Discrete Geometry II

http://wikis.fu-berlin.de/display/discgeom/Discrete+Geometry+II+SS+2014

— Preliminary Lecture Notes (without any guarantees) —

Prof. Günter M. Ziegler

Fachbereich Mathematik und Informatik, FU Berlin, 14195 Berlin
ziegler@math.fu-berlin.de

FU Berlin, Summer Term 2014

This is the second in a series of three courses on Discrete Geometry. We will get to know fascinating geometric structures such as configurations of points and lines, hyperplane arrangements, and in particular polytopes and polyhedra, and learn how to handle them using modern methods for computation and visualization and current analysis and proof techniques. A lot of this looks quite simple and concrete at first sight (and some of it is), but it also very quickly touches topics of current research.

For students with an interest in discrete mathematics and geometry, this is the starting point to specialize in discrete geometry. The topics addressed in the course supplement and deepen the understanding of discrete-geometric structures appearing in differential geometry, optimization, combinatorics, topology, and algebraic geometry. To follow the course, a solid background in linear algebra is necessary. Some knowledge of combinatorics and geometry is helpful.

–

Basic Literature

[1] Peter M. Gruber. *Convex and Discrete Geometry*, volume 336 of *Grundlehren Series*. Springer, 2007.

[2] Peter M. Gruber and Jörg Wills, editors. *Handbook of Convex Geometry*. North-Holland, Amsterdam, 1993. 2 Volumes.

[3] Branko Grünbaum. *Convex Polytopes*, volume 221 of *Graduate Texts in Math.* Springer-Verlag, New York, 2003. Second edition prepared by V. Kaibel, V. Klee and G. M. Ziegler (original edition: Interscience, London 1967).

[4] Jiří Matoušek and Bernd Gärtner. *Understanding and Using Linear Programming*. Universitext. Springer, 2007.

[5] Günter M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. Revised edition, 1998; seventh updated printing 2007.

Contents

A rough schedule, which we will adapt as we move along:

1. 1. Linear programming 1.1 On the diameter of polyhedra . . . . April 15
2. 1.2 Linear programming (Discrete geometry version) . . . . April 17
3. … and dual simplex algorithm . . . . April 22
4. 1.2.2 (Linear algebra version), LP duality . . . . April 24
5. 1.3 Further notes on linear programming . . . . April 29
6. 2. Convex bodies, volumes, and roundness 2.1 Basics . . . . May 6
7. … and examples. 2.2 Topological properties: boundary, interior . . . . (?/BMS) May 8
8. (cont.d) . . . . May 13
9. 2.3 Support and separation . . . . May 15
10. and Minkowski's theorem . . . . (?/NYC) May 20
11. 2.4 Spectrahedra . . . . May 22
12. 2.5 Löwner–John ellipsoids and roundness . . . . May 27
13. 2.6 Approximation of convex bodies by polytopes . . . . May 29
14. 2.7 On the difficulty of volume computation . . . . June 3
15. 2.8 Polarity and the Mahler conjecture . . . . June 5
16. 2.9 Volume, mixed volumes, and valuations . . . . June 10
17. 2.10 Geometric inequalities: Brunn–Minkowski and Alexandrov–Fenchel . . . . . . June 12
18. and applications . . . . June 17
19. 2.11 Isoperimetric inequalities . . . . June 19
20. 2.12 Measure concentration and high-dimensional effects . . . . June 24
21. 3. Geometry of Numbers 3.1 Lattices and lattice points . . . . June 26
22. 3.2 Minkowski's (first) theorem . . . . July 1
23. 3.3 Successive minima . . . . July 3
24. 3.4 Lattice points in convex bodies . . . . July 8
25. . . . . (?/SA) July 10
26. . . . . July 15
27. . . . . July 17

0 Introduction

What's the goal?

This is a second course in a large and interesting mathematical domain commonly known as "Discrete Geometry". This spans from very classical topics (such as regular polyhedra – see Euclid's *Elements*) to very current research topics (Discrete Geometry, Extremal Geometry, Computational Geometry, Convex Geometry) that are also of great industrial importance (for Computer Graphics, Visualization, Molecular Modelling, and many other topics).

My goal will be to develop these topics in a three-semester sequence of Graduate Courses in such a way that

- you get an **overview** of the field of Discrete Geometry and its manifold connections,
- you learn to understand, analyze, visualize, and confidently/competently argue about the basic **structures** of Discrete Geometry, which includes
  - point configurations/hyperplane arrangements,
  - frameworks
  - subspace arrangements, and
  - polytopes and polyhedra,
- you learn to know (and appreciate) the most important **results** in Discrete Geometry, which includes both simple & basic as well as striking key results,
- you get to learn and practice important **ideas and techniques** from Discrete Geometry (many of which are interesting also for other domains of Mathematics), and
- You learn about current **research topics** and problems treated in Discrete Geometry.

In this second course of the sequence, we will in particular treat the relationship between

- "discrete objects" (such as polytopes and polyhedra, but also lattices and lattice points) and
- "general objects" (such as convex bodies)

in terms of various notions of diameter, volume, and roundness.

This will not only be interesting *per se*, but also lead us to some major theorems and insight (e.g. on such fundamental notions as *volume*), but also to major applications (e.g. on sphere packings, which is in turn important for coding theory).

1 Linear programming and some applications

1.1 On the diameter of polyhedra

Let's consider a polyhedron of dimension d with n facets; let's call it an (d, n)-polyhedron.
Careful: Want to look at *pointed* polyhedron, $n \geq d$, which has a vertex, so the *lineality space* is trivial.
The *Hirsch conjecture* from 1957 is the *false* (!) statement that the edge-graph of any (d, n)-polyhedron has diameter at most $n - d$. This was disproved for unbounded polyhedra by Klee & Walkup [3] in 1967 and in general by Santos [5] in 2012. The *polynomial Hirsch conjecture* remains open: It might still be that the maximal diameter, $\Delta(d, n)$, satisfies $\Delta(d, n) \leq d(n - d)$ for all $n \geq d \geq 1$.
We will, nevertheless, see why from a "linear programming point of view" the bound $n - d$ looks natural, and even more so, why this is a relevant parameter.

Exercise 1.1. Show that $\Delta(2, n) \leq n - 2$ and $\Delta(3, n) \leq n - 3$, and that both inequalities are sharp (that is, hold with equality for $n \geq 2$ resp. $n \geq 3$).

Up to recently, the best upper bound for the diameters of polyhedra was provided by Kalai & Kleitman in a striking two page paper [2] in 1992:

$$\Delta(d, n) \leq n^{\log(d)+2},$$

which was improved only slightly by Kalai [1] to

$$\Delta(d, n) \leq n^{\log(d)+1},$$

where throughout "$\log$" denotes the binary logarithm (i.e., base 2). However, just a few weeks ago Mike Todd (Cornell University) in a 4-page paper [6] sharpened the Kalai–Kleitman analysis to obtain

$$\Delta(d, n) \leq (n - d)^{\log(d)} = d^{\log(n-d)},$$

which indeed is sharp for $d = 1$ and $d = 2$.
In class, we will go through the arguments of Todd [6] (and thus, in particular, the idea of Kalai & Kleitman [2]).

End of class on April 15

1.2 Geometry of linear programming and pivot rules

1.2.1 Linear programming (Discrete Geometry version)

Any system $Ax \leq b$ with $A \in \mathbb{R}^{n \times d}$, $b \in \mathbb{R}^n$ defines a polyhedron $P \subseteq \mathbb{R}^d$ with $\dim P \leq d$ and $\#\text{facets} \leq n$.

Without loss of generality we may assume that $\text{rank}A = d$, that is the system $Ax \leq b$ has a subsystem that defines an orthant, so in particular P is either pointed (has a vertex), or is empty.

Without loss of generality (theoretically, this may be harder to compute) we may assume that $\dim P = d$, so the polyhedron is full-dimensional. Moreover, we want to get our system into the form

$$Ax \leq b, \ -x \leq 0$$

with $b > 0$ componentwise. For this we have to solve a "Phase I" problem that finds a vertex x_0 of the polyhedron, and then do a coordinate transformation that moves the vertex x_0 to 0 and transforms a system of inequalities that are tight at x_0 to the positive orthant system $x \geq 0$.

With a linear objective function we have a system of the form

$$\begin{array}{rcl} \max\ c^t x & & \\ Ax & \leq & b \\ x. & & \end{array}$$

Example:

$$\begin{array}{lrcl} \max & y & & \\ & x - y & \leq & 2 \\ & -x + y & \leq & 1 \\ & x + 2y & \leq & 7 \\ & -x \quad & \leq & 0 \\ & y & \leq & 0. \end{array}$$

Geometric description of the polyhedron

- P is a full-dimensional polyhedron, with $\leq n$ facets, given in $\mathcal{H}$-description-
- We have a linear objective function, which might be assumed to be the last coordinate x_d, to be maximized (or in other situations: minimized).
- We assume that the polyhedron is simple, the system is in general position (this may be achieved by perturbing the right-hand sides: Exercise!).
- Any $d \times d$ full rank subsystem $A'x \leq b'$ defines a *generalized orthant*, which up to an affine transformation is equivalent to the standard positive orthant "$x \geq 0$."
- Any generalized orthant defines a point (the unique solution of $A'x = b'$) and d *rays* (by fixing all the d inequalities by one, and letting the *slack* in the last one get large).
- A generalized orthant is *feasible* if the point it defines by $A'x = b'$ is *feasible* (defines all inequalities, not only those in the subsystem). Note that this does not depend on the objective function.
- A generalized orthant is *dual feasible* if sliding along any of its rays does not improve the objective function. Note that this does not depend on the right-hand side vector b.
- A generalized orthant is *optimal* if it is both feasible and dual feasible.
- Any optimal generalized orthant defines an optimal solution of the linear program.

... and what the *primal simplex algorithm* does on it:

- We assume that after preprocessing (known as "Phase I") we have $-x \leq 0$ as a feasible generalized orthant, and in particular $x_0 = 0$ as a feasible starting vertex.

- If the generalized orthant is dual feasible, DONE with optimal solution.
- Select an improving ray, and slide along the ray. (Along the ray one inequality of the orthant is not tight any more; the objective function improves along the ray.)
- If the objective function improves without bound along the ray, DONE with optimal solution.
- Otherwise along the way we hit a bound, that is, a new facet, whose inequality completes a new feasible generalized orthant. REPEAT.

The process stops in finite time, since in every step we improve the objective function (no cycles) and there are only finitely many orthants — not more than $\binom{n}{d}$. (A better bound is obtained from the upper bound theorem — need a version for unbounded polyhedra: Exercise!)

End of class on April 17

Alternatively, here is what the *dual simplex algorithm* does on a linear program:

- We assume that after preprocessing (known as "Phase I") we have found a dual feasible generalized orthant, which in particular defines a current solution (vertex of the system, but not necessarily of the polyhedron).
- If the generalized orthant is feasible, DONE with optimal solution.
- Select an inequality violated by the current solution.
- If the violating inequality hits none of the rays of the current generalized orthant, then DONE with proof that the system is infasible.
- Otherwise construct a new dual feasible generalized orthant whose current solution gives a better upper bound on the maximum of the system. REPEAT.

The process stops in finite time, if we take care that in every step we improve the current upper bound on the objective function values on the polyhedron (no cycles) and there are only finitely many generalized orthants.

End of class on April 22

1.2.2 Linear programming (Numerical Linear Algebra version)

We write down two linear programs, in the following form.
The *primal linear program* is

$$
\begin{array}{lrcl}
(P) \qquad & \max\ c^t x & & \\
& Ax & \leq & b \\
& x & \geq & 0.
\end{array}
$$

The associated *dual linear program* is

$$
\begin{array}{lrcl}
(D) \qquad & \min\ b^t y & & \\
& A^t y & \geq & c \\
& y & \geq & 0.
\end{array}
$$

Lemma 1.2 (Weak Duality Theorem). *If for a primal-dual pair of linear programs x_0 is a feasible solution for the primal (P) and y_0 is a feasible solution for the dual (D), then*

$$c^t x_0 \leq b^t y_0.$$

In particular, the maximum of (P) is smaller or equals to the minimum of (D).

Proof. We compute

$$c^t x_0 \le (A^t y_0)^t x_0 = y_0^t (A x_0) \le y_0^t b = b^t y_0.$$

□

The linear programs are then, by introduction of *slack variables*, converted into systems of linear equations, to be solved in non-negative variables.

Thus the primal linear program becomes

$$\begin{array}{lrcl} (P) \qquad & \max\ c^t x + 0^t \hat{x} & = & \gamma \\ & Ax + I_n \hat{x} & = & b \\ & x \ge 0,\ \hat{x} \ge 0 & & \end{array}$$

This system has an "obvious" current solution, given by $x \equiv 0$ (the "non-basic variables" are set to 0: these correspond to the inequalities that define the current generalized orthant), $\hat{x} = b$ (the "basic variables" are uniquely determined). This starting solution has the value $\gamma = 0$.

These systems are manipulated by *row operations*, which do not change the solution space. Thus after a number of steps we still have the system in the form

$$\begin{array}{lrcl} (P) \qquad & \max\ \bar{c}^t x_N + 0^t x_B & = & \bar{\gamma} \\ & \bar{A}_N x_N + I_n x_B & = & \bar{b} \\ & x_N \ge 0,\ x_B \ge 0 & & \end{array}$$

Here the columns have been resorted, to keep the "basic variables" and the "non-basic variables" together, that is, the index sets B and N together give the set of all columns labelled by $B \cup N = \{1, 2, \dots, d{+}n\}$. The coefficients in the system are $\bar{A}_N = A_B^{-1} A_N$, and $\bar{b} = A_B^{-1} b$. The objective function has been rewritten in terms of the non-basic variables. Its coefficients

$$\bar{c}_N^t = c_N^t - c_B^t A_B^{-1} A_N$$

are known as the *reduced costs*: in the geometric interpretation they give the slopes of the rays of the current generalized orthant.

The current solution is given by $x_N \equiv 0$, which uniquely determines the non-basic variables to be $x_B = \bar{b} = A_B^{-1} b$.

Thus the (current solution of the) system is *feasible* if $\bar{b} \ge 0$, and it is *dual feasible* if $\bar{c}_N \le 0$.

A similar treatment/computation can be done for the dual system (D).

Lemma 1.3. *For any pair of primal linear program (P) and its dual program (D) in the equation form given above,*

- *the bases B for the system (P) are in bijection with the non-bases N of the system (D);*
- *the feasible bases for (P) are in bijection with the dual-feasible non-bases for (D);*
- *etc.*

Proof. This rests on the observation that in the $(n + d) \times (n + d)$ matrix

$$\begin{pmatrix} A & I_n \\ -I_d & A^t \end{pmatrix}$$

the row space spanned by the first n rows is the orthogonal complement of the space spanned by the last d rows. □

Theorem 1.4 (Duality Theorem for Linear Programming)**.** *If a primal linear program* (P) *and its dual* (D) *are both feasible, then they have optimal solutions* x^* *and* $y*$*, and these have the same optimal value.*
If one of the programs is not feasible, then the other one is either infeasible as well, or it is unbounded.

Proof. The optimal solutions exist, since the Simplex Algorithm will find it! □

From the geometry of an optimal basis/optimal generalized orthant, we also get *complementary slackness*: If in the optimal solution an inequality is not tight, then the corresponding variable in the dual program is zero; if a variable is positive, then the corresponding dual inequality has to be tight. This can also be seen from analysis of the inequalities in the proof of the Weak Duality Theorem.
The optimal solution to a linear program can be computed *efficiently*:

In Practice there are commercial, as well as non-commercial, software libraries for linear programming, which include implementations of the Primal Simplex Algorithm, the Dual Simplex Algorithm, as well as other methods (such as *Interior Point Methods*) which will solve to optimality practically every linear program that appears in practice.

In Theory there are two different computational models:

In the bit model the "Ellipsoid Method" (which will appear later in this course) is a polynomial time method for solving linear programs, whose running time is polynomial in the bit-size of the input. This method is theoretically very important, but has not been implemented in practice.

In the unit cost model the Simplex Algorithm with a suitable choice of variable selections ("pivot rule") may be polynomial — but this has not been proven. Indeed, we do not even know whether in general there is any short (i.e. polynomially many edges) path from a given starting vertex of the program to the optimal vertex. The best upper bound is the $n^{\log_2 d}$ upper bound discussed at the beginning of this course — and this bound is *not* a polynomial in n and d. An upper bound of the type $d(n-d)$ might exist, but has not been proven.

Thus the complexity of Linear Programming, and in particular of the Simplex Algorithm, is a major open problem both for Optimization, and for Discrete Geometry!

End of class on April 24

1.3 Further Notes on Linear Programming

Let's step away from the simplex algorithm, and let's look at the problem itsself — and let's assume we have a solution method (algorithm, perhaps software) that solves the problem, but which we can treat as a "black box." This is the *oracle* view, which has become popular in optimization, with grave consequences for (computational) discrete and convex geometry: well-defined input, well-defined output; estimate complexity
Examples:
LP-OPTIMIZATION problem/oracle:

INPUT: $d \geq 1$, $n \geq 1$, $A \in \mathbb{Q}^{n \times d}$, $b \in \mathbb{Q}^n$, $c \in \mathbb{Q}^d$

TASK: $\max c^t x : \; Ax \leq b, \; x \geq 0$

OUTPUT: optimal solution $x^* \in \mathbb{Q}^d$, with certificate (basis)
or information that problem is infeasible, with certificate (basis & inequality),
or information that problem is unbounded, with certificate (basis & ray).

LP-FEASIBILITY problem/algorithms/oracle:

INPUT: $d \geq 1$, $n \geq 1$, $A \in \mathbb{Q}^{n \times d}$, $b \in \mathbb{Q}^n$

TASK: find $x : \; Ax \leq b, \; x \geq 0$

OUTPUT: feasible solution $x^* \in \mathbb{Q}^d$, with certificate (basis)
or information that problem is infeasible, with certificate (basis & inequality).

Note: Any algorithm for solving **LP-OPTIMIZATION** can be used to solve **LP-FEASIBILITY**. We will see that the other direction "works as well."

Note: Two algorithms we know/could work out for **LP-OPTIMIZATION**: Fourier–Motzkin elimination (see Discrete Geometry I), and the Simplex Algorithm.

1.3.1 Complexity issues

Could it be that the solution exists, but it is too large (or too small) to write down in reasonable time?

Real input/solutions don't make sense, or need work to make sense of.

Recommended reading: Lovász' lecture notes [4].

Could get answer from Fourier–Motzkin elimination.

Here: get answer from simplex and Cramer's rule and Hadamard inequality.

Lemma 1.5 (Hadamard inequality)**.** *Let $A \in \mathbb{R}^{n \times n}$ be a matrix with columns $A = (A_1, \ldots, A_n)$. Then*

$$|\det A| \;\leq\; |A_1| \cdots |A_n|.$$

Lemma 1.6 (The Cramer's rule estimate)**.** *Let $A \in \mathbb{Z}^{n \times n}$, $b \in \mathbb{Z}^n$, $\det A \neq 0$ (integer data!). Then the (rational!) solution for the system of equations $Ax = b$ satisfies*

$$|x_i| \leq |A_1| \cdots |A_n| \cdot |b|.$$

Proof. Cramer's rule, together with the observation that the denominator, $\det A$, is an integer, so its absolute value is at least 1. The same is true for the length of each column $|A_i|$. $\square$

1.3.2 Feasibility

First, we should discuss the problem how to find a feasible generalized orthant for $Ax \leq b$, $x \geq 0$, in order to even *start* the simplex algorithm. Here are two solutions to that problem:

- Use the complexity estimates to get explicit upper bounds for the variables, and thus have a starting basis for the dual simplex algorithm (that is, a feasible basis for the simplex algorithm applied to the dual program).
- Phase I: Write down an artificial OPTIMIZATION program, which is feasible, and whose optimal solution (basis) will give a feasible solution (and a feasible basis!) for the FEASIBILITY problem: For example

$$\min x_0 : \ Ax - x_0\mathbf{1} \leq b, \ x \geq 0, x_0 \geq 0.$$

It is trivial that if we can solve **LP-OPTIMIZATION** then we can solve **LP-FEASIBILITY**, in a way that is completely independent of the the specific algorithm used to "implement" **LP-OPTIMIZATION**; that is, we can use any **LP-OPTIMIZATION** *oracle* to "simulate" an **LP-FEASIBILITY** algorithm; in other words, we can program a (fast) algorithm for **LP-FEASIBILITY** if we can use a (fast) subroutine for **LP-OPTIMIZATION** (e.g. by putting objective function zero).

However, note that the converse is also true: If we know how to solve **LP-FEASIBILITY**, then we can also solve **LP-OPTIMIZATION**, that is,

LP-FEASIBILITY $\Longrightarrow$ **LP-OPTIMIZATION**.

For this, note that any *feasible* solution (x, y) for the primal-dual program

$$\begin{array}{cccc} & & c^t x \geq b^t y & \\ (PD) & Ax \leq b & & A^t y \geq c \\ & x \geq 0 & & y \geq 0 \end{array}$$

1.3.3 Modelling issues

Conversion of programs from equality form to inequality form, and conversely. See the Exercises.

1.3.4 Perturbation techniques

If we replace the right-hand sides b_i by $b_i + \varepsilon^i$, for a suitably small ε, then

- the perturbed problem will be feasible if *and only if* the original problem is feasible,
- the perturbed problem will be primally *non-degenerate*, that is, it describes a simple polyhedron, and at any generalized orthant (basis), no extra inequalities are tight (that is, the non-basic variables are non-zero).

(see Exercise).

Moreover,

- similarly, by perturbing the objective function the program can be made *dually non-degenerate*, so that in particular the optimal solution is unique (if it exists), and
- the suitable $\varepsilon > 0$ can be estimated explicitly.

1.3.5 Integral solutions? An example

In general, the optimal solutions will not be integral, although many applications ask for integral solutions. Even if we find the best integral solution, this will come without a certificate, as there may be not dual constraints that are tight at the best integer solution.

However, in many combinatorial situations, we are lucky. Here is one example.

Example 1.7 (Network flows). If the bounds on each arc are integral, then the optimal solution will be integral.

(This may be seen from an algorithm by successive improvement, or from a matrix argument, see exercise.)

Interpretation of dual solutions: Max cut!

Max-Flow-Min-Cut theorem!

Exercise 1.8. Let $A \in \{0, 1, -1\}^{n \times n}$ be a $0/\pm 1$ matrix. Show that

(i) The determinant of a $0/1$-matrix A can be large, even if there are only two 1s per row.
(ii) The determinant of A is not large if there is at most one 1 and at most one -1 per row.
(iii) Use the Hadamard inequality to give an upper bound on $|\det A|$
(iv) For $A \in \{0, 1\}^{n \times n}$ give a much better upper bound, by
 — Multiplying the matrix by 2,
 — Adding a column of 0's and then a row of 1's,
 — subtracting the first row from all others
 and then applying Hadamard to the resulting ± 1-matrix.
(v) Give examples where this bound is tight.

End of class on April 29

[1] Gil Kalai. Linear programming, the simplex algorithm and simple polytopes. *Math. Programming, Ser. B*, 79:217–233, 1997. Proc. Int. Symp. Mathematical Programming (Lausanne 1997).

[2] Gil Kalai and Daniel J. Kleitman. A quasi-polynomial bound for the diameter of graphs of polyhedra. *Bulletin Amer. Math. Soc.*, 26:315–316, 1992.

[3] Victor Klee and David W. Walkup. The d-step conjecture for polyhedra of dimension $d < 6$. *Acta Math.*, 117:53–78, 1967.

[4] László Lovász. *An Algorithmic Theory of Numbers, Graphs and Convexity*, volume 50 of *CMBS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1986.

[5] Francisco Santos. A counterexample to the Hirsch conjecture. *Annals of Math.*, 176:383–412, 2012.

[6] Michael J. Todd. An improved Kalai–Kleitman bound for the diameter of a polyhedron. Preprint, April 2014, 4 pages, `http://arxiv.org/abs/1402.3579`.

2 Convex Bodies, Volumes, and Roundness

> "Although convexity is a simple property to formulate, convex bodies possess a surprisingly rich structure" (Keith Ball [1])

Archimedes book "On the sphere and the cylinder"

2.1 Some basic definitions and examples

Definition 2.1 (Linear/affine/conical/convex hulls)**.** Define in $\mathbb{R}^n$:

- *Linear subspace*, *linear hull*
- *Affine subspace* (possibly empty), *affine hull*
- *Conical subspace* (= *convex cone*, or simply *cone*), *conical hull*
- *Convex hull*, *convex set*

Definition 2.2 (Convex set, line-free, bounded, convex body)**.** Define in $\mathbb{R}^n$: A convex set is

- *line free*: does not contain an affine line
- *bounded*: does not contain a ray
- *convex body*: a closed, bounded (that is, compact) full-dimensional convex set
- *strictly convex*: if $\lambda x + (1-\lambda)y \in \mathrm{int}C$ for $0 < \lambda < 1$ and $x \neq y$.

Examples:

- linear and affine subspaces
- convex polygons in the plane
- regular polyhedra in 3-space

Example 2.3. The unit ball of $\mathbb{R}^d$ with ℓ_2 norm is a centrally-symmetric proper convex body.

Indeed, convexity follows from the triangle inequality

Thus the "theory of finite-dimensional Banach spaces" is equivalent to the "theory of centrally-symmetric convex bodies."

For example, Dvoretzky's theorem, which says that every centrally symmetric convex body in $\mathbb{R}^n$ has a central section of dimension roughly $\log n$ that is linearly approximately equivalent to some $\mathbb{R}^d$ with the Euclidean norm, is a theorem about centrally symmetric convex bodies, which have sections that are roughly ellipsoids. (Indeed, concentration of measure implies that a random subspace will do ...)

End of class on May 6

Example 2.4. The set PSD_n of positive semi-definite $(n \times n)$-matrices is a closed convex cone in $\mathbb{R}^{n\times n}$ of dimension $\binom{n+1}{2}$.

Example 2.5. If identify the N-dimensional vector space $\mathbb{R}[x_1,\ldots,x_d]_{\leq 2k}$ of real polynomials in d variables of degree less than $2k$ with $\mathbb{R}^N$. Then the set

$$\mathcal{P}_{d,2k} = \{p : \ p \in \mathbb{R}[x_1,\ldots,x_d]_{\leq 2k} \text{ and } p(x) \ \geq \ 0 \text{ for all } x \in \mathbb{R}^d\}$$

of positive polynomials is a closed convex cone in $\mathbb{R}^N$. Similarly the set

$$\Sigma_{d,2k} = \{p : \ \exists h_1,\ldots,h_n \in \mathbb{R}[x_1,\ldots,x_d]_{\leq k} \text{ such that } p(x) = h_1^2(x) + \ldots h_n^2(x)\}$$

of sums of squares (SOS) is a closed convex cone in $\mathbb{R}^N$.

Exercise: The dimension of the vector space $\mathbb{R}[x_1, \ldots, x_d]_{\leq 2k}$ is $\binom{2k+d}{d}$.

Theorem 2.6 (Gauß–Lucas Theorem)**.** *Let $p \in \mathbb{C}[z]$ be a complex polynomial in one variable with roots $r_1, \ldots, r_d$, then the roots of the derivative p' of p are contained in the convex hull* $\operatorname{conv}\{r_1, \ldots, r_d\}$.

Proof. If r_1 is a zero of p as well as of p', then $r_1 = 1r_1 + 0r_2 + \cdots + 0r_d$ is a convex combination. Assume z is a zero of p' but not of p. Write p and p' in terms of their roots (they factor over $\mathbb{C}$) and look at $\frac{p(z)}{p'(z)}$ to get a convex combination of the r_i. □

If p has only real roots, then the above result is a consequence of the Rolle's theorem (or the mean value theorem).

2.2 Topological properties

Theorem 2.7 (Carathéodory)**.** *Let $A \subseteq \mathbb{R}^d$ be a set and $x \in \operatorname{conv}(A)$ a point in the convex hull of A. Then there are $d+1$ points $p_1, \ldots, p_{d+1}$ in A such that $x \in \operatorname{conv}\{p_1, \ldots, p_{d+1}\}$.*

Proof. Write $x \in \operatorname{conv}(A)$ as convex combination of a *minimal* number of points $p_1, \ldots, p_n \in A$. If n is more than $d+1$, then there is an affine dependency, where one of the coefficients is positive. Subtract a multiple of this dependency to kill one of the coefficients of the convex combination. □

Corollary 2.8. *If $A \subset \mathbb{R}^d$ is compact, then so is $\operatorname{conv}(A)$.*

Proof. Let $\Delta_d = \operatorname{conv}\{e_1, \ldots, e_{d+1}\}$ be the standard d-simplex in $\mathbb{R}^{d+1}$. Consider a map from $A^{d+1} \times \Delta_d \to A$ given by $(p_1, \ldots, p_{d+1}, \lambda) \mapsto \sum \lambda_i p_i$. Clearly its image is contained in $\operatorname{conv}(A)$. The converse is true by Carathéodory's theorem. Hence $\operatorname{conv}(A)$ is the image of a compact set under a continuous map. □

End of class on May 8

Definition 2.9 (interior points, interior, boundary points, boundary)**.**

Definition 2.10 (relative interior, relative boundary)**.**

Proposition 2.11. *If K is convex, then $\operatorname{relint} K$ is also convex.*

Lemma 2.12. *If $x_0, \ldots, x_k$ are affinely independent, $P := \operatorname{conv}\{x_0, \ldots, x_k\}$, then $x \in \operatorname{relint} P$ if and only if $x = \lambda_0 x_0 + \cdots + \lambda_k x_k$ with all $\lambda_i > 0$.*

Corollary 2.13. *K convex, not empty, then $\operatorname{relint} K \neq \emptyset$.*

Definition 2.14 (dimension of a convex set)**.**

Theorem 2.15 (Carathéodory's Theorem — ambient space free version)**.**

Corollary 2.16. *Characterization of relative interior of a polytope $P := \operatorname{conv}\{x_0, \ldots, x_n\}$: $x = \lambda_0 x_0 + \cdots + \lambda_k x_k$ with all $\lambda_i > 0$.*

Definition 2.17 (extreme points)**.**

Theorem 2.18 (Minkowski)**.** *K closed and bounded convex set, then $K = \operatorname{conv}(\operatorname{ext} K)$.*

End of class on May 13

2.3 Support and separation

Definition 2.19 (Nearest point map)**.** Let $A \subset \mathbb{R}^d$ be a non-empty closed convex set. Then the *nearest-point map* of A is the map $\pi_A : \mathbb{R}^d \to A$ which assigns to each $x \in \mathbb{R}^d$ the point on A with the smallest (Euclidean) distance from x.

Proposition 2.20. *The map π_A of "Definition" 2.19 exists (that is, the nearest point exists, lies in A, and is unique) and the map is contractive:*

$$\|\pi_A(x) - \pi_A(y)\| \leq \|x - y\|$$

and thus in particular continuous.

Exercise 2.21. There is a converse: If for a closed set A the nearest point $\pi_A(x)$ is unique for all x, then A is convex.

Notation: H hyperplane, H^+, H^- half spaces: They are closed convex sets, their interiors are $\mathrm{inter}(H^+) = \mathbb{R}^d \setminus H^-$ and $\mathrm{inter}(H^-) = \mathbb{R}^d \setminus H^+$, their boundary is $\partial H^+ = \partial H^- = H$.

Definition 2.22 (separates)**.** If $A \subseteq \mathbb{R}^d$ is a convex set and $p \in \mathbb{R}^n$, then a hyperplane H *separates* p from A if $p \in H^+$ and $A \subseteq \mathrm{inter}(H^-)$, that is, $A \cap H^+ = \emptyset$. The hyperplane H *strictly separates* A and p if $A \subseteq \mathrm{inter}(H^-)$ and $p \in \mathrm{inter}(H^+)$.
If $A, B \subseteq \mathbb{R}^d$ are convex sets, then a hyperplane H *separates* B from A if $B \subset H^+$ and $A \subseteq \mathrm{inter}(H^-)$. The hyperplane H *strictly separates* A and B if $A \subseteq \mathrm{inter}(H^-)$ and $B \subseteq \mathrm{inter}(H^+)$.

Note that if separation implies that the sets are disjoint, and strict separation implies weak separation. However, separation is not symmetric: There may be a hyperplane that separates B from A, but none that separates A from B.

Theorem 2.23 (Separation Theorem)**.** *Let $A \subseteq \mathbb{R}^d$ be a non-empty closed convex set and $p \notin A$, then there is a hyperplane that strictly separates p and A.*

Proof. Set $q := \pi_A(p)$, $c := p - q$, $H := \{x \in \mathbb{R}^d : c^t x = c^t q\}$ and $H_{1/2} := \{x \in \mathbb{R}^d : c^t x = c^t q\}$.
An elementary geometric argument shows that $A \subset H^-$, while $p \notin H^-$, such that H separates p from A with $q \in H$, and $H_{1/2}$strictly separates p and A. $\square$

Definition 2.24 (supporting hyperplane)**.** A *supporting hyperplane* H for a convex set A satisfies $A \subseteq H^-$ and $A \cap H \neq \emptyset$.

... so this exists by the (proof of the) Separation Theorem.

Corollary 2.25. *closed convex set is intersection of half spaces given by supporting hyperplanes*

Corollary 2.26. *If A is a convex body, then for each direction $c \neq 0$ there is a unique supporting hyperplane $H = \{x : c^t x = \delta\}$.*

Definition 2.27 (support function)**.** Convex body A, define $h_A : \mathbb{R}^d \to \mathbb{R}$ by $h_A(c) := \max\{c^t x : x \in A\}$.

Corollary 2.28. *Convex body is determined by its support function.*

Definition 2.29 (Minkowski sum)**.** The *Minkowski sum* of two sets $A, B \subseteq \mathbb{R}^d$ is

$$A + B := \{x + y : x \in A,\ y \in B\}.$$

Lemma 2.30. *If A and B are convex, then so is $A + B = B + A$.*

Lemma 2.31. *K, L, M convex bodies. Then*
(i) $h_{K+L} = h_K + h_L$.
(ii) $K + M = L + M$ implies $K = L$.

Remark 2.32. We have just established that the set of convex bodies $\mathcal{K}_d$ is a cancellative commutative monoid (without neutral element).

End of class on May 15

Theorem 2.33 (Supporting Hyperplane Theorem)**.** *Let $A \subset \mathbb{R}^d$ be a closed and convex set. Then for every point p in the (relative) boundary ∂A of A there is a supporting hyperplane $H_A(p)$ for A at p, that is, $A \subseteq H_A^-(p)$ and $p \in H \cap A$.*

Proof. If A is not full-dimensional, replace the ambient space by an affine subspace of dimension $\dim(A)$. A supporting hyperlane in this subspace lies inside some (actually many) hyperplanes in $\mathbb{R}^d$, all of which are supporting. So assume A is full-dimensional. Via the nearest point map π_A we get a supporting hyperplane for A at each point $\pi(y)$ for $y \in \mathbb{R}^d \setminus A$ with unit normal vector $\frac{y-\pi_A(y)}{\|y-\pi_A(y)\|_2}$. Take a series $(y_n) \subset \mathbb{R}^d \setminus A$ that converges to p. The corresponding sequence of unit normal vectors $u_n := \frac{y_n-\pi_A(y_n)}{\|y_n-\pi_A(y_n)\|_2}$ for the supporting hyperplanes at $\pi_A(y_n)$ has, by compactness of the unit sphere, a subsequence converging to $u \in S^{d-1}$. There is a corresponding subsequence of (y_n) that also converges to p. Using convergence of the sequences and continuity of the inner product argue that $H_A(p) := \{x \in \mathbb{R}^d : u^t x = u^t p\}$ is a supporting for A at p. □

Proof of Minkowski's Theorem 2.18. The inclusion $K \supseteq \operatorname{conv}(\operatorname{ext} K)$ is trivial. For the other inclusion argue by induction on $d = \dim(C)$. The cases $d = 0, 1$ are trivial. Assume the theorem holds for all compact and convex sets of dimension less than d. Assume K has dimension d. Let $p \in \partial K$. Then, by Theorem 2.33 above, there is a supporting hyperplane $H_K(p)$ for K at p. The "face" $F := K \cap H_K(p)$ is of lower dimension and hence $p \in \operatorname{conv}(\operatorname{ext} F)$. By the homework assignment $\operatorname{ext} F \subseteq \operatorname{ext} K$ and hence $p \in \operatorname{conv}(\operatorname{ext} K)$. If $p \in \operatorname{relint}(K)$ take a line through p that intersects ∂A in two points. Argue using faces that these points are in $\operatorname{conv}(\operatorname{ext} K)$, so p must be in $\operatorname{conv}(K)$ as well. □

2.4 Spectrahedra

Definition 2.34. A *spectrahedron* S is the intersection of the cone PSD_n of symmetric positive-semidefinite matrices with a d-dimensional affine subspace V (of the space of symmetric $n \times n$ matrices). If A is positive semi-definite we write $A \succeq 0$.

Proposition 2.35. *A spectrahedron S is convex and closed. It can be written as*

$$S = \{(x_1, \dots, x_d) \in \mathbb{R}^d \colon A_0 + x_1 A_1 + \dots x_d A_d \succeq 0\},$$

for suitable symmetric matrices $A_0, \dots, A_d$ of size $n \times n$. Let $A(x) := A_0 + x_1 A_1 + \dots x_d A_d$ denote the (symmetric) matrix valued function from $\mathbb{R}^d \to \mathbb{R}^{n \times n}$.

Example 2.36. The cylinder

$$C := \{(x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 \leq 1, -1 \leq z \leq 1\}$$

is a spectrahedron. Consider the points $(x, y, z) \in \mathbb{R}^3$ such that the sum

$$A_0 + xA_1 + yA_2 + zA_3 = \begin{pmatrix} 1+x & y & 0 & 0 \\ y & 1-x & 0 & 0 \\ 0 & 0 & 1+z & 0 \\ 0 & 0 & 0 & 1-z \end{pmatrix} \succeq 0.$$

Here A_0 is the identity matrix. A_1 has a 1 in position $(1, 1)$ and a -1 at $(2, 2)$ and otherwise zeros. A_2 is zero except for 1s at $(1, 2)$ and $(2, 1)$. Finally, A_3 is zero except for a 1 at $(3, 3)$ and a -1 at $(4, 4)$. It turns out that C is the set of all points $w = (x, y, z)$ that satisfy $A(w) \succeq 0$. The cylinder C can also be viewed as the intersection of PSD_4 with the affine subspace $A_0 + \mathrm{span}\{A_1, A_2, A_3\}$.

Proposition 2.37. *Any polyhedron P is a spectrahedron.*

Proof commented, since it is a current exercise. □

Example 2.38. Any univariate sum of squares (SOS) polynomial $p \in \mathbb{R}[t]$ of degree $2n$ that can be written as

$$p = (1, t, t^2, \dots, t^n)^t \begin{pmatrix} 1 & 0 & a \\ 0 & 1-2a & 0 \\ a & 0 & 1 \end{pmatrix} (1, t, t^2, \dots, t^n)$$

defines a spectrahedron S, where S is given by all a such that the matrix is positive semi-definite. Actually $S = [-1, 1/2]$. This extends to polynomials of higher degree that can be written as $\mathbf{t}^t \, A \, \mathbf{t}$ for positive semi-definite A.

Example 2.39 (Non-example). Consider the (linear) projection of the cylinder C into the plane given by $x + 2z = 0$. What we get is the convex hull C' of two non-intersecting ellipses in the plane. Recalling that a matrix is positive semidefinite if the determinants of all of its diagonal minors are non-negative, we can conclude that any spectrahedron must be a so-called *basic semialgebraic set*, that is, a set of points satisfying finitely many polynomial inequalities where the polynomials are of finite degree. Using the fact that infinitely many points determine a polynomial of finite degree one can argue that C' is not basic semialgebraic, hence implying that C' is not a spectrahedron.

End of class on May 20 + May 22

2.5 Löwner–John ellipsoids and roundness

Definition 2.40. An *ellipsoid* $E \subseteq \mathbb{R}^d$ is the image $f(B^d)$ of the unit ball under an invertible affine transformation $f : \mathbb{R}^d \to \mathbb{R}^d$.

If the transformation is $f : x \mapsto Ax + c$, then

$$\begin{aligned} f(B^d) &= \{x \in \mathbb{R}^d : \langle A^{-1}(x-c), A^{-1}(x-c)\rangle \leq 1\} \\ &= \{x \in \mathbb{R}^d : \langle Q(x-c), x-c\rangle \leq 1\} \end{aligned}$$

for $Q = A^* A^{-1} = (AA^t)^{-1}$ positive-definite.

Lemma 2.41. *The volume of E is $|\det A| \mathrm{vol} B^d = \frac{\mathrm{vol} B^d}{\sqrt{\det Q}}$.*

Exercise 2.42. If $E = \{x \in \mathbb{R}^d : \langle Qx, x\rangle \leq 1\}$, show that the polar is $E^* = \{x \in \mathbb{R}^d : \langle Q^{-1}x, x\rangle \leq 1\}$. Deduce that $(\mathrm{vol} E)(\mathrm{vol} E^*) = (\mathrm{vol} B^d)^2$.

Exercise 2.43. If $g : \mathbb{R}^d \to \mathbb{R}^d$ is a surjective linear map, and $E \subset \mathbb{R}^d$ is an ellipsoid, then $g(E)$ is an ellipsoid in $\mathbb{R}^k$.

Lemma 2.44. *Every ellipsoid $E \subset \mathbb{R}^d$ can be written in the form $E = S(B^d) + c$, where S is a positive-definite (symmetric) matrix.*

Proof. Use the (left) polar decomposition: every invertible A can be written as $A = P'U$, where $U = A\sqrt{A^tA}^{-1}$ is a unitary matrix, and $P' = AU^{-1} = \sqrt{AA^t}$ is positive-definite. Then $A(B^d) = S(B^d)$. $\square$

Lemma 2.45. *If X, Y are positive-definite (symmetric square) matrices, then*

$$\det\left(\frac{X+Y}{2}\right) \geq \sqrt{\det(X)\det(Y)},$$

with equality if and only $X = Y$.

Proof. We can write $X = U^t D^2 U$ for unitary U and non-negative diagonal D, and with this $Y = U^t D Y' D U$. With this we obtain that without loss of generality $X = I_d$.
Furthermore, the resulting Y' can be diagonalized, and without loss of generality Y is diagonal. Then things reduce to simple inequalities of the form $\frac{1+\lambda_i}{2} \geq \sqrt{\lambda_i}$ for certain positive eigenvalues λ_i. $\square$

Theorem 2.46 (Löwner–John). *If $K \subset \mathbb{R}^d$ is a convex body, then the maximum-volume ellipsoid $E \subseteq K$ exists and is unique.*

Proof. For the existence, consider the set

$$X := \{(S, c) : S \text{ positive semidefinite},\ c \in \mathbb{R}^d, S(B^d) + c \subseteq K\}.$$

By Lemma 2.44, every ellipsoid in K is represented by a pair (S, c) in X. As K is bounded, we get that X is bounded. It is also closed, so it is compact. Moreover, the volume function on X, given by $\det(S)\mathrm{vol}(B^d)$, is continuous, so the maximum exists.

To show that it is unique, first note that from any two ellipsoids of the same maximum volume $E_1 = S_1(B^d) + c_1$ and $E_2 = S_2(B^d) + c_2$ we can construct a new one $\frac{1}{2}(E_1 + E_2)$ given by $S := \frac{1}{2}(S_1 + S_2)$ and $c := c_1 + c_2$. Lemma 2.45 now yields that if both E_1 and E_2 have maximal volume, then $S_1 = S_2$.
To see $c_1 = c_2$, we may now after a coordinate transformation assume that $S_1 = S_2 = I$ is a unit ball. So we just have to show that the convex hull of the union of two distinct unit balls contains an ellipsoid of larger volume. □

Theorem 2.47. *The minimal volume ellipsoid that contains a given convex body K is also unique.*

Theorem 2.48. *Let $K \subset \mathbb{R}^d$ be a convex body and let $E \subseteq K$ be the maximal volume ellipsoid in K, where we assume that its center is the origin 0. Then*

$$E \subseteq K \subset dE.$$

Proof. Elementary calculation. □

Theorem 2.49. *Let $K = -K \subset \mathbb{R}^d$ be a centrally-symmetric convex body and let $E \subseteq K$ be the maximal volume ellipsoid in K. Then*

$$E \subseteq K \subset \sqrt{d}E.$$

Proof. Elementary calculation. □

End of class on May 27

2.6 Volume computation and ellipsoids

Theorem 2.50 (Ernst Sas (1939)). *Let C be a convex disk (a convex body in the plane) and let $n \geq 3$ be an integer. If $P_{(n)}$ is an n-gon of maximal area contained in C, and P_n^2 is a regular n-gon inscribed into the unit disk B^2, then*

$$\frac{\mathrm{vol}(P_{(n)})}{\mathrm{vol}(C)} \geq \frac{\mathrm{vol}(P_n^2)}{\mathrm{vol}(B^2)} = \frac{n}{2\pi} \sin \frac{2\pi}{n},$$

with equality if and only if C is an ellipse.
(Extension by Alexander Macbeath (1951)) *Let C be a convex body in $\mathbb{R}^d$ and let $n \geq d + 1$ be an integer. If $P_{(n)}$ is a polytope with n vertices of maximal volume contained in C, and P_n^d is a convex polytope of maximal volume inscribed into the unit ball B^d, then*

$$\frac{\mathrm{vol}(P_{(n)})}{\mathrm{vol}(C)} \geq \frac{\mathrm{vol}(P_n^d)}{\mathrm{vol}(B^d)},$$

with equality if and only if C is an ellipsoid.

(Discussed without proof; see problem set for references.)

Theorem 2.51 (György Elekes (1986))**.** *Let $P^d_{(n)}$ be a convex d-polytope with n vertices contained in B^d. Then*

$$\frac{\mathrm{vol}(P^d_{(n)})}{\mathrm{vol}(B^d)} \leq \frac{n}{2^d}.$$

Proof. If $P = \mathrm{conv}\{v_1, \ldots, v_n\}$, show that the balls with diameter $[0, v_i]$ cover P. Each of these has volume at most $\frac{1}{2}\mathrm{vol}(B^d)$. □

Definition 2.52 (oracles)**.** MEMBERSHIP, SEPARATION
VALIDITY, VIOLATION

Definition 2.53 (guarantees)**.** A convex body is *well-guaranteed* if we know that

- $C \subseteq B(0, R)$

and if one of the (equivalent!) properties holds:

- C contains a ball of radius r_0, for a specified $r_0 > 0$.
- C has width w_0, for a specified $w_0 > 0$.
- C has volume at least v_0, for a specified $v_0 > 0$.

Corollary 2.54. *A well-guaranteed MEMBERSHIP oracle needs exponential time for any reasonable volume estimate.*

2.7 The Ellipsoid method

Lemma 2.55. *Let $B^d_+ = \{x \in \mathbb{R}^d : |x|^2 \leq 1,\ x_d \geq 0$ be the "positive half d-ball." Then the ellipsoid*

$$E \ := \ \Big\{ x \in \mathbb{R}^d : (1 - \tfrac{1}{d^2})(x_1^2 + \cdots + x_{d-1}^2) + (1 + \tfrac{1}{d})^2 (x_d - \tfrac{1}{d+1}) \leq^1 \Big\}$$

satisfies

1. $B^d_+ \subseteq E$,
2. $\dfrac{\mathrm{vol}E}{\mathrm{vol}B^d_+} \leq e^{-1/(2(d+1))}$.

Proof. Simple calculations. For (2) use $1 + x \leq e^x$. □

End of class on June 3

Theorem 2.56 (Ellipsoid Method: Khatchian, Grötschel–Lovász–Schrijver)**.** *If a convex body is given by a well-guaranteed SEPARATION oracle, e.g. by the guarantee that if C is not empty then it satisfies*

$$B(x_0, r) \subseteq C \subseteq B(x, R)$$

for known $r \leq R$ but unknown x_0, then there is an algorithm that decides that C is empty or finds a point $x \in C$ after at most

$$2d(d+1) \ln \frac{R}{r}$$

calls to the oracle.

Proof. Start with $E_0 := B(0,r)$, and construct a sequence of ellipsoids $E_0, E_1, \dots$ by querying the center c_i of E_i. If $c_i \in C$ we are done, otherwise Lemma 2.55 yields E_{i+1} such that we have

(1) If C is not empty, then $B(x_0, r) \subseteq E_{i+1}$,

(2) $\mathrm{vol}\frac{\mathrm{vol}(E_{i+1})}{\mathrm{vol}(E_i)} \le e^{-y\frac{1}{2(d+1)}}$.

Thus the sequence breaks off at some ellipsoid E_n with

$$n \le 2(d+1)\frac{\mathrm{vol}(B(0,R))}{\mathrm{vol}(B(x_0,r))} = 2(d+1)\frac{R^d}{r^d}.$$

□

2.8 Polarity, and the Mahler conjecture

Definition 2.57 (Polar dual)**.** For $\emptyset \ne K \subset \mathbb{R}^d$:

$$K^* := \{c \in \mathbb{R}^d : c^t x \le 1 \text{ for all} x \in K\}$$

is the *polar* of the set K.

Lemma 2.58. *Let* $\emptyset \ne K \subset \mathbb{R}^d$.

(1) $0 \in K^*$*; the set* K^* *is closed and convex.*

(2) $(\mathbb{R}^d) = \{0\}$, $\{0\}^* = \mathbb{R}^d$*; for any linear subspace* $L \subset \mathbb{R}^s$, $L^* = L^\perp$.

(3) $K \subseteq L$ *implies* $L^* \subseteq K^*$.

(4) $\left(\bigcup_{i \in I} K_i\right)^* = \bigcap_{i \in I} K_i{}^*$.

(5) $(\alpha K)^* = \frac{1}{\alpha}K^*$ *for* $\alpha > 0$.

(6) $(AK)^* = (A^t)^{-1}K^*$ *for any invertible* $(d \times d)$*-matrix* A.

(7) $K = \mathrm{conv}\{v_1, \dots, v_n\}$ *implies* $K^* = \{y \in \mathbb{R}^d : y^t v_i \le 1$ *for* $1 \le i \le n\}$.

(8) $K \subseteq K^{**}$.

Theorem 2.59 (Bipolar theorem)**.** *If* $K \subseteq \mathbb{R}^d$ *is closed, convex, and contains* 0*, then* $K = K^{**}$.

Note: If K is a $\mathcal{V}$-polytope, then K^* is an $\mathcal{H}$-polyhedron, etc.

End of class on June 5

Definition 2.60. The *Hanner polytopes* are the polytopes that can be generated from the interval $I := [-1, +1]$ by any two of the three operations product, direct sum $\oplus$, and product $\times$.

(Any two of the operations allow us to also "simulate" the third one here, as $I^* = I$ and $P \oplus Q = (P^* \times Q^*)^*$ and $P \times Q = (P^* \oplus Q^*)^*$.)

Proposition 2.61 (On Hanner polytopes)**.** (0) *The number of combinatorial types for* $d \geq 1$ *grows exponentially:* $n(d) = 1, 1, 2, 4, 8, 18, 40, 94, \ldots$

(1) *All Hanner polytopes have* 3^d *non-empty faces,* $f_0 + f_1 + \cdots + f_d = 3^d$.

(2) *All Hanner polytopes satisfy* $\mathrm{vol}(P)\mathrm{vol}(P^*) = \frac{4^k}{d!}$

Conjecture 2.62 (The Mahler conjecture: Kurt Mahler, 1939; the 3^d conjecture: Kalai 1988)**.** *Let* K *be a convex body in* $\mathbb{R}^d$ *with* $K = -K$*, then*

$$\mathrm{vol}(K)\mathrm{vol}(K^*) \;\geq\; \frac{4^k}{d!}$$

with equality exactly for the Hanner polytopes.
Let P *be a* d*-polytope in* $\mathbb{R}^d$ *with* $P = -P$*, then*

$$f_0 + f_1 + \cdots + f_d \;\geq\; 3^d$$

with equality exactly for the Hanner polytopes.

Note that for the (long-standing) Mahler conjecture, it is not even clear that the extremal objects are polytopes, as we are searching in the class of convex bodies. It is also not trivial that objects (convex bodies) achieving the minimum even exist! (This needs a compactness result such as the Blaschke selection principle, to be discussed later.)
For a recent overview/discussion of the Mahler conjecture [4], see Tao's blog [6].
For proofs/details on Hansen polytopes and the 3^d conjecture, see Hansen [2], Kalai [3], as well as Sanyal et al. [5].

End of class on June 10

[1] Keith Ball. An elementary introduction to modern convex geometry. In S. Levy, editor, *Flavors of Geometry*, volume 31 of *Publ. MSRI*, pages 1–58. Cambridge University Press, 1997.

[2] Olof Hanner. Intersections of translates of convex bodies. *Math. Scand.*, 4:67–89, 1956.

[3] Gil Kalai. The number of faces of centrally-symmetric polytopes. *Graphs and Combinatorics*, 5:389–391, 1989. (Research Problem).

[4] Kurt Mahler. Ein Übertragungsprinzip für konvexe Körper. *Casopis Pest. Mat. Fys.*, 68:93–102, 1939.

[5] Raman Sanyal, Axel Werner, and Günter M. Ziegler. On Kalai's conjectures about centrally symmetric polytopes. *Discrete Comput. Geometry*, 41:183–198, 2009.

[6] Terry Tao. Open question: the Mahler conjecture on convex bodies. Blog page started March 8, 2007, http://terrytao.wordpress.com/2007/03/08/open-problem-the-mahler-conjecture-on-convex-bodies/.

3 Geometric inequalities, mixed volumes, and isoperimetric problems

3.1 Introduction: Arithmetic inequalities

Lemma 3.1. *Among all rectangles with area A a square has the minimal inequality.*

Proof. This translates into

$$2(a+b) \geq 4\sqrt{ab}$$

for $a, b \geq 0$, with equality if and only if $a = b$, where the left-hand side is the perimeter, the right-hand side is 4 times the area. The inequality is equivalent to $\sqrt{ab} \leq \frac{a+b}{2}$, that is, geometric mean is smaller or equals to arithmetic mean, which is proved by squaring, where $ab \geq (\frac{a+b}{2})^2$ is equivalent to $(a-b)^2 \geq 0$. □

Theorem 3.2 (Arithmetic-Geometric Mean inequality)**.** *For $z_1, \ldots, z_n \geq 0$,*

$$\frac{z_1 + \cdots + z_n}{n} \geq \sqrt[n]{z_1 \cdots z_n}$$

with equality only if all z_i are equal.

Proof. We discussed two proofs. The first one noted that all this is trivial if one of the z_i is zero, and otherwise with the substitution $z_i = e^{y_i}$ this can be derived from convexity of the function $y \mapsto e^y$.
Second proof: by a non-standard induction, taken from [1]. □

Lemma 3.3 ("Minkowski's inequality")**.** *For $x_1, \ldots, x_n, y_1, \ldots, y_n \geq 0$, we have*

$$\sqrt[n]{(x_1+y_1)\cdots(x_n+y_n)} \quad \geq \quad \sqrt[n]{x_1 \cdots x_n} + \sqrt[n]{y_1 \cdots y_n},$$

with equality if and only if

- $x_i = \lambda y_i$ *for all i and a fixed λ_i,*
- $x_1 = \cdots = x_n = 1$*,*
- $y_1 = \cdots = y_n = 1$*, or*
- $x_i = y_i = 0$ *for some value of i.*

Proof. In the case that $x_i = y_i = 0$ for some value of i the inequlity is clearly true with equality. Otherwise we have $x_i + y_i > 0$ for all i and thus can set

$$x_i' := \frac{x_i}{x_i + y_i}, \qquad y_i' := \frac{y_i}{x_i + y_i}$$

for all i, so we have to prove that

$$\sqrt[n]{x_1' \cdots x_n'} + \sqrt[n]{y_1 \cdots y_n} \leq 1,$$

which is obtained from a simple calculation using the AGM inequality as well as $x_i' + y_i' = 1$. The remaining equality cases are also obtained from the AGM inequality. □

3.2 Brunn's Slice Inequality and the Brunn–Minkowski Theorem

For the following, let $K \subset \mathbb{R}^{d+1}$ be a $(d+1)$-dimensional convex body. For $c \neq 0$ we slice it by the parallel hyperplanes $H_t := \{x \in \mathbb{R}^{t+1} : c^t x = t$, and consider the volume of the slices, measured by the function

$$f_K(t) := \mathrm{vol}(K \cap H_t).$$

Definition 3.4 (unimodal/concave function)**.** A function $f : \mathbb{R} \to \mathbb{R}$ is *unimodal* if $a < b < c$ implies that $f(b) \geq \min\{f(a), f(c)\}$.
f is *concave* if $f(b) \geq \frac{b-c}{a-c} f(a) + \frac{b-a}{c-a} f(c)$ for $a < b < c$, or equivalently if $f((1-\lambda)a + \lambda f(c)) \geq (1-\lambda)f(a) + \lambda f(c)$.

Important observation: $f_K(t)$ is in general *not* concave!

Theorem 3.5 (Brunn's slice inequality)**.** *Let $K \subset \mathbb{R}^{d+1}$ be a $(d+1)$-dimensional convex body, and let $f_K(t) := \mathrm{vol}(K \cap H_t)$ be the slice function for the parallel hyperplanes $H_t := \{x \in \mathbb{R}^{t+1} : c^t x = t$, then*

$$\sqrt[d]{f_K} : \quad t \longmapsto \sqrt[d]{\mathrm{vol}(K \cap H_t)}$$

is concave on the interval $[t_{\min}, t_{\max}] = \{t \in \mathbb{R} : K \cap H_t \neq 0\}$. Thus, in particular, f_K is unimodal on all of $\mathbb{R}$.

End of class on June 12

Theorem 3.6 ("Brunn–Minkowski inequality")**. (1)** *Let $K, L \subset \mathbb{R}^d$ be convex bodies, then*

$$\sqrt[d]{\mathrm{vol}(K+L)} \quad \geq \quad \sqrt[d]{\mathrm{vol}(K)} + \sqrt[d]{\mathrm{vol}(K)}$$

with equality if and only if K and L are positively homothetic, that is, $K = \mu L + x_0$ for a positive factor $\mu > 0$ and a translation vector $x_0 \in \mathbb{R}^d$.
(2) *Let $K, L \subset \mathbb{R}^d$ be nonempty compact (closed bounded) convex sets, then the same inequality holds, with equality if and only if*

- *K and L are positively homothetic,*
- *K and L lie in parallel hyperplanes, or*
- *one of K and L is a point.*

(3) *Let $K, L \subset \mathbb{R}^d$ be compact and nonempty, then the inequality above still holds.*

Remark 3.7. An equivalent version writes the Brunn–Minkowski Inequality (BMI) as

$$\sqrt[d]{\mathrm{vol}((1-\lambda)K_0 + \lambda K_1)} \quad \geq \quad (1-\lambda)\sqrt[d]{\mathrm{vol}(K_0)} + \lambda \sqrt[d]{\mathrm{vol}(K_1)}$$

for $0 \leq \lambda \leq 1$.

Proof that the Brunn–Minkowski inequality 3.6 implies the Brunn Slice Theorem 3.5. Without loss of generality we may assume that $c^t x = x_{d+1}$.
Furthermore without loss of generality we use $a = 0$ and $b = 1$.
Define $K_0 \times \{0\} := K \cap H_0$ and $K_1 \times \{1\} := K \cap H_1$. Then K contains the so-called *Cayley embedding* of K_0 and K_1 into parallel hyperplanes,

$$C(K_0, K_1) = \mathrm{conv}\{(K_0 \times \{0\}) \cap (K_1 \times \{1\})\},$$

with

$$K \cap H_\lambda \supseteq C(K_0, K_1) \cap H_\lambda = ((1-\lambda)K_0 + \lambda K_1) \times \{\lambda\})\}.$$

For this, the BMI yields

$$\sqrt[d]{\mathrm{vol}(K_\lambda)} \geq \sqrt[d]{\mathrm{vol}((1-\lambda)K_0 + \lambda K_1)} \geq (1-\lambda)\sqrt[d]{\mathrm{vol}(K_0)} + \lambda \sqrt[d]{\mathrm{vol}(K_1)}$$

and we are done. □

Now we tackle the Brunn–Minkowski Inequality (Theorem 3.6), where we use a combinatorial approach, which yields the most general part (3), however without the characterization of the cases of equality.

Lemma. *BMI holds for nonempty convex sets if it holds for polyboxes.*

Here we use knowledge from Measure Theory: We can approximate any compact set in $\mathbb{R}^d$ arbitrarily well with finite unions of axis-parallel rectangular boxes, in such a way that in the limit the measure of the boxes yields the measure of the convex set.
A *polybox consisting of n boxes* is a union n axis-parallel rectangular boxes in $\mathbb{R}^d$ with disjoint interiors. (The condition of "disjoint interiors" is irrelevant for the types of subsets we obtain, but it is relevant for the number n of boxes needed to get a set.)

Proof of the Brunn–Minkowski inequality 3.6, part (3), for polyboxes. Let $S, T \subset \mathbb{R}^d$, which together have $n \geq 2$ polyboxes. We will use induction on the number n of boxes.
The case of $n = 2$ is precisely given by the Minkowski inequality, Lemma 3.3.
For $n > 2$ we may assume that K contains of at least 2 boxes.
We can now find a coordinate hyperplane, w.l.o.g. $H = \{x \in \mathbb{R}^d : x_d = 0\}$, which separates two boxes of S, such that there are less than n boxes of S and T that have volume above H and also less than n boxes below.
Let $p := \frac{\mathrm{vol}(S^+)}{\mathrm{vol}(S)}$ be the fraction of volume of K above H, so $0 < p < 1$.
Translate T so that it has the *same* volume fraction $p = \frac{\mathrm{vol}(T^+)}{\mathrm{vol}(T)}$.
Now we compute, using in the first step that $S^+ + T^+ \subseteq \{x \in \mathbb{R}^d : x_d \geq 0$ and $S^- + T^- \subseteq \{x \in \mathbb{R}^d : x_d \leq 0$ lie in opposite halfspaces, so their interiors don't overlap, and in the second step the BMI,

$$\begin{aligned}
\mathrm{vol}(S+T) &\geq \mathrm{vol}(S^+ + T^+) + \mathrm{vol}(S^- + T^-) \\
&\geq (\sqrt[d]{\mathrm{vol}(S^+)} + \sqrt[d]{\mathrm{vol}(T^+)})^d + (\sqrt[d]{\mathrm{vol}(S^-)} + \sqrt[d]{\mathrm{vol}(T^-)})^d \\
&= (\sqrt[d]{p}\sqrt[d]{\mathrm{vol}(S)} + \sqrt[d]{p}\sqrt[d]{\mathrm{vol}(T)})^d + (\sqrt[d]{1-p}\sqrt[d]{\mathrm{vol}(S)} + \sqrt[d]{1-p}\sqrt[d]{\mathrm{vol}(T)})^d \\
&= p(\sqrt[d]{\mathrm{vol}(S)} + \sqrt[d]{\mathrm{vol}(T)})^d + (1-p)(\sqrt[d]{\mathrm{vol}(S)} + \sqrt[d]{\mathrm{vol}(T)})^d \\
&= (\sqrt[d]{\mathrm{vol}(S)} + \sqrt[d]{\mathrm{vol}(T)})^d.
\end{aligned}$$

□

3.3 Minkowski's existence and uniqueness theorem

Theorem 3.8 (Minkowski's existence and uniqueness theorem)**.** *Let $d \geq 1$, $a_1, \ldots, a_n \in \mathbb{R}^d$ distinct unit vectors, spanning, and $\alpha_1, \ldots, \alpha_n > 0$.*
Then a d-polytope $P \subset \mathbb{R}^d$ with unit facet normals a_i and facet volumes α_i exists if and only if

$$\alpha_1 a_1 + \cdots + \alpha_n a_n = 0.$$

This is trivial for $d = 1$ and elementary (Exercise) for $d = 2$.

Proof. For the "only if" part we consider an arbitrary projection along a vector c, and find $\mathrm{vol}(\overline{F_i}) = \langle c, a_i \rangle \mathrm{vol}(F_i)$, and in the projection the volumes (with signs!) add to zero, so

$$\langle c, \alpha_1 a_1 + \cdots + \alpha_n a_n \rangle = 0.$$

As this holds for every c, we are done.
For the "if" part we have to construct a suitable polytope for given data a_i and α_i.

Proposition 3.9. *If the rows of A are spanning and positively dependent, then*

$$\mathcal{B}_A = \mathrm{im}(A) + \mathbb{R}^n_{\geq 0}.$$

where $\mathrm{im}(A)$, the image of $x \mapsto Ax$, is the column span of the matrix A. In particular, $\mathcal{B}_A$ is a convex polyhedral cone, and its lineality space is $\mathrm{im}A$, that is, the complete lines in $\mathcal{B}_A$ correspond to translations in $\mathbb{R}^d$.

(The proof of the proposition is routine.)

End of class on June 17

We now define the matrix $A \in \mathbb{R}^{n \times d}$ with rows $a_1^t, \ldots, a_n^t$, and the vector of right-hand sides $b \in \mathbb{R}^n$. Consider the polyhedron $P_A(b)$ as a function of the right-hand sides, which yields

$$P_A(b) := \{x \in \mathbb{R}^d : Ax \leq b\},$$

a non-empty polyhedron on the cone $\mathcal{B}_A = \{b \in \mathbb{R}^n : P_A(b) \neq \emptyset\}$, and

$$\mathcal{M}_A := \{b \in \mathbb{R}^n : \mathrm{vol}(P_A(b)) \geq 1\}.$$

Proposition 3.10. *If the rows of A are spanning and positively dependent, then $\mathcal{M}_A$ is a convex set with lineality space is $\mathrm{im}A$. That is, $\overline{\mathcal{M}_A} := \mathcal{M}_A/\mathrm{im}(A)$ is a strictly convex closed convex set.*

Proof. Let $b', b'' \in \mathcal{M}_A$ be right-hand sides that yield polyhedra $P_A(b'), P_A(b'')$ of volume at least 1, and $b := (1-\lambda)P_A(b') + \lambda P_A(b'')$.
Then we find

$$(1-\lambda)P_A(b') + \lambda P_A(b'') \subseteq P_A(b).$$

(Check this!)
Applying the BMI now yields

$$\begin{aligned} \mathrm{vol}(P_A(b)) &\geq \mathrm{vol}((1-\lambda)P_A(b') + \lambda P_A(b'')) \\ &\geq (1-\lambda)\mathrm{vol}(P_A(b')) + \lambda \mathrm{vol}(P_A(b'')) \geq 1. \end{aligned}$$

Equality here means that we need that both $P_A(b')$ and $P_A(b'')$ have volume 1 (for the third inequality), where $P_A(b')$ and $P_A(b'')$ need to be positive homothets to get equality in the second inequality (the BMI), so as they have the same volume they need to be translates, which implies that $b' - b'' \in \mathrm{im}(A)$. $\square$

Proposition 3.11. *On the interior of $\mathcal{B}_A$, the function $b \mapsto \mathrm{vol}(P_A(b))$ is differentiable (it is piecewise-polynomial), with*

$$\frac{\partial}{\partial b_i}\mathrm{vol}(P_A(b)) \;=\; \mathrm{vol}_{d-1}(P_A(b)^{a_i}) = \mathrm{vol}_{d-1}(F_i(P_A(b))).$$

Proof. Elementary geometry: If we vary b_i a bit, $P_A(b)$ changes by moving the facet hyperplane of F_i, and the volume of the difference to first order is the volume of the facet F_i times the height of variation. $\square$

End of class on June 19

Corollary 3.12. *In every boundary point $b^0 \in \partial\mathcal{M}_A$, there is a unique supporting hyperplane, which is given by*

$$H \;=\; \{y \in \mathbb{R}^n : \frac{1}{d}\sum_i \mathrm{vol}(F_i(P_A(b^0)))y_i = 1\}.$$

Proof. This relies on the volume formula for $P_A(b)$, which is

$$\mathrm{vol}(P_A(b)) = \tfrac{1}{d}\mathrm{vol}(F_i(P_A(b)))b_i$$

which is elementary. (For this consider first x as an interior point of $P_A(b)$, then $P_A(b)$ decomposes into pyramids with base $F_i(P_A(b))$ and height $h_i = b_i - a_i^t x$. The volume thus is $\mathrm{vol}(P_A(b)) = \frac{1}{d}\mathrm{vol}(F_i)(b_i - a^t x)$, which gives the correct result, which of course has to be independent of x, by the "only if" part of the Existence and Uniqueness theorem, for which we had established that $\sum_i \mathrm{vol}(F_i)a^i = 0$.) $\square$

Now to proceed with the proof of Minkowski's Existence and Uniqueness Theorem, we consider the optimization problem

$$\begin{aligned} \text{minimize } \phi(b) \;&:=\; \sum_{i=1}^{n} \alpha_i b_i \\ \text{subject to} \quad & b \in \mathcal{M}_A \end{aligned}$$

Here we minimize a linear function over a closed convex set. One can check that the minimum exists, is positive, and is assumed at a point B^* that is unique up to translation of the polytope $P_A(b)$ (as $\mathcal{M}_A/\mathrm{im}(A)$ is strictly convex).

At the point b^*, we know that the gradient of the volume function $\mathrm{vol}(P_A(b))$ coincides with linear function we are trying to minimize. From this we get that b^* lies on the hyperplane

$$\frac{1}{d}\sum_i \mathrm{vol}(F_i(b^*))y_i \;=\; 1$$

as it lies on the support hyperplane of $\mathcal{M}_A$ at the point b^*, and it lies on the hyperplane

$$\frac{1}{d}\sum_i \alpha_i y_i \ = \ \phi_{\min}$$

by construction, where the normal vectors to the hyperplanes must be multiples of each other. From this we see that

$$\mathrm{vol}(F_i(P_A(b^*))) = \frac{\alpha_i}{\phi_{\min}}$$

holds for all i. This yields that

$$P_A(\tfrac{1}{\sqrt[d]{\phi_{\min}}})$$

is the polytope we were looking for, to complete the proof of Minkowski's Existence and Uniqueness Theorem. $\square$

Note: this is constructive "in principle."

Applications, for example: If all polytopes in a dissection $P = P_1 \cup \cdots \cup P_m$ are centrally symmetric, then so is P.

End of class on June 24

3.4 Application: Sorting partially ordered sets

Definition 3.13. $(X, \preceq)$ a finite partially ordered set, then $e(X, \preceq)$ is the number of *linear extensions* of $(X, \preceq)$.

Clearly $1 \leq e(X, \preceq) \leq n!$, with equality for a *linear order* (also known as *chain* or *total order*) resp. for an anorderd set (*antichain*).

Theorem 3.14 (Efficient comparison theorem)**.** *Let $(X, \preceq)$ be a finite partial order that is not linear. Then there are elements $a, b \in X$ such that*

$$\delta \leq \frac{e(X, \preceq +(a,b))}{e(X, \preceq)} \leq 1 - \delta,$$

where δ is a constant.

Here $e(X, \preceq +(a,b))$ denotes the partial ordering we obtain from $e(X, \preceq)$ if we are given the additional information that $a \leq b$.

We will sketch a proof by Kahl & Linial (1991) which yields this for .... $\delta = \frac{1}{2e} \approx 0.1840$.
The original proof by Kahn & Sachs (1984) yielded .... $\delta = \frac{3}{11} \approx 0.2727$.
The current best bound by Brightwell, Felsner & Trotter (1995) is .... $\delta = \frac{5-\sqrt{5}}{10} \approx 0.2764$.
The *conjecture* is that this should be true for .... $\delta = \frac{1}{3} \approx 0.3333$,
which would be optimal. See: Matoušek [3, Sect. 12.3].

Remark 3.15 (The complexity of sorting a partially ordered set). If we are supposed to sort a partially ordered set by pairwise comparisons, that is, find the unknown linear order from partial data by queries to a comparison oracle, the worst case complexity is certainly at least $\log_2 e(X, \preceq)$. For example, this yields the well-known lower bound of $\log_2 n \approx n \log_2 n$ for sorting without previous information.

The "efficient comparison theorem" yields an upper bound: If we choose our comparisons judiciously, $\log_{1/(1-\delta)} e(X, \preceq)$ steps will be enough.

Definition 3.16. For a given partial order $(X, \preceq)$ on a set X of size n, which for simplicity we identify with $\{1, \dots, n\}$, the *order polytope* is

$$P(X, \preceq) := \{x \in [0,1]^n : x_a \le x_b \text{ for all } a, b \in X \text{ with } a \preceq b\}.$$

Lemma 3.17. *The number of vertices of $P(X, \preceq)$ is the number of order ideals (a.k.a. down-sets) of $(X, \preceq)$. Indeed, the vertices are the characteristic vectors of the dual order ideals (a.k.a. up-sets) of $(X, \preceq)$.*
The volume of $P(X, \preceq)$ is $\frac{1}{n!}e(X, \preceq)$.

Proof. $P(X, \preceq)$ has a canonical triangulation into simplices of determinant 1 (that is, volume $\frac{1}{n!}$) corresponding to the linear extensions. □

Definition 3.18 (height)**.** Let X be a finite set and $a \in X$.
For a linear ordering $(X, \le)$, the *height* of a in $(X, \le)$ is defined as the number of elements below a, $h_{\le}(a) := |\{x \in X : x \le a\}|$.
For a partial ordering $(X, \preceq)$, the *height* of a in $(X, \preceq)$ is definied as the average number of elements below a in the linear extensions of $(X, \preceq)$, that is,

$$h_{\preceq}(a) := \frac{1}{e(X, \preceq)} \sum_{\le \in E(X, \preceq)} h_{\le}(a).$$

Lemma 3.19. *For any poset $(X, \preceq)$ on an n-element set X, the center of gravity of its order polytope $P(X, \preceq)$ has the coordinates $c_a = \frac{1}{n+1}h_{\preceq}(a)$.*

Proof. The center is the average of the centers of the simplices in the triangulation. □

Proof of the Efficient Comparison Theorem 3.14.
(1) Pick elements $a \neq b$ in X with $|h_{\preceq}(a) - h_{\preceq}(b)| < 1$. In particular, a and b are not comparable in $\preceq$. If $(X, \preceq)$ is not a linear order, these exist (Exercise!)
We want to show that (a, b) solves the problem. For this we have to show that the hyperplane $x_a = x_b$ splits the polytope $P(X, \le)$ into two parts that each have a constant fraction of the volume of the whole polytope.
(2) Choose a new orthonormal coordinate system $y_1, \dots, y_n$, where $y_1 = x_a - x_b$.
(*Note:* The coordinate transformation can be obtained by an orthogonal transformation followed by a rescaling by factor $\frac{1}{2}\sqrt{2}$.)
In these new coordinates, the splitting hyperplane is given by $y_1 = 0$. The polytope $P = P(X, \preceq)$ in these new coordinates has two properties:

- The projection of P to the first coordinate is $[-1, 1]$.
  (There are vertices corresponding to up-sets that contain a but not b, and the other way around.)
- The center of gravity satisfies $-\frac{1}{n+1} < c_1 < \frac{1}{n+1}$.
  (Indeed, $c_1 = \frac{1}{n+1}(h_{\preceq}(b) - h_{\preceq}(a))$ with $|h_{\preceq}(a) - h_{\preceq}(b)| < 1$)

We want to show that these two properties already imply what we need, namely *every* convex body with these two properties satisfies $\mathrm{vol}(P_{y_1 \ge 0}) \ge \frac{1}{2e}\mathrm{vol}(P)$ and $\mathrm{vol}(P_{y_1 \le 0}) \ge \frac{1}{2e}\mathrm{vol}(P)$.

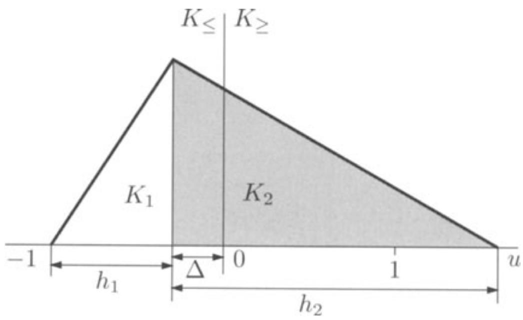
(3) The y_1-coordinate of the center of gravity is given by the volume of the slices by

$$c_1(P) = \frac{1}{\mathrm{vol}(P)} \int_{-1}^{1} t\,\mathrm{vol}_{n-1}(P_t)\mathrm{d}t.$$

Thus we can replace P by a rotationally symmetric convex body R with the same properties and with the same center of gravity, given by the radius function $r(t)$. By Brunn's Slice Inequality (Theorem 3.5), the function $r(t)$ is convex, and thus the resulting body R is convex.

(4) Replace R by a double cone K, determined by radius function $\kappa(t)$, with the following properties

- $\mathrm{vol}(K_{\geq 0}) = \mathrm{vol}(R_{\geq 0})$, while the center of gravity of $R_{\geq 0}$ moves to the right, if at all ("move mass".)
- $\mathrm{vol}(K_{\leq 0}) = \mathrm{vol}(R_{\leq 0})$, while the center of gravity of $R_{\geq 0}$ moves to the right, if at all ("move mass".)



(5) Computations for the double cone K: It is determined by the y_1-coordinate of the "base," called $-\Delta$, and by the heights $h_1 = 1 - \Delta$ and $h_2 = u + \Delta \geq 1 + \Delta$.
The barycenter is computed to satisfy $c_1(K) = \frac{h_2 - h_1}{n+1} - \Delta$, which yields $\frac{u}{h_2} \geq 1 - \frac{1}{n}$.
And from this we get the volume estimate

$$\begin{aligned}
\mathrm{vol}(K_{\geq 0}) &= \frac{u}{u+1}\left(\frac{u}{h_2}\right)^n \mathrm{vol}(K_2) \\
&= \frac{u}{u+1}\left(\frac{u}{h_2}\right)^n \frac{h_2}{h_1+h_2}\mathrm{vol}(K) \\
&\geq \frac{u}{u+1}\left(1 - \frac{1}{n}\right)^{n-1} \mathrm{vol}(K) \;\;\geq\;\; \frac{2e}{\mathrm{vol}}(K).
\end{aligned}$$

See Matoušek [3, Sect. 12.3] for details. □

End of class on June 26

3.5 Mixed subdivisions and Mixed volumes

Definition 3.20. Let again $A \in \mathbb{R}^{n \times d}$ have disjoint rows of length 1 and let $b \in \mathbb{R}^n$.
The *closed inner region* $\mathcal{B}_A^\circ \subseteq \mathcal{B}_A$ is the set of all right-hand sides b such that all inequalities define a non-empty face:

$$\mathcal{B}_A^\circ = \{b \in \mathbb{R}^n : P_A(b) \cap \{x \in \mathbb{R}^d : a_i^t x = b_i\} \neq \emptyset \text{ for all } i\}.$$

Note that $\mathcal{B}_A^\circ$ is a closed polyhedral cone again! In the following for a polyhedron $P \subseteq \mathbb{R}^d$ and a vector $c \in \mathbb{R}^d$, the expression P^c denotes the face of P in direction of c.

Definition 3.21. Define $P \leq_w Q$ if $\dim P^c \leq \dim Q^c$ for all $c \in \mathbb{R}^d$.
Define $P \sim_w Q$ if $\dim P^c = \dim Q^c$ for all $c \in \mathbb{R}^d$.
In the latter case, P and Q are called *normally equivalent*.
The *type cone* of b is

$$\mathcal{T}_A(b) := \{b' \in \mathbb{R}^d : P_A(b') \sim_w P_A(b)\}$$

Its closure is called the *closed type cone* of b.

Proposition 3.22. *The type cone $\mathcal{T}_A(b)$ is a relatively open polyhedral cone.*
Thus the closure of the type cone $\mathcal{T}_A(b)$ is a polyhedral cone.
The type cones define a polyhedral subdivision of $\mathcal{B}_A^\circ$.
Its maximal cells correspond to the types of simple polytopes $P_A(b)$ for which each a_i defines a facet.

Proposition 3.23. *$P \leq_w Q$ holds if and only if $\lambda Q = P + R$ for some $\lambda > 0$ and a polytope R.*
$P_A(b' + b'') = P_A(b') + P_A(b'')$ holds if and only if b' and b'' lie in the same closed type cone.

Corollary 3.24. *Restricted to a type cone $\mathcal{T}_A(b)$, the volume function is given in the form*

$$\mathrm{vol}(P_A(b)) = \mathrm{vol}(\lambda_1 P_1 + \cdots + \lambda_n P_n),$$

where P_i span the rays of the type cone $\mathcal{T}_A(b)/\mathrm{im}(A)$.

Theorem 3.25 (Minkowski's Theorem). *If $K_1, \ldots, K_n \subseteq \mathbb{R}^d$ are compact convex sets, then for $\lambda_1, \ldots, \lambda_n \geq 0$*

$$\mathrm{vol}(\lambda_1 K_1 + \cdots + \lambda_n K_n)$$

is a homogeneous polynomial of degree d.

Notation:

- $r := (r_1, \ldots, r_n) \in \mathbb{N}_0^n$,
- $|r| := r_1 + \cdots + r_n$,
- $\mathbb{N}_0^n(d) := \{r \in N_0^n : |r| = d\}$,
- $\lambda^r := \lambda_1^{r_1} \cdots \lambda_n^{r_n}$.

With this notation, the homogeneous polynomial of degree d in Minkowski's theorem can compactly be written as

$$V(\lambda_1, \ldots, \lambda_n) = \mathrm{vol}(\lambda_1 K_1 + \cdots + \lambda_n K_n) = \sum_{r \in \mathbb{N}_0^n(d)} c_r \lambda^r.$$

Proof. We proceed in three steps.

(1) It suffices to prove the theorem for the case when the K_i are convex polytopes. For this, we can use the following convergence result:

If $f_1, f_2, \ldots$ are homogeneous polynomials of degree d in n variables, and $\lim_{s\to\infty} f_s(\lambda) = f(\lambda)$ for all $\lambda \geq 0$, then $f(\lambda_1, \ldots, \lambda_n)$ is also a homogeneous polynomial of degree d.

Thus we can approximate the K_i by polytopes P_i better and better and use continuity.

(2) If the Minkowski sum is a direct sum ("the Minkowski sum is *exact*"),

$$\dim(P_1 + \cdots + P_n) = \dim(P_1) + \cdots + \dim(P_n),$$

then

$$\mathrm{vol}(\lambda_1 P_1 + \cdots + \lambda_n P_n) = \prod_{i=1}^{n} \lambda_i^{\dim(P_i)} \mathrm{vol}(P_1 + \cdots + P_n).$$

End of class on July 1

(3) To prove this part, we use "exact mixed subdivisions."

Notation:

- $P_i = \mathrm{conv}(V_i)$ for a finite set V_i,
- $S = (S_1, \ldots, S_n)$ with $S_i \subset V_i$,
- $\langle S \rangle := \mathrm{conv}(S_1) + \cdots + \mathrm{conv}(S_n) \subseteq P_1 + \cdots + P_n$,
- The *type* of S is $d(S) = (d_1, \ldots, d_n) = (\dim \mathrm{conv}(S_1), \ldots, \dim \mathrm{conv}(S_n))$,
- $\lambda \cdot S := (\lambda_1 S_1 \ldots, \lambda_n S_n)$,
- so $\langle \lambda \cdot S \rangle = \lambda_1 \mathrm{conv}(S_1) + \cdots + \lambda_n \mathrm{conv}(S_n)$.

Definition 3.26. A *mixed subdivision* of $P = P_1 + \cdots + P_n$ is a collection $\mathcal{S} \subseteq 2^{V_1} \times \cdots \times 2^{V_n} = \{(S_1, \ldots, S_n) : S_i \subseteq V_i\}$ if the polytopes

$$\langle S \rangle = \mathrm{conv}(S_1) + \cdots + \mathrm{conv}(S_n) \quad \text{for } S \in \mathcal{S}$$

form a subdivision of P.

In this subdivision, we also require that "faces fit together", that is, that $\langle S \rangle \cap \langle S' \rangle$ is of the form $\langle T \rangle$, there each $\mathrm{conv}(T_i)$ is a face of both S_i and S_i'.

The subdivision is *exact* if $\langle S \rangle$ is exact for all $S \in \mathcal{S}$.

It is called *fine* if it is exact and additionally the $\mathrm{conv}(S_i)$ are simplices with vertex set S_i.

Thus a mixed subdivision of $P = P_1 + \cdots + P_n$ consists of pieces of the form $F_1 + \cdots + F_n$ for $F_i = \mathrm{conv}(S_i)$ and $S_i \subseteq V_i$ with $\mathrm{conv}(V_i) = P_i$.

Examples! Examples:

- not mixed, not exact
- mixed, not exact
- not mixed, exact
- mixed, exact.

Now if we assume that we have an exact mixed subdivision $\mathcal{S}$ of $P = P_1 + \cdots + P_n$, and $\lambda_1, \ldots, \lambda_n > 0$, then also $\lambda \cdot \mathcal{S}$ is a mixed subdivision of $\lambda_1 P_1 + \cdots + \lambda_n P_n$ (check this: This uses the "faces fit together"-condition!).

Thus we get that

$$\mathrm{vol}(\lambda_1 P_1 + \cdots + \lambda_n P_n) = \sum_{\substack{S\in\mathcal{S}\\ \dim\langle S\rangle = d}} \mathrm{vol}_d(\langle \lambda \cdot S\rangle) = \sum_{\substack{S\in\mathcal{S}\\ \dim\langle S\rangle = d}} \lambda^{d(S)}\mathrm{vol}_d(\langle S\rangle).$$

This also holds for $\lambda \geq 0$ by continuity. This completes the proof of Minkowski's theorem, modulo existence of mixed subdivision. $\square$

Definition 3.27 (The Cayley embedding)**.** For polytopes $P_1, \dots, P_n \subset \mathbb{R}^d$ the Cayley embedding is

$$C(P_1, \dots, P_n) := \mathrm{conv}\Big(P_1 \times \{e_1\}, \dots, P_n \times \{e_n\}\Big) \subset \mathbb{R}^{d\times n}.$$

In particular, if $V_i \subseteq P_i$ are finite subsets with $\mathrm{conv}(V_i) = P_i$, this defines a subset $V(V_1, \dots, V_n) \subseteq C(P_1, \dots, P_n)$ by

$$V(P_1, \dots, P_n) := (V_1 \times \{e_1\}) \cup \cdots \cup (P_n \times \{e_n\}) \subseteq C(P_1, \dots, P_n)$$

with $\mathrm{conv}V(P_1, \dots, P_n) = C(P_1, \dots, P_n)$.

Theorem 3.28 (The Cayley trick)**.** *Let $P_1, \dots, P_n \subset \mathbb{R}^d$ be polytopes, then*

(i) *the mixed subdivisions of $P_1 + \cdots + P_n$ are in bijection with the subdivisions of $C(P_1, \dots, P_n)$ with vertex set contained in $V(P_1, \dots, P_n)$,*

(ii) *the fine mixed subdivisions of $P_1 + \cdots + P_n$ are in bijection with the triangulations of $C(P_1, \dots, P_n)$ with vertex set $V(P_1, \dots, P_n)$.*

In particular, as such triangulations of $C(P_1, \dots, P_n)$ exist, there are fine (and hence exact) mixed subdivisions of $P_1 + \cdots + P_n$.

Proof. We do not provide the proof here, but refer to De Loera et al. [2]. Note that the "easy" part is to verify that

(i) subdivisions of the Cayley polytope give mixed subdivisions of the (rescaled) Minkowski sum $\frac{1}{n}P_1 + \cdots + \frac{1}{n}P_n$,

(ii) triangulations of the Cayley polytope give mixed subdivisions of the (rescaled) Minkowski sum,

as they "give" this plainly by intersection of $C(P_1, \dots, P_n)$ with the subspace $H_{(\frac{1}{n}, \dots, \frac{1}{n})} := \mathbb{R}^d \times \{(\frac{1}{n}, \dots, \frac{1}{n}\}$ — and this "easy" part is exactly what we need to complete the proof of Minkowski's theorem. $\square$

End of class on July 3

3.6 The mixed volumes

Definition 3.29 (Mixed volume)**.** Let $K_1, \dots, K_d \subset \mathbb{R}^d$ be nonempty compact convex sets. The *mixed volume* $\mathrm{MV}(K_1, \dots, K_d)$ of $K_1, \dots, K_d$ is defined as the coefficient of $\lambda_1 \cdots \lambda_d$ in the polynomial $\mathrm{vol}_d(\lambda_1 K_1 + \cdots + \lambda_d K_d)$, that is, by

$$\mathrm{MV}(K_1, \dots, K_d) = \frac{\partial^d}{\partial\lambda_1 \cdots \partial\lambda_d} \mathrm{vol}_d(\lambda_1 K_1 + \cdots + \lambda_d K_d).$$

The mixed volume $\mathrm{MV}(K_1, \dots, K_d)$ of $K_1, \dots, K_d$ is 0 if $\dim(K_1 + \cdots + K_d) < d$.

Proposition 3.30 (Properties of the mixed volume)**.** *Let $K_1, \dots, K_d \subset \mathbb{R}^d$ be compact convex sets with $\dim(K_1 + \cdots + K_d) = d$.*

(i) $\mathrm{MV}(K_1, \dots, K_d)$ *is symmetric in the arguments.*
(ii) $\mathrm{MV} : \mathcal{K}_d \times \cdots \times \mathcal{K}_d \to \mathbb{R}$ *is continuous (in the space $\mathcal{K}_d$ of compact convex sets with a suitable metric, to be detailed later)*
(iii) $\mathrm{MV}(K_1, \dots, K_d) \geq 0$.
(iv) $\mathrm{MV}(K, \dots, K) = d!\mathrm{vol}_d(K)$.
(v) MV *is invariant under rigid motions T.*

Proof. (i) clear by definition.
(ii) approximation: volume is continuous.
(iii) approximate by polytopes, then note that MV is given by volumes of pieces in an exact mixed subdivision.
(iv) compute: $\mathrm{vol}_d(\lambda_1 K + \cdots + \lambda_d K) = (\lambda_1 + \cdots + \lambda_d)^d \mathrm{vol}_d(K)$.
(v) clear by definition. $\square$

Definition 3.31 (Mixed volumes with multiplicities)**.** Let $K_1, \dots, K_n \subset \mathbb{R}^d$ be compact convex sets with $\dim(K_1 + \cdots + K_n) = d$ and $r_1, \dots, r_n \in \mathbb{N}_0$ with $r_1 + \cdots + r_n = d$. Then

$$\mathrm{MV}(K_1[r_1], \dots, K_n[r_n]) := \mathrm{MV}(\underbrace{K_1, \dots, K_1}_{r_1}, \dots, \underbrace{K_n, \dots, K_n}_{r_n}).$$

Proposition 3.32 (The coefficients of the Minkowski polynomial are mixed volumes with multiplicities)**.** *For $K_1, \dots, K_n \subset \mathbb{R}^d$ be compact convex sets with $\dim(K_1 + \cdots + K_n) = d$,*

$$\mathrm{vol}_d(\lambda_1 K_1 + \cdots + \lambda_n K_n) = \sum_{r \in \mathbb{N}_0^n(d)} \mathrm{MV}(K_1[r_1], \dots, K_n[r_n]) \lambda_1^{r_1} \cdots \lambda_n^{r_1}.$$

Proof. For the n compact convex sets $K_1, \dots, K_n \subset \mathbb{R}^d$ let the Minkowski polynomial $\mathrm{vol}_d(\lambda_1 K_1 + \cdots + \lambda_n K_n)$ of degree d be $\sum_r c_r \lambda^r$. The coefficient c_r of $\lambda^r = \lambda_1^{r_1} \cdots \lambda_n^{r_1}$ in this polynomial can be obtained as

$$\frac{\partial^{r_1 + \cdots + r_n}}{\partial^{r_1}\lambda_1 \cdots \partial^{r_n}\lambda_n} \mathrm{vol}_d(\lambda_1 K_1 + \cdots + \lambda_n K_n) = r_1! \cdots r_n! c_r.$$

If on the other hand, let us take r_i copies of each K_i, and consider the resulting $r_1 + \cdots + r_n = d$ convex sets (with multiplicities). Their Minkowski polynomial is

$$\mathrm{vol}_d(\lambda_{11} K_1 + \cdots + \lambda_{1r_1} K_1 + \cdots + \lambda_{n1} K_n + \cdots + \lambda_{nr_n} K_n)$$

and the coefficient of $\lambda_{11} \cdots \lambda_{1r_1} \cdots \lambda_{n1} \cdots \lambda_{nr_n}$ in this polynomial is the mixed volume $\mathrm{MV}(K_1[r_1], \dots, K_r$
by definition. On the other hand we can compute the Minkowski polynomial as

$$\begin{aligned} & \mathrm{vol}_d(\lambda_{11}K_1 + \cdots + \lambda_{1r_1}K_1 + \cdots + \lambda_{n1}K_n + \cdots + \lambda_{nr_n}K_n) \\ = & \mathrm{vol}_d((\lambda_{11} + \cdots + \lambda_{1r_1})K_1 + \cdots + (\lambda_{n1} + \cdots + \lambda_{nr_n})K_n). \\ = & \sum_{r \in \mathbb{N}_0^n(d)} c_r(\lambda_{11} + \cdots + \lambda_{1r_1})^{r_1} \cdots (\lambda_{n1} + \cdots + \lambda_{nr_n})^{r_n} \end{aligned}$$

And the coefficient of $\lambda_{11} \cdots \lambda_{1r_1} \cdots \lambda_{n1} \cdots \lambda_{nr_n}$ in this polynomial is clearly also $r_1! \cdots r_n! c_r$. □

Corollary 3.33 (Multilinearity of mixed volumes)**.** *For $\alpha, \beta \geq 0$, and compact convex sets $K_1', K_1'', K_2, \dots, K_d$,*

$$\mathrm{MV}(\alpha K_1' + \beta K_1'', K_2, \dots, K_d) = \alpha \mathrm{MV}(K_1', K_2, \dots, K_d) + \beta \mathrm{MV}(K_1'', K_2, \dots, K_d).$$

Proof. Exercise? □

Corollary 3.34 (Mixed volumes are valuations)**.** *For $K_2, \dots, K_d$ compact convex sets, the function*

$$K \mapsto \mathrm{MV}(K, K_2, \dots, K_d)$$

on compact convex sets is a valuation*, that is,*

$$\mathrm{MV}(K, K_2, \dots, K_d) + \mathrm{MV}(L, K_2, \dots, K_d) = \mathrm{MV}(K \cap L, K_2, \dots, K_d) + \mathrm{MV}(K \cup L, K_2, \dots, K_d)$$

whenever $K \cup L$ convex.

End of class on July 8