span of S is d-dimensional in Definition 1.1 is not fulfilled.

Corollary 2.2. *Let $\sigma_0, \sigma_1, \ldots, \sigma_{d-k}$ be a collection of simplices in $\mathbf{R}^d$ and D a k-dimensional affine subspace of $\mathbf{R}^d$. If $D \cap \sigma_i \neq \emptyset$ for each $i = 0, 1, \ldots, d-k$, then one can choose a vertex x_i of σ_i for each i such that*

$$\text{conv}\,\{x_0, x_1, \ldots, x_{d-k}\} \cap D \neq \emptyset.$$

This is an immediate consequence of Theorem 2.1 applied on simplices $\hat{\sigma}_0 := \pi(\sigma_0), \hat{\sigma}_1 := \pi(\sigma_1), \ldots, \hat{\sigma}_{d-k} := \sigma_{d-k}$, where $\pi\colon \mathbf{R}^d \to D^\perp$ is the orthogonal projection on the orthogonal complement $D^\perp$ of D.

Corollary 2.3. *Let $[x_0^+, x_0^-], [x_1^+, x_1^-], \ldots, [x_{d-k}^+, x_{d-k}^-]$ be a collection of line segments (one-dimensional simplices) in $\mathbf{R}^d$ such that some k-dimensional, affine subspace $D \subset \mathbf{R}^d$ intersects all of them. Then for some choice $\epsilon_i \in \{+,-\}$ of signs*

$$\Delta \cap D \neq \emptyset \quad \text{where} \quad \Delta := \mathrm{conv}\,\{x_0^{\epsilon_0}, x_1^{\epsilon_1}, \ldots, x_{d-k}^{\epsilon_{d-k}}\}.$$

3. Radon's theorem

A 4-element set S in $\mathbf{R}^2$ can be divided into two subsets such that the corresponding convex hulls have a non-empty intersection. Indeed, either the convex hull $\mathrm{conv}\,(S)$ is a triangle (possibly degenerate), or a rectangle. In both cases the required division is easily found (Figure 1). This is a special case of a classical result, known as Radon's theorem.

Theorem 3.1. *Suppose that $X = \{x_j\}_{j=1}^{d+2} \subset \mathbf{R}^d$ is a $(d+2)$-element set in a d-dimensional, Euclidean space. Then there exists a partition $X = X_1 \cup X_2$ of X such that*

$$\mathrm{conv}\,(X_1) \cap \mathrm{conv}\,(X_2) \neq \emptyset. \tag{2}$$

Theorem 3.1 can be restated as a result about shadows of $(d+1)$-dimensional simplices in d-dimensional Euclidean spaces. It may be instructive to look first at the planar case $d = 2$ and briefly inspect Figure 1 displayed in the Introduction (Section 1). This picture should convey the idea that the convex hull of a 4-element set of points in $\mathbf{R}^2$ is the image (shadow) of a 3-dimensional simplex (tetrahedron). This is true in general. Indeed, if $\Delta = \Delta^{d+1} = \mathrm{conv}\,\{a_1, a_2, \ldots, a_{d+2}\}$ is a $(d+1)$-dimensional simplex with vertices $a_1, a_2, \ldots, a_{d+2}$, than for any $(d+2)$-element set $X = \{x_j\}_{j=1}^{d+2} \subset \mathbf{R}^d$ there is a unique affine map $\pi\colon \Delta \to \mathbf{R}^d$ such that $\pi(a_j) = x_j$ for each j. Suppose that Δ_1 (respectively Δ_2) are the collections of vertices of Δ such that $\Delta_1 := \{a_j \mid x_j \in X_1\} = \pi^{-1}(X_1)$, similarly $\Delta_2 := \pi^{-1}(X_2)$. Then $\mathrm{conv}\,(X_i) = \pi(F_i)$ where F_i (for $i \in \{1,2\}$) is the face of Δ spanned by Δ_i and Radon's theorem is the statement that for each affine (linear) map $\pi\colon \Delta^{d+1} \to \mathbf{R}^d$ there exist two vertex disjoint faces F_1 and F_2 of $\Delta = \Delta^{d+1}$ such that $\pi(F_1) \cap \pi(F_2) \neq \emptyset$.

The usual proof of Radon's theorem is based on simple manipulation with affine dependences of points in $\mathbf{R}^d$, see [M02] or [Zi06]. Here we offer a different proof based on Colorful Carathéodory theorem which introduces a beautiful idea due to Karanbir Sarkaria [Sar00].

3.1. A criterion for $A \cap B \neq \emptyset$

It may look silly to ask for a criterion when two sets have a non-empty intersection. Indeed, it appears that one way or another this can achieved only by proving the existence of a point that belongs to both of these sets. Therefore it may indeed come as a surprise that a straightforward reformulation of the condition $A \cap B \neq \emptyset$ can be substantially easier to prove.

Since our main motivation is to prove statements like Radon's theorem (relation (2)), we assume that both A and B are subsets of $\mathbf{R}^d$.

Our first criterion is the obvious equivalence

$$A \cap B \neq \emptyset \Longleftrightarrow (A \times B) \cap D \neq \emptyset \tag{3}$$

where $D := \{(x,x)\}_{x \in \mathbf{R}^2}$ is the diagonal in $\mathbf{R}^2$. Geometrically more useful reformulation is given by the following equivalence

$$A \cap B \neq \emptyset \Longleftrightarrow (A * B) \cap D \neq \emptyset \tag{4}$$

where $A * B$ is the "join" of sets A and B defined as a subset of the join $\mathbf{R}^d * \mathbf{R}^d$ and D is again the diagonal $D \subset \mathbf{R}^d \times \mathbf{R}^d \subset \mathbf{R}^d * \mathbf{R}^d$. The join $L_1 * L_2$ of two subsets of an ambient Euclidean space $\mathbf{R}^N$ is the union

$$L_1 * L_2 := \bigcup_{x \in L_1, y \in L_2} [x, y] \tag{5}$$

of all line segments connecting a point $x \in L_1$ with a point $y \in L_2$. We always assume that L_1 and L_2 are "in general position" in the sense that different line segments $[x, y]$ and $[x', y']$ can intersect only in the end-points. Consequently, $\mathbf{R}^d * \mathbf{R}^d$ will be defined if we choose two isometric copies L_1 and L_2 of $\mathbf{R}^d$ in some ambient space $\mathbf{R}^N$, which are in general position, and declare that

$$\mathbf{R}^d * \mathbf{R}^d := L_1 * L_2. \tag{6}$$

A convenient way to make all this very concrete and easy to visualize is to choose the vector space $Mat_{(d+1)\times 2}(\mathbf{R})$ of all real $(d+1)\times 2$ matrices as the ambient space $\mathbf{R}^N$. Define the first copy L_1 of $\mathbf{R}^d$ inside $Mat_{(d+1)\times 2}(\mathbf{R})$ as the (affine) subspace of all matrices of the form

$$v_{(1)} := \begin{bmatrix} v & 0 \\ 1 & 0 \end{bmatrix}$$

where v is a (column) vector in $\mathbf{R}^d$. L_2 is defined similarly as the set of all matrices $v_{(2)}$ obtained from $v_{(1)}$ by interchanging the columns.

Then a typical element x of $L_1 * L_2 \subset Mat_{(d+1)\times 2}(\mathbf{R})$ is

$$x = tu_{(1)} + (1-t)v_{(2)} = \begin{bmatrix} tu & (1-t)v \\ t & 1-t \end{bmatrix}$$

for some elements $u, v \in \mathbf{R}^d$ and some $t \in [0,1]$. The product $L_1 \times L_2 \subset Mat_{(d+1)\times 2}(\mathbf{R})$ is defined as the set of all matrices of the form

$$y = \frac{1}{2}u_{(1)} + \frac{1}{2}v_{(2)} = \begin{bmatrix} \frac{1}{2}u & \frac{1}{2}v \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

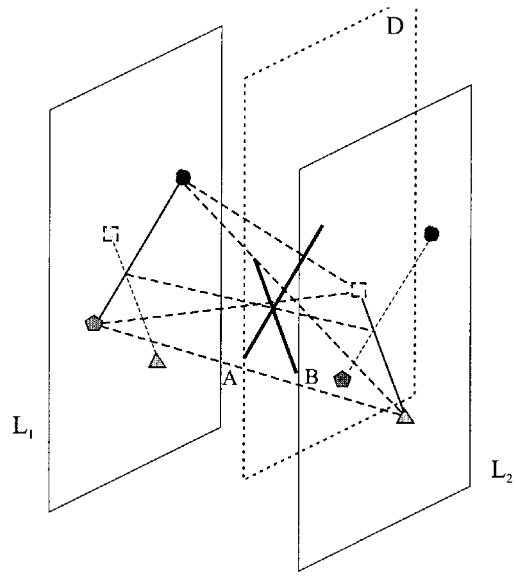


Fig. 4: An illustration of equivalence (4)

while the diagonal $D \subset L_1 \times L_2$ is described as the set of all elements of the form $\frac{1}{2}v_{(1)} + \frac{1}{2}v_{(2)}$ for some $v \in \mathbf{R}^d$.

After this clarification, the reader should be able, in light of (6), to use (or at least tolerate) expressions like $tu + (1-t)v \in \mathbf{R}^d * \mathbf{R}^d$ as abbreviations for $tu_{(1)} + (1-t)v_{(2)} \in L_1 * L_2$, the expression $\frac{1}{2}u + \frac{1}{2}u$ as an abbreviation for the diagonal element $\frac{1}{2}u_{(1)} + \frac{1}{2}u_{(2)} \in D$ etc. More importantly, the true meaning of the equivalence (4) now becomes perfectly clear.

As a test of understanding, the reader is invited to inspect Figure 4 where the non-empty intersection of two line segments A and B in the plane $\mathbf{R}^2 \equiv D$ is confirmed/detected as the non-empty intersection of the tetrahedron $A * B$ with the diagonal D.

Remark 3.2. A more natural ambient (affine) space for $L_1 * L_2$ then $Mat_{(d+1)\times 2}(\mathbf{R})$ is the space $Mat^{(-1)}_{(d+1)\times 2}(\mathbf{R})$ of all matrices with the sum of all entries in the last row equal to 1. Indeed, it is not difficult to show that this space is precisely the affine span of $L_1 * L_2$.

3.2. Proof of Radon's theorem

After necessary preparations we are ready to give a "one-line proof" of Radon's theorem.

Proof of Theorem 3.1. Given $X = \{x_j\}_{j=1}^{d+2} \subset \mathbf{R}^d$, let $(x_j)_{(1)}$ and $(x_j)_{(2)}$ be the corresponding elements in L_1 and L_2 respectively. Since

$$\frac{1}{2}x_j + \frac{1}{2}x_j = \frac{1}{2}(x_j)_{(1)} + \frac{1}{2}(x_j)_{(2)} \in D,$$

we observe that $D \cap [(x_j)_{(1)}, (x_j)_{(2)}] \neq \emptyset$ for each $j = 1, \dots, d+2$. Since

$$d + 1 = \dim(Mat^{-1}_{(d+1)\times 2}(\mathbf{R})) - \dim(\mathrm{D})$$

we are allowed to apply Corollary 2.3. Hence, for some $\epsilon_j \in \{+, -\}$,

$$\Delta \cap D \neq \emptyset \quad \text{where} \quad \Delta := \mathrm{conv}\,\{x_1^{\epsilon_1}, x_2^{\epsilon_2}, \dots, x_{d+2}^{\epsilon_{d+2}}\}. \tag{7}$$

Define $X^+ := \{x_j \mid \epsilon_j = +\}$ and $X^- := \{x_j \mid \epsilon_j = -\}$. Then $\{X^+, X^-\}$ is a partition of the set X which, in light of (7), satisfies the criterion (4). It follows that

$$\mathrm{conv}\,(X^+) \cap \mathrm{conv}\,(X^-) \neq \emptyset$$

as claimed by Radon's theorem. ∎

Remark 3.3. The reader familiar with the usual, elementary linear algebra proof of Radon's theorem, my wonder why we have decided to present a nice but more complicated proof above. The answer is that this proof can be easily modified to yield a proof of Tverberg's theorem!

4. Tverberg's theorem

Theorem 4.1. *A set $S \subset \mathbf{R}^d$ of size $(q-1)(d+1)+1$ can always be partitioned into q non-empty, disjoint subsets, $S = S_1 \cup \dots \cup S_q$, such that*

$$\bigcap_{j=1}^{q} \mathrm{conv}\,(S_j) \neq \emptyset. \tag{8}$$

The case $q = 2$ is Radon's theorem. We outline the proof of the case $q = 3$ and $d = 2$ which exhibits all important features of the general case. It can be easily upgraded, with minor modifications, to a complete proof of Theorem 4.1.

In the case $q = d+1 = 3$, Tverberg's theorem claims that each 7-element set S of points in the plane can be partitioned into 3 (disjoint) subsets, $S = S_1 \cup S_2 \cup S_3$, such that

$$\mathrm{conv}\,(S_1) \cap \mathrm{conv}\,(S_2) \cap \mathrm{conv}\,(S_3) \neq \emptyset. \tag{9}$$

In the spirit of Section 3.1, it would be reasonable to expect that there exists a criterion for testing the relation $A \cap B \cap C \neq \emptyset$. Indeed, such a criterion is provided by the equivalence

$$A \cap B \cap C \neq \emptyset \iff (A * B * C) \cap D \neq \emptyset. \tag{10}$$

In order for (10) to make sense, we should carefully reexamine the steps used for the development of criterion (4). For example, we use now the space $Mat_{3\times 3}(\mathbf{R})$ of all 3×3-matrices (or the corresponding codimension one subspace $Mat^{-1}_{3\times 3}(\mathbf{R})$) as

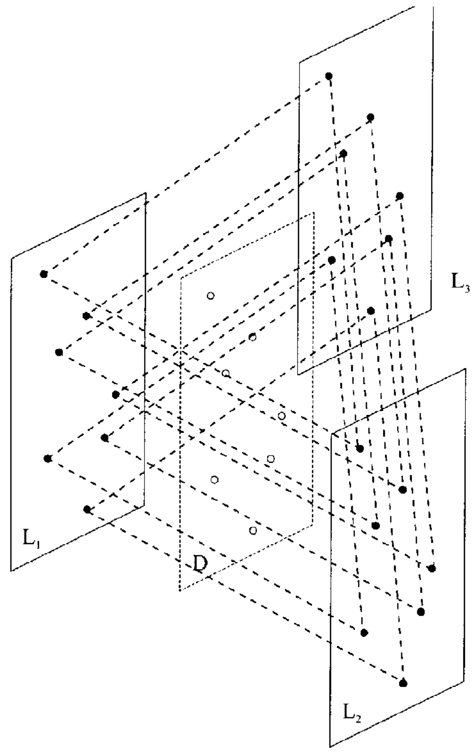


Fig. 5: The proof of the $q = d + 1 = 3$ case of Tverberg's theorem

the ambient space for $\mathbf{R}^2 * \mathbf{R}^2 * \mathbf{R}^2 \cong L_1 * L_2 * L_3$. The subspace L_1 of $Mat_{3\times 3}(\mathbf{R})$ is defined as the set of all matrices of the form

$$v_{(1)} := \begin{bmatrix} v & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

for some $v \in \mathbf{R}^2$. The subspaces L_2 and L_3 are defined similarly by interchanging the corresponding columns. The vectors in L_2 (respectively L_3), that correspond to $v \in \mathbf{R}^d$ are denoted by $v_{(2)}$ and $v_{(3)}$ respectively. Finally, the space $Mat^{-1}_{3\times 3}(\mathbf{R})$ is defined as the space of all (3×3)-matrices such that the sum of all entries in the bottom row is equal to 1.

Leaving out some details, we focus on the critical part of the argument. Let $S = \{x_1, x_2, \ldots, x_7\} \subset \mathbf{R}^2$. Then $\hat{S} = \{\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_7\} \subset D$ is the corresponding subset of the diagonal where

$$\hat{x}_j = \frac{1}{3}x_j + \frac{1}{3}x_j + \frac{1}{3}x_j = \frac{1}{3}(x_j)_{(1)} + \frac{1}{3}(x_j)_{(2)} + \frac{1}{3}(x_j)_{(3)}.$$

Let us apply Colorful Carathéodory theorem, in the form of Corollary 2.2, to the collection of 7 triangles $\Delta_j := \mathrm{conv}\,\{(x_j)_{(1)}, (x_j)_{(2)}, (x_j)_{(3)}\}$ and the subspace D, inside the ambient space $Mat^{-1}_{3\times 3}(\mathbf{R})$, Figure 5. By construction $\Delta_j \cap D = \{\hat{x}_j\} \neq \emptyset$ and since the dimensions match,

$$6 = \dim(Mat^{-1}_{3\times 3}(\mathbf{R})) - \dim(\mathrm{D}) = 8 - 2,$$

we observe that this application is legitimate. By Corollary 2.2, there is a function $f\colon \{1,2,3,4,5,6,7\} \to \{1,2,3\}$ such that

$$\operatorname{conv}\{(x_j)_{(f(j))}\}_{j=1}^{7} \cap D \neq \emptyset. \tag{11}$$

Let $S = S_1 \cup S_2 \cup S_3$ be the partition of S defined by $S_1 := \{x_j \in S \mid f(j) = 1\}, S_2 := \{x_j \in S \mid f(j) = 2\}, S_3 := \{x_j \in S \mid f(j) = 3\}$. This is one of Tverberg partitions. Indeed, in light of the criterion (10), the desired relation

$$\operatorname{conv}(S_1) \cap \operatorname{conv}(S_1) \cap \operatorname{conv}(S_1) \neq \emptyset$$

is a consequence of (11). ∎

5. Shadows of convex bodies

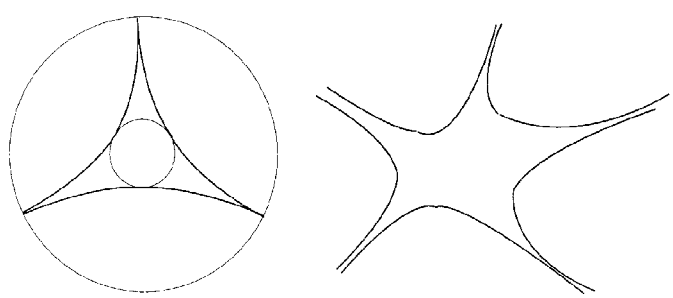
Both Radon's and Tverberg's theorem, as well as Colored Carathéodory theorem, are combinatorial statements about shadows (projections) of simplices. For example Tverberg's theorem says that for each affine map $A\colon \Delta^{(q-1)(d+1)} \to \mathbf{R}^d$ of a $(q-1)(d+1)$-dimensional simplex into a d-dimensional affine space, there exist q, vertex disjoint faces $\Delta_1, \ldots, \Delta_q$ of $\Delta^{(q-1)(d+1)}$ such that

$$\bigcap_{i=1}^{q} A(\Delta_i) \neq \emptyset.$$

There are other results about shadows of complex bodies which put emphasis on metric, or measure-theoretic aspect of the phenomenon. Here we briefly outline, in the form of a guide to the literature, a few of the most striking results of this type.

5.1. Milman amoebas

A very simple experiment shows that our usual intuition about higher dimensional (convex) bodies, based solely on our experience from dimensions 2 and 3, may not be fully adequate, especially when we are interested in metric and measure-theoretic properties of these objects.



(a) (b)

Fig. 6: With the increase of dimension simplices go "hyperbolic"!

Let $\Delta^d := \mathrm{conv}\,(E)$ be the d-simplex obtained as the convex hull of the standard orthonormal basis $E = \{e_0, e_1, \dots, e_d\}$ in $\mathbf{R}^d$. Then, Figure 6(a), the ratio $\frac{R_d}{r_d}$ tends to infinity as $d \mapsto +\infty$, where R_d and r_d are respectively the *circumradius* and the *inradius* of Δ^d. This indicates that the simplices in higher dimensions are much thinner and slimmer than one would expect.

A much more dramatic evidence for the "hyperbolicity" of high-dimensional convex bodies is provided by Vitali Milman [Mi96]. Milman begins with a question of how to draw an "accurate" 2 or 3-dimensional picture of a high-dimensional convex body K. He argues that the behavior (the rate of decay) of the volume of a parallel intersection $K \cap H$, with a moving hyperplane H, is of utmost importance for the geometry of K. His conclusion is that since this volume decays exponentially after passing the median level, a correct 2-dimensional picture is a star-shaped amoeba, as in Figure 6(b), which again emphasizes the hyperbolic distribution of the "substance" a convex body is made of. The reader will find much more detailed and convincing presentation of these ideas in [Mi96] and subsequent papers by the author.

5.2. Johnson-Lindenstrauss flattening lemma

While we are still under influence of Milman's vision of higher-dimensional convex bodies, let us test our "higher-dimensional geometric intuition" one more time with a phenomenon of different nature.

It is indeed not a surprise that a 3-dimensional simplex cannot be accurately (preserving the distances) represented in the plane. More generally, it is easy to verify that there does not exist an isometric embedding of vertices of a d-dimensional simplex into a k-dimensional Euclidean space for $k < d$.

Question. What if we allow a "mild distortion" say by 10%. In other words is it possible to find a $(1+\epsilon)$-embedding of all vertices $V = Vert(\Delta^d)$ of a d-simplex Δ^d in $\mathbf{R}^k$, for $k < d$. By definition, a map $f\colon V \to \mathbf{R}^k$ is a $(1+\epsilon)$-embedding if

$$1 - \epsilon \leq \frac{d(f(u), f(v))}{d(u, v)} \leq 1 + \epsilon$$

for each pair of distinct vertices $u, v \in V$.

A (perhaps) surprising answer is given by the so called "Johnson-Lindenstrauss flattening lemma" [M02].

Theorem 5.1. *Let X be an n-point set in a Euclidean space R^{n-1} and let $\epsilon \in (0, 1]$ be given. Then there exists a $(1+\epsilon)$-embedding of X into the Euclidean space $\mathbf{R}^k$ where $k = O(\epsilon^{-2} \log n)$.*

The reader is referred to [M02] for the proof, related facts and references. Here are some of the consequences which illuminate the surprising nature of this phenomenon.

The result says that, if we do not worry about a small distortion (say by 10%), any question about the distribution of distances between n points in $\mathbf{R}^n$ can be studied on n-element subsets of the space $\mathbf{R}^{O(\log n)}$ of (asymptotically) much smaller dimension. For example (cf. [M02], Section 15.2) to represent n points of $\mathbf{R}^n$ in a computer requires n^2 numbers. To store all the distances, the number is quadratic in n as well. A consequence of the flattening lemma is that we can store only $O(n \log n)$ numbers, and still be able to reconstruct any of the n^2 distances with the error at most 10%.

5.3. Dvoretzky, Brunn and slices of convex bodies

Projecting convex bodies from a high dimensional Euclidean space to a space of smaller dimension, that is "smashing convex bodies" by allowing them to fall on the "ground", is far from being the only amusing thing to do with convex objects. Equally interesting is *slicing of convex bodies*, typically by affine and in particular by hyperplane cuts. Perhaps two of the most famous results of this type are Dvoretzky's theorem and Brunn's inequality [M02].

Recall that a simplest instance of Ramsey's combinatorial theorem says that in each group of 6 or more people there must be either 3 persons that know each other or 3 persons so no two of them are acquaintances. More generally, for any integer n there is the smallest number $R(n)$ so that in each party of $R(n)$ people there is either a *clique* of size n, i.e. a group of n persons so that any two them are acquainted, or totally the opposite, a group of n people no two of them are known to each other. Let us mention that $R(5)$ is somewhere between 43 and 49 while $R(6)$ is according to some authors a reasonably small number which will remain unknown to humans until the end of our civilization!

Ramsey's theorem is a result of quite general type saying that some regularity must be observed in an object (i.e. a set with some structure) if it is of sufficiently large size. Once such an "order from chaos"-type result is observed, one would often like to know what is the "size" of the chaos that will produce the desired amount of regularity (e.g. what is the behavior of the function $R(n)$ in Ramsey's theorem).

Dvoretzky's theorem can be seen as a geometric instance of this general "order from chaos" phenomenon. We say that a convex body N is t-spherical (where $t > 1$) if there exists a ball $B = \{x \in \mathbf{R}^d \mid \|x - a\| \leq r\}$ such that $B \subset N \subset B'$ where $B' = \{x \in \mathbf{R}^d \mid \|x - a\| \leq r \cdot t\}$.

THEOREM 5.2. *For each natural number k and any real number $\epsilon > 0$, there exists an integer $D = D(k, \epsilon, d)$ such that for any d-dimensional, centrally symmetric convex body $K \subset \mathbf{R}^d$, there exists a k-dimensional linear subspace $L \subset \mathbf{R}^d$ such that the section $K \cap L$ is $(1 + \epsilon)$-spherical.*

As an illustration of how big the dimension d of the convex body must be in order to achieve 2-spherical sections ($\epsilon = 1$), it is known that under the most unfavorable circumstances k cannot be expected to be larger than $C \cdot \log d$ (for some constant C).

Contrary to Dvoretzky's theorem, Brunn's (theorem) inequality can be given a form of a result about quite concrete convex objects from our every day lives.

Consider a 3-dimensional (convex) loaf of bread (Figure 7) and slice it by parallel cuts into thin pieces. Measure the area of each of the pieces and replace it by a new piece which has the shape of a disc (more accurately a very thin cylinder) with the same area as the original piece. Put these new pieces back together in the same order along the same axes of symmetry.

Brunn's theorem claims that the new loaf of bread will be again a convex body. This amounts to saying that the radii of new disc-shaped slices behave as a concave function.

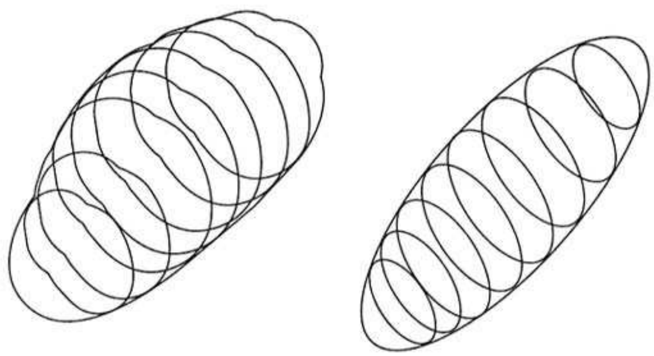


Fig. 7: Symmetrization of the slices of a convex bread

In the general case the "loaf of bread" is an n-dimensional convex body sliced by affine, k-dimensional planes orthogonal to a $(n-k)$-dimensional affine plane $L \cong \mathbf{R}^{n-k}$.

Theorem 5.3. (Brunn's theorem) *Let $K \subset \mathbf{R}^n$ be a compact convex body and $P(K)$ its projection where $P\colon \mathbf{R}^n \to \mathbf{R}^{n-k}$ is a non-degenerate linear map. Let $Vol_k(D)$ be the k-dimensional volume (Lebesgue measure) of a (measurable) set $D \subset \mathbf{R}^k$. Then the function $\phi^{1/k}\colon P(K) \to \mathbf{R}$ is concave where $\phi\colon P(K) \to \mathbf{R}$ is defined by*

$$\phi(x) := Vol_k(K \cap P^{-1}(x)).$$

The proof of Theorem 5.3 is given at the end of this section. The function ϕ defined in Theorem 5.3 is a special case of the "push-down" construction. Suppose that $P\colon \mathbf{R}^m \to \mathbf{R}^p$ is a linear projection map and let $P(K) \subset \mathbf{R}^p$ be the associated projection of a convex body $K \subset \mathbf{R}^m$. Given a function $f\colon K \to \mathbf{R}$, the associated "push down" is the function $Pf : P(K) \to \mathbf{R}$ defined by

$$Pf(x)\colon = \int_{P^{-1}\cap K} f\, d\mu_x$$

where μ_x is the Lebesgue measure defined on the (Euclidean) space $P^{-1}(x) \cong \mathbf{R}^{m-p}$. A more general version of Brunn's inequality says that the push down of a concave function is even more concave in a very precise sense.

DEFINITION 5.4. *A (non-negative) function $f\colon K \to \mathbf{R}$, defined on a convex set $K \subset \mathbf{R}^m$, is α-concave ($\alpha > 0$) if $f^{1/\alpha}$ is concave i.e. if*

$$f^{1/\alpha}(\lambda x + (1-\lambda)y) \geq \lambda f^{1/\alpha}(x) + (1-\lambda) f^{1/\alpha}(y)$$

for each $x, y \in K$ and $\lambda \in [0,1]$. As usual, if f is a continuous function it is sufficient to check the inequality

$$f^{1/\alpha}\left(\frac{x_1+x_2}{2}\right) \geq \frac{1}{2}(f^{1/\alpha}(x_1) + f^{1/\alpha}(x_2))$$

for each pair of points x_1, x_2 in the domain.

PROPOSITION 5.5. *If $f\colon K \to \mathbf{R}$ is α-concave then $Pf\colon P(K) \to \mathbf{R}$ is $(\alpha + m - p)$-concave where $P\colon \mathbf{R}^m \to \mathbf{R}^p$ is a linear projection map.*

Proof. We give a proof only in the case $m = 2$ and $p = 1$. This is not really a serious loss of generality considering that one has to check the inequality (evaluate the function) at points x_1, x_2 and $1/2(x_1 + x_2)$ which are collinear, i.e. one can focus on the affine line which contains these points (and this is where the condition $p = 1$ comes from).

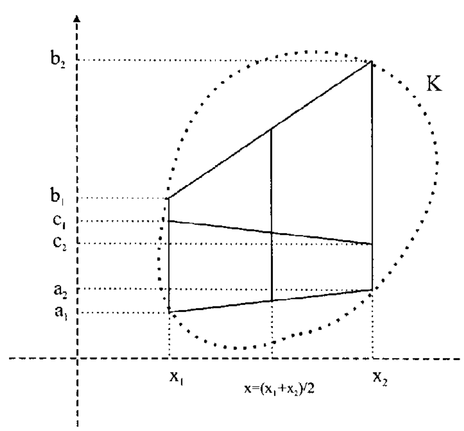


Fig. 8

The projection $P(K)$ is an interval. For $t \in P(K)$, let $I_t := K \cap P^{-1}(t)$. Given $x_1, x_2 \in P(K)$ let $I_{x_i} = [a_i, b_i] \times \{x_i\}$ (Figure 8). If $x := (x_1+x_2)/2, a := (a_1+a_2)/2$ and $b := (b_1 + b_2)/2$, then from convexity of K follows that $[a, b] \times \{x\} \subset I_x$. We are supposed to show that if $f\colon K \to \mathbf{R}$ is α-convex, then $Pf\colon P(K) \to \mathbf{R}$ is $(1+\alpha)$-convex function which amounts to proving the inequality

$$Pf^{1/1+\alpha}\left(\frac{x_1+x_2}{2}\right) \geq \frac{1}{2}(Pf^{1/1+\alpha}(x_1) + Pf^{1/1+\alpha}(x_2)). \tag{12}$$

This means that without any loss of generality we may assume that K is equal to the trapeze depicted in Figure 8, and as a consequence $I_x = [a, b] \times \{x\}$.

Choose $c_i \in [a_i, b_i]$ so that

$$\int_{a_i}^{c_i} f(t, x_i)\, dt = \int_{c_i}^{b_i} f(t, x_i)\, dt = \frac{1}{2}\int_{a_i}^{b_i} f(t, x_i)\, dt.$$

Divide the convex body K into the upper part K_{up}, above the line determined by points (x_i, c_i), $i = 1, 2$ and respectively the lower part K_{down} below this line.

Let $f_1 = f|_{K_{up}}$ and $f_2 = f|_{K_{down}}$ be the restrictions of the function f on these convex bodies, in particular $Pf_i(x_j) = (1/2)Pf(x_j)$. Let us observe that the inequality (12) follows from the corresponding inequalities for functions (convex bodies) f_1 and f_2 (respectively K_{up} and K_{down}). Indeed,

$$Pf^{1/1+\alpha}(x) = (Pf_1(x) + Pf_2(x))^{1/1+\alpha} \geq 2^{1/1+\alpha}\frac{Pf_1^{1/1+\alpha}(x) + Pf_2^{1/1+\alpha}(x)}{2}$$

by the concavity of the function $y = x^{1/1+\alpha}$ and, since by the assumption the functions Pf_i are $(1+\alpha)$-concave,

$$Pf_j^{1/1+\alpha}(x) \geq \frac{Pf_j^{1/1+\alpha}(x_1)) + Pf_j^{1/1+\alpha}(x_1)}{2} = \frac{Pf^{1/1+\alpha}(x_1) + Pf^{1/1+\alpha}(x_2)}{2\cdot 2^{1/1+\alpha}}$$

which leads to the desired inequality (12).

The process can be iterated, i.e. both convex bodies K_{up} and K_{down} can be further subdivided which leads to the sequence of thinner and thinner trapezes where the inequality analogous to (12) has to be established.

This means that eventually, by an approximation argument, the function f can be assumed to be constant along the vertical intervals I_t which means that $Pf(t) = L(t)\cdot f(t)$ where L is the linear function $L(t) := m(I_t)$ measuring the length of intervals I_t. Being linear, the function L is concave (or 1-concave) while f is by assumption α-concave. Consequently the desired conclusion about Pf follows from the following lemma.

Lemma 5.6. *Suppose that $f, g\colon K \to \mathbf{R}$ are both defined on a compact, convex set $K \subset \mathbf{R}^m$. Moreover assume that the function f is α-concave and the function g is β-concave for some $\alpha, \beta > 0$. Then the function $h := f\cdot g$ is $(\alpha+\beta)$-concave.*

Proof. An easy application of Hölder's inequality

$$a_1b_1 + a_2b_2 \leq \left(\frac{a_1^p + a_2^p}{2}\right)^{1/p}\left(\frac{b_1^q + b_2^q}{2}\right)^{1/q}$$

where $p = (\alpha+\beta)/\alpha$ and $q = (\alpha+\beta)/\beta$, leads to the desired result. Indeed, if $x_1, x_2 \in K$ then

$$\frac{(f(x_1)g(x_1))^{\frac{1}{\alpha+\beta}} + (f(x_2)g(x_2))^{\frac{1}{\alpha+\beta}}}{2} \leq$$

$$\leq \left[\frac{f(x_1)^{1/\alpha} + f(x_2)^{1/\alpha}}{2}\right]^{\frac{\alpha}{\alpha+\beta}}\left[\frac{f(x_1)^{1/\alpha} + f(x_2)^{1/\alpha}}{2}\right]^{\frac{\alpha}{\alpha+\beta}}$$

which in light of the fact that f is α-concave and g is β-concave yields,

$$\frac{(f(x_1)g(x_1))^{\frac{1}{\alpha+\beta}} + (f(x_2)g(x_2))^{\frac{1}{\alpha+\beta}}}{2} \leq$$

$$\leq \left[f\left(\frac{x_1+x_2}{2}\right)\right]^{1/(\alpha+\beta)} \left[g\left(\frac{x_1+x_2}{2}\right)\right]^{1/(\alpha+\beta)}.$$

Proof of Theorem 5.3. Characteristic function χ_K of K is α-concave for each $\alpha > 0$. By Proposition 5.5 the function ϕ is $(k+\alpha)$-concave for each $\alpha > 0$, hence the function $\phi^{1/k}$ is concave. ∎

6. Appendix

For the benefit of a less experienced reader, we include here a short glossary of some basic terms that are used in this article. Such reader should also be informed that a harmonic balance between geometric intuition and elements of linear algebra will suffice for the study of the geometry of convex sets in the linear (Euclidean) space $\mathbf{R}^n$. A non-initiated reader will soon discover that drawing pictures in the plane and using them for the interpretation of higher dimensional phenomena is a legitimate strategy, if we don't forget that the intuitive and suggestive geometric language can be (and occasionally must be) supplemented with the precise algebraic calculations. In time, these calculations become partly routine and the reader will feel more and more comfortable when relying on the geometric descriptions alone.

Glossary:

Euclidean space $\mathbf{R}^d$: A d-dimensional, real, linear (vector) space equipped with the Euclidean metric. Points in $\mathbf{R}^d$ are n-tuples $(a_1, a_2, \dots, a_n)$ of real numbers (row vectors) with the addition and multiplication by scalars defined by point-wise addition and multiplication,

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n),$$
$$\lambda(a_1, \dots, a_n) = (\lambda a_1, \dots, \lambda a_n).$$

The Euclidean metric (norm) is defined by

$$d(a, b) = \sqrt{(a_1 - a_2)^2 + \cdots + (a_n - b_n)^2}.$$

For $d = 1, 2$ or 3 we obtain a line, a plane and the 3-space.

Linear combination: A linear combination of vectors $x_1, x_2, \dots, x_m \in \mathbf{R}^d$ is the vector $x = \alpha_1 x_1 + \cdots + \alpha_m x_m$ for some choice of scalars (real numbers) $\alpha_1, \dots, \alpha_m$.

Barycenter: The barycenter of a finite collection of points $x_1, x_2, \dots, x_m \in \mathbf{R}^d$ is the point $x := (1/m)(x_1 + x_2 + \cdots + x_m)$, that is the linear combination of vectors $x_1, \dots, x_m$ where $\alpha_1 = \cdots = \alpha_m = 1/m$. More generally, the weights α_i may be allowed to be different (see convex combination).

Convex combination: A linear combination $v = \lambda_0 v_0 + \lambda_1 v_1 + \cdots + \lambda_k v_k$ of vectors $v_0, v_1, \ldots, v_k$ in some Euclidean space $\mathbf{R}^d$ in which the sum of the coefficients λ_j is equal to 1 and $\lambda_j \geq 0$ for each $j = 0, 1, \ldots, k$. The numbers λ_j can be interpreted as *weights* put at the corresponding points v_j in which case the convex combination v is the associated barycenter.

Affine combination: If the condition that the "weights" λ_j are positive is removed from the definition of the convex combination, the corresponding linear combination of vectors is called an *affine combination*.

Lines and line segments: Given two distinct points $x, y \in \mathbf{R}^d$, the unique line $p = p(x, y)$ which contains x and y is described as the set $p := \{\lambda x + (1 - \lambda)y \mid \lambda \in \mathbf{R}\}$. The associated line segment with end-points x and y is the set $[x, y] := \{\lambda x + (1 - \lambda)y \mid 0 \leq \lambda \leq 1\}$.

Convex set, convex hull: A set $K \subset \mathbf{R}^d$ is *convex* if together with each two points $x, y \in K$ it contains the associated line segment $[x, y] = \{\lambda x + (1-\lambda)y \mid 0 \leq \lambda \leq 1\}$. Convex hull $\mathrm{conv}\,(S)$ of a set $S \subset \mathbf{R}^d$ is the minimal convex set K such that $S \subset K$. A basic fact is that $x \in \mathrm{conv}\,(S)$ if and only if x is a convex combination $x = \lambda_0 x_0 + \lambda_1 x_1 + \cdots + \lambda_k x_k$ for some collection of points $x_j \in S$ or in other words, $\mathrm{conv}\,(S)$ is the set of all convex combinations of elements from S.

Affine set (space), affine hull: An affine subspace $L \subset \mathbf{R}^d$ is defined as the set which has the property that for each pair of distinct points $x, y \in L$, the associated line $p(x, y)$ is a subset of L. The affine hull (affine span) $\mathrm{affine}(S)$ of a set $S \subset \mathbf{R}^d$ is the minimal affine space containing S as a subset. It is not difficult to show that $x \in \mathrm{affine}(S)$ if and only if x is an affine combination of some points from S.

Affine (in)dependence: A collection of points $x_0, x_1, \ldots, x_k$ in $\mathbf{R}^d$ is affinely independent if a point $x \in \mathrm{affine}\{x_j\}_{j=0}^k$ can be expressed as an affine combination of points $x_0, x_1, \ldots, x_k$ in only one way. This is equivalent to the condition that if $0 = \lambda_1 x_1 + \cdots + \lambda_k x_k$ is a linear combination such that $\lambda_1 + \cdots + \lambda_k = 0$, then $\lambda_j = 0$ for each j. This condition is easily checked to be equivalent to the linear independence of vectors $\{x_1 - x_0, \ldots, x_k - x_0\}$ which in turn implies that $\mathrm{affine}\{x_j\}_{j=0}^k$ is a translate of a linear, k-dimensional subspace of $\mathbf{R}^d$.

Simplex: The definition of simplex is already given in Section 1. According to this definition, a m-dimensional simplex $\Sigma = \Sigma^m := \mathrm{conv}\,\{a_0, a_1, \ldots, a_m\}$ in $\mathbf{R}^d$ is well defined only if the corresponding *vertices* a_j are affinely independent. In this case for each point $x \in \Sigma$ there is a unique choice of weights $\alpha_i \in [0, 1]$, $\alpha_0 + \cdots + \alpha_m = 1$, such that x is the corresponding barycenter (convex combination) $x = \alpha_0 a_0 + \cdots + \alpha_m a_m$.

Degenerate simplex: The convex hull $\mathrm{conv}\,\{a_0, a_1, \ldots, a_m\}$ of a finite sets of points is often referred to as a *degenerate simplex* if the points a_j are not affinely independent.

Join $A * B$ of sets (spaces): Formal definition of the join of two (or more) subsets of some Euclidean space is given in Sections 3.1 and 4. The definition of $A * B$ can be simplified if both A and B are convex subset of $\mathbf{R}^d$ such that no two line segments $[a_1, b_1]$ and $[a_2, b_2]$, where $a_i \in A$ and $b_i \in B$, have a common interior point. It turns out that in this case

$$A * B = \mathrm{conv}\,(A \cup B).$$

What we described and used here is often referred to as the *geometric join* of subsets A and B of an Euclidean space $\mathbf{R}^d$. There are other, closely related concepts, ([M03] Section 4.2), among them the join of simplicial complexes, the join of topological spaces etc. The reader is referred to [Živ96], [Živ98], and [M03] for a more complete discussion, examples and applications.

REFERENCES

[Bar82] I. Bárány, *A generalization of Carathéodory's theorem,* Discr. Math. **40** (1982), 141–152.

[GrMi87] M. Gromov, V. D. Milman, *Generalization of spherical isoperimetric inequality to uniformly convex Banach spaces,* Compositio Math. **62**:3 (1987), 263–282.

[Sar00] K. S. Sarkaria, *Tverberg theorem via number fields*, Israel J. Math. **79** (1992), 317–320.

[Tve] H. Tverberg, *A generalization of Radon's theorem*, J. London Math. Soc. **41** (1966), 123–128.

[T-V] H. Tverberg and S.T. Vrećica, *On generalizations of Radon's theorem and the ham sandwich theorem*, Europ. J. Comb. **14** (1993) 259–264.

[M02] J. Matoušek, *Lectures on Discrete Geometry*, Graduate Texts in Mathematics 212, Springer, Berlin 2002.

[M03] J. Matoušek, *Using the Borsuk-Ulam Theorem; Lectures on Topological Methods in Combinatorics and Geometry*, Springer Universitext, Berlin 2003.

[Mi96] V. D. Milman, *Surprising geometric phenomena in high dimensional convexity theory*, A plenary talk on European Congress of Mathematicians, Budapest 1996.

[Kal01] G. Kalai, *Combinatorics with geometric flavor: Some examples*, In: *Visions in Mathematics Towards 2000 (GAFA, special volume), part II*, pp. 742–792, Birkhäuser, Basel 2001.

[V-Ž] S. T. Vrećica and R. T. Živaljević, *New cases of the colored Tverberg theorem*, Jerusalem Combinatorics '93 (H. Barcelo, G. Kalai, eds.), Contemporary Mathematics, A.M.S., Providence, 1994.

[Ž-V] R. T. Živaljević and S.T. Vrećica, *The colored Tverberg problem and complexes of injective functions*, J. Comb. Theory, Ser. A **61** (1992), 309–318.

[Živ96] R. T. Živaljević, *User's guide to equivariant methods in combinatorics*, Publ. Inst. Math. **73** (1996), 114–130.

[Živ98] R. T. Živaljević. *User's guide to equivariant methods in combinatorics II,* Publ. Inst. Math. **78** (1998), 107–132.

[Živ04] R. T. Živaljević. *Topological methods*, In: *CRC Handbook of Discrete and Computational Geometry* (new edition), J.E. Goodman, J. O'Rourke (eds.), Boca Raton 2004.

[Zi06] G. M. Ziegler. *Lectures on Polytopes*, Graduate Texts in Mathematics 152, Springer-Verlag, Berlin 1995, Revised edition 2006.

Mathematical Institute SANU,
Knez Mihailova 35/1, P.O.Box 361,
11001 Belgrade, Serbia.

E-mail: `rade@turing.mi.sanu.ac.yu`