Geometric and Topological Combinatorics in Economics, Game theory, & Optimization.

J. A. De Loera

February 16, 2016

1 Lecture 1: Overview and Goals of the Course

These 20 lectures tell the story of an intellectual trip that starts in areas of mathematics, geometry and topology, that are known for their pure nature, but the trip ends with central concerns in discrete applied mathematics and mathematical economics. The ultimate goal is to revisit some beautiful problems in game theory, and optimization. Indeed, geometric and topological theorems have been used, among many others places, for understanding optimality criteria of mathematical optimization, equilibrium theorems for games, and in algorithms for fair-division and voting. Although some of the theorems presented are well-over hundred-years-old we try to visit new generalizations and extensions that promise future new applications. We imagine the typical reader is a graduate student, who may not have seen this part of mathematics, but we hope you will find the trip pleasurable. Let us mention five motivating applications examples.

- *Envy-free Cake-cutting, Rental Harmony problems.*

  Say n people wish to divide a rectangular cake. Each of them has an idea of what they consider fair or sufficient. We will discuss a model and an algorithm that proves the following fact, known to economists as the *fair-division theorem.*

  Theorem 1.1. *Any group of n people can divide a cake amongst themselves in an envy-free way, i.e., the division is done so that each person believes her piece is the biggest. More precisely: Consider a one-dimensional cake (line segment), it can be divided into n pieces by $n-1$ cuts in such a way that each player receives his/her preferred piece.*

First of all, preferences are expressed by a measure function for each of the players. We make some reasonable assumptions, such as no player will ever prefer an empty piece of cake. A more technical assumption is that if a player prefers the same piece within an infinite set of possible divisions, because of the nature of the space of all possible divisions there must be at least one limit division. We make a continuity assumption that the limiting divisions

Here is a perhaps surprising variation of the fair-division theorem

Theorem 1.2. *Consider a house with n rooms and a total rent amount to be divided among n roommates. Assume that for each possible division of the rent amount any roommate can point to one or more room as preferred. Then there exists a division of the rent and an assignment of rooms to each participant, such that each player receives one of his/her preferred rooms.*

The same assumptions we made in the previous theorem hold to make this happen. The proofs we will provide of these two theorems are grounded in Sperner's lemma. An algorithmic solution follows.

- *Necklace-splitting among thieves, the Ham-sandwich theorem, geometric data processing.*

  k thieves have stolen a necklace with n different types of precious stones (think rubies, diamonds, etc). Luckily the number of times each type occurs is divisible by k. The thieves intend to divide the necklace in such a way that they do as few cuts as possible and each gets the same number of jewels of each type. The question is can this always be done and what is the smallest possible number of cuts possible? We will investigate this problem using the Borsuk-Ulam theorem a topological consequence of Fan-Tucker's lemma.

  Another consequence of the Borsuk-Ulam theorem is the *ham sandwich theorem.* Consider a sandwich consisting of two slices of bread and a slice of ham in between. The theorem asserts that one can always make a straight cut in such a way that each of the three pieces is perfectly divided in half. More generally one can prove (adding layers of cheese, salami, etc. works too!):

  Theorem 1.3. *Given n continuous probability measures in $\mathbb{R}^n$ there exist an an affine hyperplane such that each measure takes value $1/2$ on each of the two half-spaces defined by the hyperplane.*

There are recent advanced generalizations of the ham-sandwich theorem. In [] the authors show that, for any prime power n and any compact convex set with interior $K \subset \mathbb{R}^d$, there exists a partition of K into n convex sets with equal volumes and equal surface areas. The goal is to have an equi-partition among n players of a divisible good. In the plane this boils down to the following fact proven in []. Any convex body in the plane can be partitioned into n convex regions with equal areas and equal perimeters. The theorem has received the funny name the *spicy chicken theorem* because it can be applied to equi-partition of a (perfectly convex) chicken among guests. You will cut the raw chicken fillet with a sharp knife, marinate each of the pieces in a spicy sauce, and then fry the pieces. The surface of each piece will be crispy and spicy, so the challenge is to cut the chicken so that all your guests get the same amount of crispy (surface) crust and the same volume amount of chicken meat. The theorem shows it is possible! But it is an open question on how to carry on the partition.

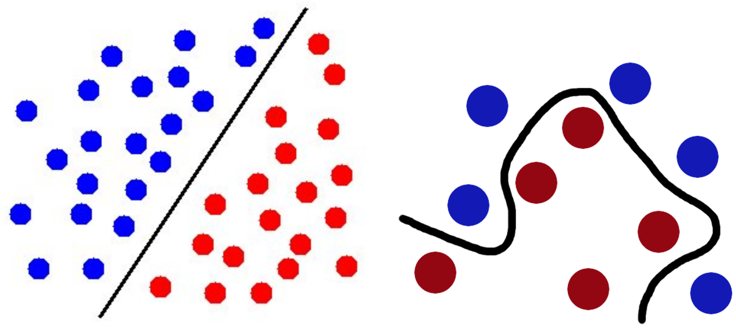
Another fascinating result, which aims to find *good portion* divisions rather than envy-free divisions, is the following proposition, which we state for the plane but generalizes to all dimensions.

Theorem 1.4. *Given any compact set K in the plane, there exist a point p in K such that no matter which line one traces passing through p leaves at least $\frac{1}{3}$ of the area of K in each side of the line.*

The techniques to prove the above result, and algorithmically find the point too, are a consequence of Helly's theorem. Helly's theorem has many surprising applications. Let us mentioned two more.

Suppose that X is set of discrete data points made of red points R and blue points B. suppose that for any subset $S \subset X$ in $\mathbb{R}^d$ of cardinality no more than $d+2$, there exists a hyperplane which strictly separates and $S \cap B$ and $S \cap R$. Then there exists a hyperplane which strictly separates the sets R and B. This is a result useful for the classification of data points (e.g., bad vs good drivers? democrats vs republicans voters?).

Similarly, in some situations we are interested on the clustering of data by points that are nearby or are similar. We may wish to find the smallest enclosing ball that contains the points. One can prove that, given a finite set X of points in $\mathbb{R}^d$, X is contained in a ball of radius r if every $d+1$ of its points are contained in a ball of radius



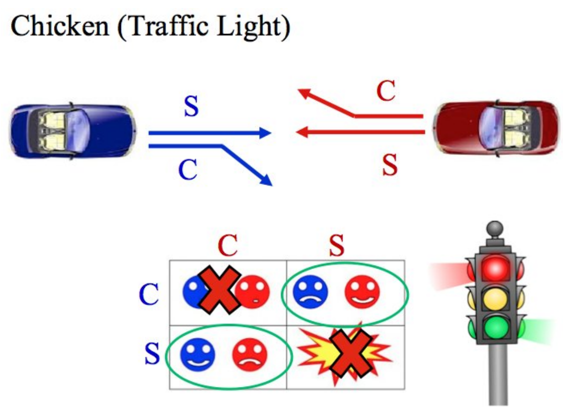
r. For example,if you are given s points in the plane such that every three of them are contained in a disk of radius 1, then all s points are contained in a disk of radius 1. The techniques are stronger and one can use them to solve the following problem:

Suppose the points represent bad objects that need to be contained in a smallest-radius enclosing ball, but we have lack of exactly how many and where the points are. There is a risk. You are given points $(u_1, u_2, \ldots, u_d) \in R^d$, belonging to an unknown measurable set. Your goal is to find the center $\mathbf{x}$ the ball of smallest radius R that contains a "large proportion" of those points. E.g., if these are cancerous cells, you do not wish to loose more than 1 percent of the bad cells. But the bad news you may not know explicitly the probability measure. The are sampling algorithms to attempt this problem. Clearly this is an stochastic optimization problem that can be formulated as

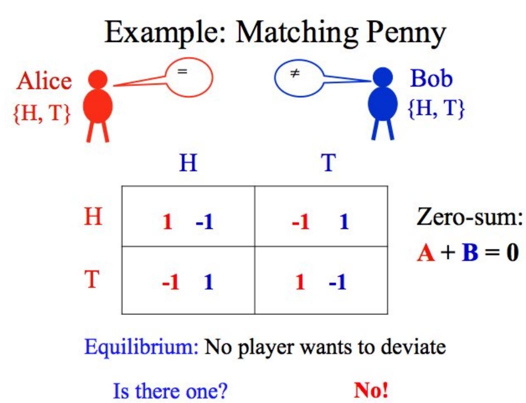
$$\begin{aligned} \min \quad & R \\ \text{subject to} \quad & Pr[\{\sqrt{\sum_1^d (x_i - u_i)^2} - R \leq 0\}] \geq 1 - \epsilon, \mathbf{x} \in \mathbb{R}^{d+1}. \end{aligned}$$

- *The games of chicken, matching pennies & Nash Equilibria.*

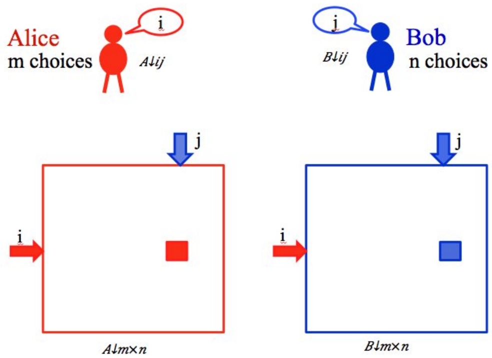
  The purpose of theory of games is to analyze or design systems where rational selfish agents interact to achieve certain goals. Consider two examples. First we have the game of chicken, where two people drive in a one lane street facing each other. The idea is that if one is a chicken you swerve the car before impact. If both drivers swerve away then they are both chickens (not a very pleasant nickname). Of course if nobody is chicken the outcome is fatal. What is the best strategy to follow? Clearly, no *pure strategy* will suffice to please both players. A key idea to find a compromise, some kind of stable solution, is to have a traffic light that takes turns assigning whose turn is it to swerve away before from crashing. This method has a superior payoff over the pure strategies (and preserves the lifes of players).



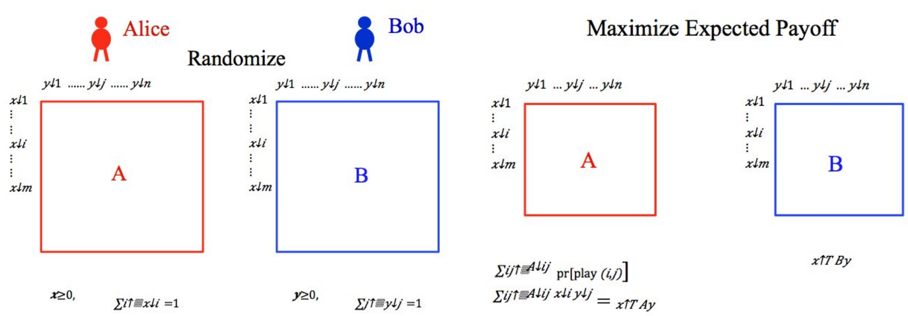
A similar game is *matching penny game*. Players Alice and Bob simultaneously showing heads or tails of a coin. If the choice is the same Bob wins one penny Alice looses a penny, if they choose differently then Alice wins a penny and Bob looses a penny. These are the payoffs of the game.



The payoffs for each player can be recorded in two matrices (A and B, for each player). In the match penny game the payoff matrices can be put together to show $A + B = 0$. We call such games *zero-sum games*. For a game with two players this is not always the case (e.g., for game of chicken!) A *pure strategy* consists of Alice selecting a single row to be played and Bob a single column to be played. An *equilibrium strategy* is one no player wants to deviate from.



The problem is that, as we saw twice already, there are no pure strategies that are equilibria! But as in the case of the game of chicken mixing pure strategies is gives a suitable compromise. *mixed strategies* are linear combinations of the pure strategies proposed in the payoff matrices. We think of the choices being made randomly. Alice has a choice x of probabilities in which to choose each pure strategy, Bob has a vector y of probablities in which to choose his pure strategies. They will have an expected payoffs $x^T Ay$, and $x^T By$ respectively.



John von Neumann was a pioneer on the study of equilibria for zero-sum games. He showed equilibria existed in the 1920's. Later in 1950's, with the emergence of linear optimization and the simplex method Dantzig showed that zero-sum games are special in that equilibria can be computed solving a linear optimization problem. In 1949, John Nash showed that in any game there is always at least one mixed strategy that is an equilibrium solution. We call them Nash equilibria. Later we will prove Nash's theorem for which he received the Nobel prize in Economics. The existence of Nash equilibria is proven using fixed-point theorem, such as Brouwer's and Kakutani's theorem. They in turn are consequences of Sperner's lemma. We will look at the details of these theorems and try to discuss some of the computational methods.

- *Coin-exchange and bin-packing problems.*

  Suppose we are given some **coins** of different denominations. One can ask the following natural questions (try to answer them for the two coins in the picture):

  1. How many ways are there to give change for b cents?
  2. What is the smallest number of coins necessary to do so?
  3. What is largest quantity b that cannot be expressed using the coins?
  4. Which values of b have exactly 20 ways to be broken in change?

  Another similar family of problems is that for the bin-packing problems The problem is we are given n items and n bins. Item j has size s_j and bin i has capacity c. The goal is to assign each item to a bin (to pack the bins!) so that the total weight of the items in each bin does not exceed c, but the number of bins used is smallest possible. This is a difficult problem and for the most part one is interested on good approximation algorithms.

Let us consider one instance of the problem. Let (s, a) be an instance for *bin packing* with *item sizes* $s_1, \ldots, s_d \in [0,1]$ and a vector $a \in \mathbb{Z}^d_{\geq 0}$ of *item multiplicities*. In other words, our instance contains a_i many copies of an item of size s_i. (we assume that s_i is given as a rational number and Δ is the largest number appearing in the denominator of s_i or the multiplicities a_i.)

Consider the polytope $\mathcal{P} := \{x \in \mathbb{Z}^d_{\geq 0} \mid s^T x \leq 1\}$. Now the bin packer's objective is to select a minimum number of vectors from $\mathcal{P}$ that sum up to a,

$$\min \left\{ \mathbf{1}^T \lambda \mid \sum_{x \in \mathcal{P}} \lambda_x \cdot x = a; \;\; \lambda \in \mathbb{Z}^{\mathcal{P}}_{\geq 0} \right\} \tag{1}$$

where λ_x is the *weight* that is given to $x \in \mathcal{P}$. This special case is known as the *(1-dimensional) cutting stock problem.*

Bin packing and the cutting stock problem belong to a family of problems that consist of selecting integer points in a polytope with multiplicities. In fact, several scheduling problems fall into this framework as well, where the polytope describes the set of jobs that are admissible on a machine under various constraints. Recently Goemans and Rothvoss showed

Theorem 1.5. *For any Bin Packing instance (s, a) with $s \in [0,1]^d$ and $a \in \mathbb{Z}^d_{\geq 0}$, an optimum integral solution can be computed in time $(\log \Delta)^{2^{O(d)}}$ where Δ is the largest integer appearing in a denominator s_i or in a multiplicity a_i.*

Fundamentally, one is interested on finding the sparsest representation of a vector b from a list of vectors $X = (x_1, \ldots, x_t) \subset \mathbb{R}^d$ which we think of as the columns of the matrix A In the case of the cutting stock problem the vectors x_i are the possible packing patterns. They define the following *sparse representation* problem.

$$\min \|x\|_0, \; Ax = b, x \geq 0, x \in \mathbb{Z}^t. \tag{2}$$

Here, $\|\cdot\|_0$ denotes the 0-norm, which counts the cardinality of the *support* of x, i.e. $\mathrm{supp}(x) = \{i : x_i \neq 0\}$. In other words, the value of

$\|x\|_0$ equals the number of non-zero entries in the vector x. Problem (2) aims to find the vector of minimal support.

There is a rich literature about this problem. The sparsest solution of a system of linear equations is quite important in applications to signal processing [?], cryptography and coding theory [?]. Sparse integer solutions also appear in the context of finding guarantees for bin-packing problems via the Gilmore-Gomory formulation [?], as first suggested in [?]. More generally, upper bounds given for the size of the sparsest integer solution indicate that if there exists an optimal solution to such an integer program, then there exists one which is polynomial in the number of equations and the maximum binary encoding length among integers in the objective function vector and the constraint matrix (see Section 3 in [?]). The sparsity of the solution is also strongly connected to the *integer Carathéodory problem* (see [?] and the references there).

It is known that even for real variables, the 0-norm minimization is NP-hard [?], but that the greedy algorithm provides a guaranteed approximation in this setting. Moreover, when one looks at random matrices, one can prove nice properties for the size of the solution [?]. It is precisely such structural differences that we wish to study here for integer solutions.

There are two important geometric objects associated with the sparse representation problem. First, the *conic hull* of X is the set

$$\operatorname{cone}(X) = \{\lambda_1 x_1 + \cdots + \lambda_t x_t : x_1, \ldots, x_t \in X, \lambda_1, \ldots, \lambda_t \in \mathbb{R}_{\geq 0}\},$$

and the *semigroup* of X or the *integer conic hull* of X is the set

$$\operatorname{Sg}(X) = \{\lambda_1 x_1 + \cdots + \lambda_t x_t : x_1, \ldots, x_t \in X, \lambda_1, \ldots, \lambda_t \in \mathbb{Z}_{\geq 0}\}.$$

For each $b \in \operatorname{Sg}(X)$, we are interested in finding upper bounds and the asymptotic behavior of the function

$$\mathsf{m}_0(b) \quad = \quad \min\{\|a\|_0 : a \in \mathsf{P}_X(b)\},$$

where $\mathsf{P}_X(b) = \{a \in \mathbb{Z}^t_{\geq 0} : a_1 x_1 + \cdots + a_t x_t = b\}$ is the solution set for b. Note that the convex hull of $\mathsf{P}_X(b)$ is a lattice polytope.

In the context of the applications, the upper bounds for $\mathsf{m}_0(b)$ are of special interest. The problem of estimating $M_0(X) := \max_{b \in \operatorname{Sg}(X)} \mathsf{m}_0(b)$

goes back to classical results on the integer Carathéodory problem. Cook, Fonlupt, and Schrijver [?] showed that $M_0(X) \leq 2d - 1$ if $C = \text{cone}(X)$ is pointed and X forms a Hilbert basis of C. This result was later improved by Sebő [?] to $M_0(X) \leq 2d-2$. It remains an open question to determine the exact value even when X is a Hilbert basis. For an arbitrary set $X \subset \mathbb{Z}^d$, Eisenbrand and Shmonin [?] obtained the bound

$$M_0(X) \leq 2d \log(4d\|X\|_\infty), \tag{3}$$

where $\|X\|_\infty = \max_{x \in X} \|x\|_\infty$.

There are two challenging open problems: What are the optimal bounds for $\mathsf{m}_0(b)$ in terms of the generating set X? What is the asymptotic behavior of the univariate function $f_0(\lambda) := \mathsf{m}_0(\lambda b)$ obtained from successive dilations of the vector b?

Let us conclude this introduction with a bird-view of the four foundational combinatorial theorems from geometry and topology. First, two from combinatorial topology, Sperner's and Fan-Tucker's lemmas, and then, second, two corner stones of combinatorial convex geometry, Carathéodory, and Helly theorems. We will see how these four "mathematical cornerstones" are deeply interrelated among themselves (e.g., the topological results imply the the geometric statements, while they imply each other in some form or another), and each theorem, is in fact a family of theorems of certain type, with many generalizations, corollaries and extensions. The landscape is so lovely one is often tempted to stay longer at many stops, but remember our final destination is the world of applications!

First, Sperner's Lemma is a combinatorial statement about labelings of triangulated simplices. It is quite well-known as equivalent with the topological fixed-point theorem of Brouwer [?, ?].

Theorem (Sperner's lemma, 1928 []). *Let T be a triangulation of a $(n-1)$-simplex, and suppose that the vertices of T have a labeling satisfying these conditions: each vertex of the triangulation is assigned a unique label from the set $\{1, 2, \ldots, n\}$, and each other vertex v of T is assigned a label of one of the vertices of P in carr$(\{v\})$.*

Any Sperner labeling of a triangulation T of the d-simplex must contain an odd number of cells for which all their labels are distinct. In particular, there is at least one such cell.

Second, the Fan's lemma is a combinatorial analogue of the famous Borsuk-Ulam theorem (see []). We need to start with a little bit of terminology and notation. Denote by S^d be the unit d-sphere, the set of all points of unit Euclidean distance from the origin in $\mathbb{R}^{d+1}$. Any pair of points in S^d of the form $x, -x$ is a pair of *antipodes* in S^d. A triangulation of S^n has an *anti-symmetric* labeling ℓ if $\ell(-v) = -\ell(v)$ for all vertices v. A labeling has a *complementary edge* if some adjacent pair of vertices has labels that sum to zero. A simplex is *alternating* if its vertex's labels are distinct in magnitude and alternate signs, when arranged in order of increasing value.

Theorem (Fan's lemma 1946 []). *Let T be a symmetric triangulation of S^n with an m-labeling that is anti-symmetric and has no complementary edge. Then has at least one positive alternating n-simplex.*

We recall now Carathéodory's, and Helly's theorems. They are clearly among the most important theorems in convex geometry.

Theorem (C. Carathéodory 1911 [**?**]). *Let S be any subset of $\mathbb{R}^d$. Then each point in the convex hull of S is a convex combination of at most $d+1$ points of S.*

Theorem (E. Helly, 1913 [**?**]). *Let $\mathcal{F}$ be a finite family of convex sets of $\mathbb{R}^d$. If $\bigcap \mathcal{K} \neq \emptyset$ for all $\mathcal{K} \subset \mathcal{F}$ of cardinality at most $d+1$, then $\bigcap \mathcal{F} \neq \emptyset$.*

We chose these four theorems because they are centrally located and essential! One can sense this because they imply many of the later results and they have many corollaries generalizations and extensions. Our course presents a interconnected theory and we can see what theorems imply others, E.g., Sperner's lemma implies Helly theorems. Fan-Tucker implies Carathéodory.

(see the diagram of implications below).

These four fantastic theorems and their variations are key for discrete applied mathematics today; a fact we will demonstrate with plenty of examples.

2 Combinatorial Topology Tools: Midterm 1

A *triangulation* is a subdivision by simplices that meet either face-to-face or not at all. Each simplex is the affine hull of its vertices; these are the vertices of the triangulation. We are interested on coloring or labeling the vertices of a triangulated manifold, most often a triangulated ball or a sphere following certain rules and then make conclusions about multicolored simplices inside the triangulation.

2.1 Lectures 2: Preliminaries

The **line segment** joining two points $x, y \in \mathbb{R}^d$ is given by the set of all points of the form

$$[x, y] := \{\gamma x + (1 - \gamma)y : 0 \leq \gamma \leq 1\}.$$

A set $A \subset \mathbb{R}^d$ is **convex** if it contains the line segment joining two of its points, for every pair of points in the set. A lot of our arguments will rely on convex sets.

Example 1. *Figure 1 shows some convex sets in* $\mathbb{R}^3$*, whereas Figure* **??** *are instances of nonconvex sets in* $\mathbb{R}^2$*.*

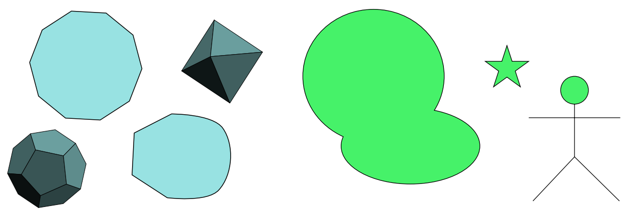


Figure 1: Examples of convex figures (left) and non-convex figures (right)

It may be easy for humans to determine whether or not a given figure is convex through straightforward observation. However, in practice, a computer is only able to understand the convexity of a figure by a set of inequalities, and as the examples above show, whether or not a set is convex is not always immediately clear from input inequalities. It is therefore desirable to have a clear criterion that can be implemented to determine whether or not a given figure in $\mathbb{R}^d$ is convex.

Proposition 2.1. *The intersection of a (possibly infinite, possibly uncountable) collection convex sets is convex.*

Proof. Let $\{C_\beta : \beta \in I\}$ be a (possibly infinite) collection of convex sets in $\mathbb{R}^d$ and let $\gamma \in [0, 1]$.

Suppose that $x, y \in \bigcap_{\beta \in I} C_\beta$. Then $x, y \in C_\beta$ for each $\beta \in I$, and since each C_β is convex,

$$z = \gamma x + (1 - \gamma)y \in C_\beta$$

for each $\beta \in I$.

Thus, $z \in \bigcap_{\beta \in I} C_\beta$, so that the intersection of these collection of convex sets is again convex. $\square$

Hyperplanes and Half-Spaces

For every nonzero $c \in \mathbb{R}^d$, we may associate c with a linear functional $f : \mathbb{R}^d \to \mathbb{R}$. An example of this is $f : \mathbb{R}^d \to \mathbb{R}$ where $f(x) = c \cdot x$, with the usual dot product in $\mathbb{R}^d$.

Definition 2.2. *For $\alpha \in \mathbb{R}$, we say that $H_\alpha = \{x \in \mathbb{R}^d : f(x) = \alpha\}$ is an* ***affine hyperplane*** *or simply a* ***hyperplane****.*

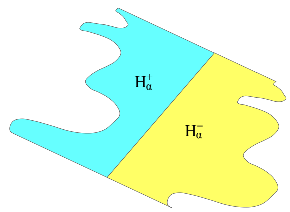


Figure 2: A hyperplane dividing $\mathbb{R}^2$ into two half-spaces.

An ***affine space*** *is an intersection of finitely many hyperplanes. The* ***affine hull*** *of a set $A \subset \mathbb{R}^d$, denoted* aff(A) *is the smallest affine space containing A.*

Notice that affine planes are always convex, as they are finite intersection of hyperplanes, which are themselves convex. Thus, affine spaces provide important examples of convex sets as they provide meaning to the dimension of a convex set. Hyperplanes correspond to what are called **level sets** of linear functions. The **dimension** of an affine space A in $\mathbb{R}^d$ is the largest number of **affinely independent points** in A minus one. The **dimension** of a convex set C in $\mathbb{R}^d$ is the dimension of aff(C).

Observe that every hyperplane in $\mathbb{R}^d$ divides $\mathbb{R}^d$ into two **half-spaces**, namely

$$H_\alpha^+ = \{x \in \mathbb{R}^d : f(x) \geq \alpha\} \text{ and } H_\alpha^- = \{x \in \mathbb{R}^d : f(x) \leq .\alpha\}$$

Note that equalities may be used to define half-spaces, as they can be decomposed into two inequalities. This will come up later when we discuss polytopes.

We can formally denote half-spaces as convex sets defined by a **linear inequality** of the form

$$a_1x_1 + ... + a_dx_d \leq b.$$

Given a convex set C in $\mathbb{R}^d$, a linear inequality $f(x) \leq \alpha$ is said to be **valid** on C if every point in C satisfies that inequality. A subset F a convex set C is a **face** of C if there exists a linear inequality $f(x) \leq \alpha$ which is valid on C and that $F = \{x \in C : f(x) = \alpha\}$. Whenever this is the case, the hyperplane defined by $f(x) = \alpha$ is a **supporting hyperplane** of F.

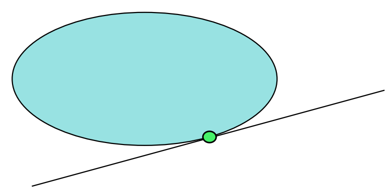


Figure 3: A convex set in $\mathbb{R}^2$ with a supporting hyperplane and a vertex as its face.

Definition 2.3. *A face of dimension zero is called a* ***vertex****. A face of dimension one is called an* ***edge****, and a face of dimension* $\dim(C) - 1$ *is called a* ***facet****.*

The following are some standard properties of well-behaved convex sets.

Lemma 2.4. *Let C be a closed and bounded convex set in $\mathbb{R}^d$. Let $x_0 \notin K$. Then*

1. *There is a unique nearest point x_1 of C to x_0.*
2. *The hyperplane H through x_1 orthogonal to $x_1 - x_0$ is a supporting hyperplane of C.*

Theorem 2.5. *A convex set C is the intersection of its closed supporting half-spaces. In other words Convex sets are the sets of solutions of (possibly infinite) systems of linear inequalities.*

Theorems 2.5 will not be proved here, but shall be assumed as true from this point onward. Both theorems suggest that it is sufficient to consider intersection of regions specified by linear inequalities to describe convex sets Convex sets arise naturally in the discussion of linear programming and optimization. Suppose you have the problem of finding a vector $(x_1, ..., x_d)$

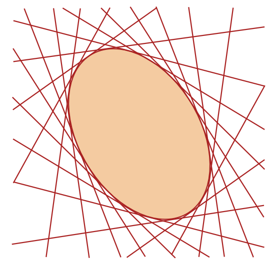


Figure 4: Convex set as an intersection of half-spaces.

that satisfies all of the following:

$$\begin{aligned} a_{1,1}x_1 + ... + a_{1,d}x_d &\leq b_1 \\ a_{2,1}x_1 + ... + a_{2,d}x_d &\leq b_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{k,1}x_1 + ... + a_{k,d}x_d &\leq b_k \end{aligned}$$

This is called the **linear feasibility problem**, and is an example of a standard problem in the field. The set of points satisfying the constraints constitute the **feasible** region or feasible set. If the problem has no solution, it is called **infeasible**. If the feasible set is convex, solutions may be combined to produce new solutions, so this is considered desirable. A feasible set may be bounded or unbounded. For a problem with n variables, it is a necessary but not sufficient condition that the number of constraints at least n + 1 in order for the feasible set to be bounded.

In our future considerations, we will be interested in **polyhedra**, convex figures that are the intersection of finitely many half-spaces. An example illustrating a **polyhedron** is shown in Figure 5.

Example 2. *The following are instances of a* ***polyhedron****:*

1. *The d-dimensional unit cube*

$$C_d = \{x \in \mathbb{R}^d : 0 \leq x_i \leq 1, i = 1 \ldots d\}.$$

2. *The* $(d-1)$*-dimensional standard simplex*

$$\Delta_{n-1} = \{x \in \mathbb{R}^d : \sum_{i=1}^{d} x_i = 1, x_i > 0\}.$$

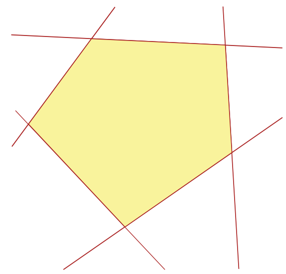


Figure 5: Convex set described with finitely many linear inequalities.

3. *The d-dimensional cross-polytope*

$$O_d = \{x \in \mathbb{R}^d : \sum_{i=1}^{d} |x_i| \leq 1\}.$$

4. *The* ***simplotope****, which is the Cartesian product of several simplices*

$$\Delta_{m_1} \times \Delta_{m_2} \times \ldots \times \Delta_{m_r}$$

All of the above are indeed intersections of finitely many half-spaces. C_d can be seen as the intersection of all half-spaces of the form $x_i \leq 1$ and $x_j \geq 0$, where $i, j = 1, \ldots, d$. Every equality $f(x) = \alpha$ can be seen as the conjunction of the two inequalities $f(x) \leq \alpha$ and $f(x) \geq \alpha$, so δ_{n-1} is indeed a polytope.

Meanwhile, there are 2^d possible linear inequalities describing O_d. For example, in $\mathbb{R}^3$, O_3 is specified as the following set of 8 inequalities:

$$-1 \leq x_1 + x_2 + x_3 \leq 1,$$
$$-1 \leq x_1 + x_2 - x_3 \leq 1,$$
$$-1 \leq x_1 - x_2 + x_3 \leq 1,$$
$$-1 \leq x_1 - x_2 - x_3 \leq 1.$$

Each inequality specifies a supporting hyperplane for a face of O_3. The reader is invited to verify that O_3 is then a regular octahedron as shown at top right of Figure 1. On the other hand, it is not immediately clear that simplotopes are polyhedra. This is indeed the case, and we leave it as an exercise for the reader to verify the more general claim that the Cartesian product of finitely many polyhedra is again a polyhedron.

Convex and Affine Combinations

Even though not every shape in nature appears as a convex set, we may always use convex sets to approximate these shapes!

Definition 2.6. *Let $A \subset \mathbb{R}^d$. The **convex hull** of A, denoted by* $\operatorname{conv}(A)$*, is the intersection of all the convex sets containing A, that is, it is the smallest convex set that contains A.*

We usually denote by $\operatorname{conv}(a_1, \ldots, a_n)$ *for the convex hull of* $\{a_1, \ldots, a_n\} \subset \mathbb{R}^d$.

Definition 2.7. *A **polytope** is the convex hull of a finite set of points in $\mathbb{R}^d$. It is the smallest convex set containing the points.*

Definition 2.8. *Given finitely many points $A = x_1, x_2, ..., x_d$, we say the linear combination $\sum_{i=1}^{d} \gamma_i x_i$ is*

- *a **conic combination** if all γ_i are nonnegative.*
- *an **affine combination** if $\sum_{i=1}^{d} \gamma_i = 1$*
- *a **convex combination** if it is both a conic and affine combination.*

Lemma 2.9. *For a set of points A in $\mathbb{R}^d$ we have that conv(A) equals the set of all finite convex combinations of points in A.*

Sketch of Proof Denote B as the set of all finite convex combinations of points in A. In other words, $x \in B$ if and only if there is a finite subset $S = \{x_1, \ldots, x_n\}$ of A with $x = \sum_{i=1}^{n} \gamma_i x_i$, $\gamma_i \geq 0$ for all i and $\sum_{i=1}^{n} \gamma_i = 1$. We need to prove that $B = \operatorname{conv}(A)$.

Let us first prove that B is a convex set containing A. Obviously, each $x \in A$ can be expressed as a finite convex combination, that is $x = 1 \cdot x$, so $A \subset B$. Now, if $u, v \in B$, then we may assume that there is a common subset $S = \{x_1, \ldots, x_n\}$ of A such that u and v can be expressed as convex combinations of points in S (Why?). Suppose that

$$u = \sum_{i=1}^{n} \gamma_i x_i$$

and

$$v = \sum_{i=1}^{n} \gamma_i' x_i,$$

where

$$\gamma_i, \gamma_i' \geq 0$$

and

$$\sum_{i=1}^{n} \gamma_i = \sum_{i=1}^{n} \gamma_i' = 1.$$

Thus, if $\lambda \in [0, 1]$, then we obtain

$$\begin{aligned} \lambda u + (1-\lambda) &= \sum_{i=1}^{n} \lambda \gamma_i x_i + (1-\lambda)\gamma_i' x_i \\ &= \sum_{i=1}^{n} \lambda \gamma_i + (1-\lambda)\gamma_i']x_i. \end{aligned}$$

For each i, we have

$$\lambda \gamma_i + (1-\lambda)\gamma_i' \geq \lambda(0) + (1-\lambda)(0) = 0,$$

and moreover,

$$\begin{aligned} \sum_{i=1}^{n} \lambda \gamma_i + (1-\lambda)\gamma_i' &= \lambda \sum_{i=1}^{n} \gamma_i + (1-\lambda) \sum_{i=1}^{n} \gamma_i' \\ &= \lambda(1) + (1-\lambda)(1) \\ &= 1. \end{aligned}$$

Thus, $\lambda u + (1-\lambda)v$ is also a finite convex combination of points in A, so B is indeed a convex set containing A. We then have $\mathrm{conv}(A) \subset B$ by the definition of convex hull of A.

The interested reader is then invited to prove that $B \subset \mathrm{conv}(A)$ by showing that if C is any convex set containing A, then $B \subset C$. It then follows that $B \subset \mathrm{conv}(A)$, so $\mathrm{conv}(A)$ is precisely the set of all finite convex combinations of points in A. □

Definition 2.10. *A set of points $\{x_1, ..., x_k\}$ is **affinely dependent** if there is a nontrivial linear combination $\sum_{i=1}^{k} \gamma_i x_i = 0$ with $\sum_{i=1}^{k} \gamma_i = 0$. Otherwise, it is said to be **affinely independent**.*

Lemma 2.11. *A set of $d+2$ or more points in $\mathbb{R}^d$ is affinely dependent.*

Proof. Suppose that $x_1, \ldots, x_k$ are vectors in $\mathbb{R}^d$, and consider the vectors $x_1', \ldots, x_k'$ in $\mathbb{R}^{d+1}$ where x_i' is the vector whose first d components coincide with those of x_i.

Since $k > d + 1$, then this collection of vectors is linearly dependent, that is, there are $\gamma_i \in \mathbb{R}$ with $\sum_{i=1}^k \gamma_i x_i' = 0$ and $\gamma_1, \ldots, \gamma_k$ are not all zero. Thus, we have both $\sum_{i=1}^k \gamma_i x_i = 0$ and $\sum_{i=1}^k \gamma_i = 0$ by considering the linear dependence on the first d components and on the last component of x_i' separately.

Furthermore, this linear combination is nontrivial, because, some γ_j is nonzero. By definition, $x_1, \ldots, x_k$ forms an affinely dependent set. $\square$

We have the following characterization of affinely independent sets in $\mathbb{R}^d$.

Lemma 2.12. *A set $B \subset \mathbb{R}^d$ is affinely independent if and only if every point in* $\operatorname{aff}(B)$ *has a unique representation as an affine combination of points in B.*

Proof. Suppose that $B = \{x_1, \ldots, x_k\}$ and that some point x in $\operatorname{aff}(B)$ admits two different affine combinations

$$x = \sum_{i=1}^{k} \beta_i x_i,$$

and

$$x = \sum_{i=1}^{k} \gamma_i x_i,$$

with $\beta_j \neq \gamma_j$ for some j and

$$\sum_{i=1}^{k} \beta_i = \sum_{i=1}^{k} \gamma_i = 1.$$

Then, by taking differences,

$$\sum_{i=1}^{k} (\beta_i - \gamma_i) x_i = 0,$$

with

$$\sum_{i=1}^{k} (\beta_i - \gamma_i) = \sum_{i=1}^{k} \beta_i - \sum_{i=1}^{k} \gamma_i = 1 - 1 = 0.$$

This linear combination of 0 is nontrivial, because $\beta_j - \gamma_j \neq 0$, so that B is an affinely dependent set. Hence, if every point in $\operatorname{aff}(B)$ is uniquely expressed as an affine combination of points in B, then B must be affinely independent.

On the other hand, if $x \in \operatorname{aff}(B)$, then it can be shown that x can be expressed as an affine combination of some points in B. In other words, $x = \sum_{i=1}^{m} \alpha_i x_i$ and $\sum_{i=1}^{m} \alpha_i = 1$ where $\{x_1, \ldots, x_m\} \subset B$. If B is affinely independent, then B cannot have more than $d+1$ points by Lemma 2.11, let $B = \{x_1, \ldots, x_{k+1}\}$ for some $k \leq d$.

We claim that the set $\{x_1 - x_{k+1}, \ldots, x_k - x_{k+1}\}$ is linearly independent over $\mathbb{R}$. Suppose that $\sum_{i=1}^{k} \gamma_i (x_1 - x_{k+1}) = 0$, for some $\gamma_i \in \mathbb{R}$, $i = 1, \ldots, k$. Thus, $\sum_{j=1}^{k+1} \beta_j x_j = 0$, where $\beta_j = \gamma_j$ for $j = 1, \ldots, k$ and $\beta_{k+1} = -\sum_{j=1}^{k} \gamma_j$, so that $\sum_{j=1}^{k+1} \beta_j = 0$. The affine independence of B forces $\beta_j = 0$ for all $j = 1, \ldots, k+1$, so we have $\gamma_i = 0$ for all $i = 1, \ldots, k$. This demonstrates the linear independence of the set $\{x_1 - x_{k+1}, \ldots, x_k - x_{k+1}\}$.

Now, if $x \in B$, then by the remark above, $x = \sum_{i=1}^{k+1} \alpha_i x_i$ and $\sum_{i=1}^{k+1} \alpha_i = 1$. We then obtain

$$x - x_{k+1} = \sum_{i=1}^{k} \alpha_i (x_i - x_{k+1})$$

As we know that the set $\{x_1 - x_{k+1}, \ldots, x_k - x_{k+1}\}$ is linearly independent, the coefficients α_i must be unique. $x = \sum_{i=1}^{k+1} \alpha_i x_i + x_{k+1}$ $\square$

The proof of Lemma 2.11 offers a test for affine dependence of points $\{x_1, \ldots, x_k\}$, one may treat them as vectors in $\mathbb{R}^d$ and write them as columns of a matrix, and append a row of 1's to the bottom as follows:

$$M = \begin{bmatrix} x_{1,1} & x_{2,1} & \ldots & x_{k,1} \\ x_{1,2} & x_{2,2} & \ldots & x_{k,2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1,d} & x_{2,d} & \ldots & x_{k,d} \\ 1 & 1 & \ldots & 1 \end{bmatrix}$$

If the null space of the M is trivial, the points are then affinely independent. Otherwise, they are affinely dependent.

Triangulations of Convex Polytopes and Point Configurations

A **point configuration** is a finite set of points A in $\mathbb{R}^d$. As noted earlier in the lectures, a convex polytope is the convex hull of finitely many points. Recall that from Lemma 2.9, we may describe a convex polytope P with vertices $p_1, \ldots, p_n$ as

$$P = \operatorname{conv}(p_1, ..., p_n) := \left\{ \sum_{i=1}^{n} \gamma_i p_i : \gamma_i \geq 0,\ \forall i = 1, \ldots, n,\ \text{and}\ \sum_{i=1}^{n} \alpha_i = 1 \right\}.$$

Recall a k-dimensional **simplex** is the convex hull of a set of $k+1$ affinely independent points.

Definition 2.13. *A* ***triangulation*** *of a point configuration A is a finite collection T of simplices σ that partitions* $\operatorname{conv}(A)$ *such that*

1. $\displaystyle\bigcup_{\sigma \in T} \sigma = \operatorname{conv}(A)$.

2. *Any pair of simplices intersects at a (possibly empty)* ***common face****.*

3. *Every vertex of T is contained in A.*

Observe that we need not use all points in the interior of $\operatorname{conv}(A)$. However, we are not allowed to add points that were absent in A into a triangulation of A. Figure 6 illustrates an example of valid triangulation that does not use all points of A, whereas Figure 7 displays examples that are not triangulations.

A common face in the definition refers to the intersection between two simplices being a face for both of the simplices. For example, on the left part of Figure 8, although an edge is a 1-dimensional face for any one of the the triangles on the right, the same edge is not a face for the triangle on the left.

On the other hand, two simplices are not allowed to intersect at the interior for both of them, as portrayed on the right part of Figure 8.

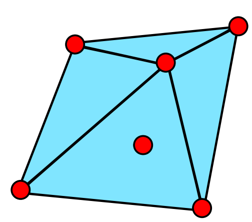


Figure 6: A permissible triangulation.

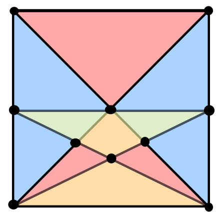


Figure 7: A subdivision that is not a triangulation.

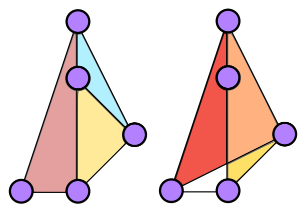


Figure 8: Forbidden cases in a would-be triangulation.

Definition 2.14. *Given a triangulation T of a polytope, the **diameter** of a simplex $\sigma \in T$ is given by*

$$diam(\sigma) = \max\{||x - y|| : x, y \in \sigma\}.$$

*The **mesh size** of the triangulation T is given by*

$$mesh(T) := \max\{diam(\sigma) : \sigma \in T\}.$$

We remark that the mesh size does not necessarily shrink if we only introduce extra simplices using vertices from the same point configuration A. This may be due to the fact that the line segment whose length equals the mesh size does not contain any point in A strictly in between. Thus, if we wish to reduce the mesh size of a triangulation T of point configuration A, then we must introduce more points to A and refine the existing triangulation.

Given a point configuration A that is finite, there may be multiple possible triangulations, but only finitely many of them. Figure 9 portrays 4 possible triangulations of the the same set of 5 points. It is not an easy task to determine the set of all triangulations, let alone determine a possible underlying structure.

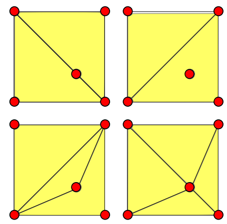


Figure 9: Four possible triangulations for the same point configuration.

Having covered the basic terminology for triangulations, we are ready to define barycentric triangulations.

Definition 2.15. *The **barycenter** of a simplex $\sigma = \mathrm{conv}(a_1, \ldots, a_{k+1})$ where $\{a_1, \ldots, a_{k+1}\}$ is affinely independent, is the point*

$$\sum_{i=1}^{k+1} \frac{1}{k+1} a_i.$$

*The **barycentric triangulation** T_b of a simplex, denoted $\sigma = \mathrm{conv}(a_1, \ldots, a_{k+1})$ has the following characterization:*

1. *The vertices of the triangulation T_b are the barycenters of faces of σ (including the vertices $a_1, \ldots, a_{k+1}$ themselves).*

2. *If $a_{i_1}, a_{i_2}, \ldots, a_{i_d}$ are the d vertices contained in a face F of σ, then associate the barycenter of F with the subset $\{i_1, i_2, \ldots, i_d\}$ of $\{1, 2, \ldots, k+1\}$. Observe that this now gives a bijection between the set of faces of σ with the subsets of $\{1, 2, \ldots, k+1\}$.*

3. *The simplices in T_b are the convex hulls of the subsets of barycenters $b_1, b_2, \ldots, b_{k+1}$ such that corresponding subsets form a full chain of containment in the indices:*

$$\{i_1\} \subset \{i_1, i_2\} \subset \ldots \subset \{i_1, i_2, \ldots, i_{k+1}\}.$$

   In other words, there cannot be further insertions of subsets within this chain. An example is shown for a tetrahedron, a 3-dimensional simplex, in Figure 10.

Observation 2.16. *There are $(k+1)!$ simplices in the barycentric triangulation of a k-dimensional simplex.*

Proof. As in the remark made in the previous definition, every simplex in the barycentric is in one-to-one correspondence with a full chain of subsets of $\{1, 2, \ldots, k+1\}$.

Now, there are $k+1$ possible choices for an element in $\{1, 2, \ldots, k+1\}$ to be included in the first subset in the chain, followed by k possible choices for a different element to be included in the second subset in the chain, and so on. As these choice can be made independently, $(k+1)!$ such chains are possible, which corresponds to the same number of simplices in a barycentric division. $\square$

Barycentric triangulations provide a useful tool at our disposal, because of these two reasons. Firstly, barycentric triangulations always exist for an arbitrary simplex and yields a possible triangulation of the simplex. Besides that, by repeated use of barycentric triangulations, as shown in Figure 11, we may successively construct a sequence of triangulations whose mesh size tend to 0.

Definition 2.17. *Denote the **first barycentric subdivision** of a simplex σ as the barycentric triangulation of σ.*

*Fix an integer $j > 1$. The **jth barycentric subdivision** of a simplex σ is the triangulation that results from the barycentric triangulation of each simplex from the $(j-1)$th barycentric subdivision of σ.*

Lemma 2.18. *The mesh size of the jth barycentric subdivision of a simplex σ approaches 0, as $j \to \infty$.*

Proof. For a given k-dimensional simplex Δ, $\operatorname{diam}(\Delta)$, the longest distance between two points in Δ is always the distance of some of its two vertices. The barycentric subdivision of Δ would introduce a barycenter G of Δ. If we take the line segment joining a vertex of Δ to the barycenter of the

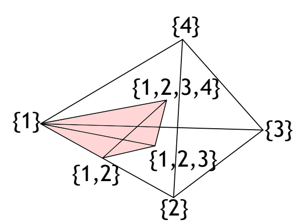


Figure 10: An example of a simplex in the barycentric triangulation of a tetrahedron.

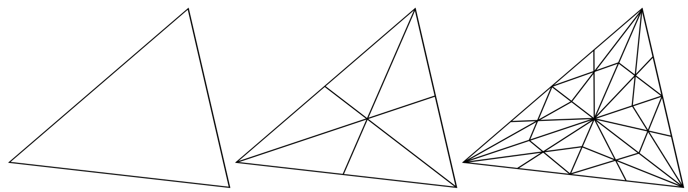


Figure 11: A sequence of barycentric subdivisions.

facet not containing this vertex, then the line segment passes through G and G divides the line segment in the ratio of $k : 1$. In addition, it can be shown inductively that this is the highest ratio possible for all vertices in the triangulation lying on a line segment joining other two vertices in the triangulation. Thus, the mesh size of this barycentric subdivision reduces by a factor of at most $\frac{k}{k+1}$.

Therefore, $\text{mesh}(T_1) \leq \frac{k}{k+1}\text{diam}(\sigma)$ and $\text{mesh}(T_j) \leq \frac{k}{k+1}\text{mesh}(T_{j-1})$, so that inductively

$$\text{mesh}(T_j) \leq \left(\frac{k}{k+1}\right)^j \text{diam}(\sigma).$$

Since $\text{diam}(\sigma)$ is constant, mesh size is always nonnegative and $\lim_{j\to\infty}\left(\frac{k}{k+1}\right)^j \text{diam}(\sigma) = 0$, we conclude that

$$\lim_{j\to\infty} \text{mesh}(T_j) = 0$$

□

Now, given a triangulation T of a k-dimensional simplex Δ and $V(T)$ denotes the set of vertices of T, the function $l : V(T) \to \{1, 2, \ldots, k+1\}$ is called a **coloring** or a **labeling function**.

2.2 Lecture 3,4: Sperner's lemma and its relatives

In this lecture we state and proof the important Sperner's Lemma. We first introduce some convenient notation.

Let Δ_d be a d - dimensional simplex with vertices $\{e_1, \cdots, e_d, e_{d+1}\}$. Furthermore, let T be a triangulation of Δ_d together with a labeling of its vertices, i.e. a function $l : V(T) \longrightarrow \{1, 2, \cdots, d+1\}$.

Definition 2.19. *Denote by F_i the facet in Δ_d opposite to e_i. A labeling l is **good** if for any vertex in T contained in F_i we have $l(v) \neq i$.*

Note that this is equivalent to the property that for any vertex v contained in a face $\text{conv}(e_s : s \in S)$ of Δ_d, we must have $l(v) \in S$. In particular, this forces $l(e_i) = i$ for all $1 \leq i \leq d+1$.

Definition 2.20. *Given a simplex Δ_d and a triangulation T with labeling l, a d-dimensional simplex $\sigma \in T$ with vertices $v_1, \cdots, v_{d+1}$ is **full** if all vertices use different labels.*

Theorem 2.21 (Sperner's lemma)**.** *Let Δ_d be a d-dimensional simplex with triangulation T, and $l : V(T) \to \{1, \cdots, d+1\}$ a good labeling on $V(T)$. Then there exists a full d-dimensional simplex $\sigma \in T$. Furthermore, the number of full simplices is odd.*

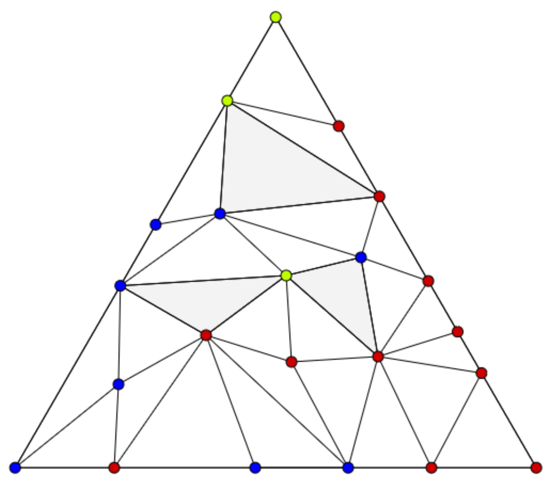


Figure 12: An example in dimension 2 of Sperner's lemma.

The proof will be by induction in the dimension d.

Proof. We begin with the base case $d = 1$. In this case we have a string of digits 0 and 1, starting with a 0 and ending with a 1. On each pair of consecutive digits the string either remains the same (00 and 11) or change (10 and 01). Since it starts with a 0 and ends with a 1, there must be an odd number of changes, which is precisely what we needed to check.

We now assume is true for any simplex Δ of dimension less than d. We are going to use the following graph G. The vertex set is the set of all $(d-1)$ simplices in T whose vertices uses all the labels in $\{1, 2, \cdots, d\}$. There is an edge between two such simplices if they are both facets of a common d simplex in T.

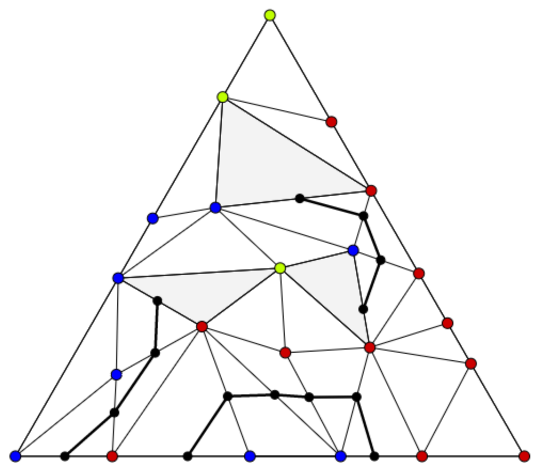


Figure 13: An example in dimension 2 of the graph.

In any dimension the only possible degrees for any vertex are 0,1, or 2:

1. Degree 0. There are two options. Either the vertex is at the facet F_{d+1} and contained in a full simplex, or is contained in two full simplices.

2. Degree 1. There are two options. Either the vertex is at the facet F_{d+1} or it is contained in (only one) full simplex.

3. Degree 2. Everything else.

Any graph with degrees at most 2 is the union of disjoint paths and cycles. We consider isolated vertices as paths of length zero. Note that the full simplices are always the endpoint of some path.

The restriction of the triangulation and the labeling to the facet/simplex F_{d+1}, is again a good labeling (see the remarks after Definition 2.19). By the inductive hypothesis, there is an **odd** number of $d-1$ simplices in T contained in F_{d+1} whose vertices uses all the labels in $\{1, 2, \cdots, d\}$. These are precisely all vertices in G lying on F_{d+1}.

Furthermore, all the vertices lying in T are all endpoints of a path (possibly of length zero). If we follow the path they begin, then we will either end up in a full simplex or back to F_{d+1}. The latter can happen just an even number of times, so it must happen an odd number of times that we end up in a full simplex. Cycles do not add any full simplices. Also a path whose two endpoints are both full simplices adds two to the number of full simplices, so it does not affect parity either, hence we are done. $\square$

Sperner's lemma's key importance is its role on the computation of fixed points of continuous maps. This is a topological challenge with applications in game theory and economics where the notion of equilibrium is very important [**?**]. Mathematically an equilibrium is a fixed point of a continuous mapping. Finding a fixed point is then an issue of practical importance. A wide variety of algorithms have been proposed and there is an extensive literature in the mathematical programming community.

One of the most famous theorems about fixed points is due to the Dutch mathematician L. E. J. Brouwer:

Theorem (Brouwer)**.** *If C is a topological d-dimensional ball and $f : C \mapsto C$ is a continuous function, then f has a fixed point, namely, there is a point x^* in C with $f(x^*) = x^*$.*

Recall that a *homeomorphism* is a one-to-one and onto continuous function whose inverse is also continuous. A topological d-ball is the image of the standard unit ball $B^d = \{x \in \mathbb{R}^d : \sum_i x_i^2 \leq 1\}$ under a homeomorphism. A simplex is our favorite example of a topological ball. Brouwer's original proof says nothing about how to find the fixed point or a good approximation to a fixed point, not even in the case when C is a simplex. In the case of a simplex Brouwer's theorem may be demonstrated via a combinatorial result about labeling triangulations due to Sperner. A non-obvious consequence of Sperner's Lemma is the famous Brouwer Fixed Point Theorem.

Theorem 2.22 (Brouwer's fixed point theorem, 1912)**.** *Let C be a nonempty, compact, convex subset of $\mathbb{R}^n$, and $f : C \to C$ a continuous function. Then there is an $x \in C$ such that $f(x) = x$, i.e. a fixed point of f.*

The next goal will be showing how the Brouwer's theorem is a consequence of Sperner's Lemma. We begin this journey with the Knaster-Kuratowski-Mazurkiewicz Lemma. This set-covering variant of Sperner's lemma is known as the KKM Lemma.

Theorem 2.23 (Knaster–Kuratowski–Mazurkiewicz (KKM) lemma, 1929)**.** *Let* Δ *be an* $(n-1)$*-dimensional simplex with vertices labeled* $1, \ldots, n$. *Let* $C_1, \ldots, C_n$ *be closed sets such that for any* $I \subseteq \{1, \ldots, n\}$,

$$\operatorname{conv}(I) \subseteq \bigcup_{i \in I} C_i, \tag{4}$$

where $\operatorname{conv}(I)$ *is the convex hull of the vertices in* I. *Then* $\bigcap_{i=1}^{n} C_i$ *is nonempty.*

Before we give a proof, let's illustrate the KKM lemma with an example. Figure 1 has a two-dimensional simplex with vertices labeled 1 through 3, and three closed sets C_1, C_2, C_3 satisfying (1). The nonempty intersection $C_1 \cap C_2 \cap C_3$ is highlighted in red.

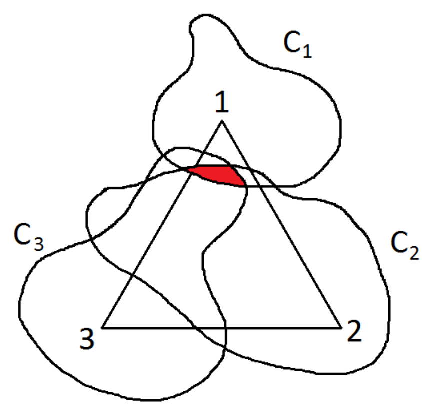


Figure 1: An illustration of the KKM lemma in two dimensions.

Our proof of KKM will use the following nice fact about simplices.

Lemma 2.24. *Let* Δ *be an* $(n-1)$*-dimensional simplex with vertices labeled* $1, \ldots, n$. *If* $I_1, I_2 \subseteq \{1, \ldots, n\}$, *then*

$$\operatorname{conv}(I_1 \cap I_2) = \operatorname{conv}(I_1) \cap \operatorname{conv}(I_2).$$

Proof. **(Lemma 3)** Let $x \in \operatorname{conv}(I_1 \cap I_2)$. Then x can be written as a convex combination of vertices belonging to $I_1 \cap I_2$ (this was a homework problem), which is therefore a convex combination of vertices in I_1 (so $x \in \operatorname{conv}(I_1)$) and a convex combination of vertices in I_2 (so $x \in \operatorname{conv}(I_2)$). Thus $x \in \operatorname{conv}(I_1) \cap \operatorname{conv}(I_2)$.

Now let $y \in \operatorname{conv}(I_1) \cap \operatorname{conv}(I_2)$. Then we can write y as two convex combinations:

$$y = \sum_{i \in I_1} \gamma_i x_i, \qquad y = \sum_{i \in I_2} \lambda_i x_i,$$

where x_i is the vertex of Δ corresponding to $i \in \{1, \ldots, n\}$, all $\gamma_i \geq 0$, all $\lambda_i \geq 0$, $\sum_{i \in I_1} \gamma_i = 1$ and $\sum_{i \in I_2} \lambda_i = 1$. Then

$$\sum_{i \in I_1} \gamma_i x_i - \sum_{i \in I_2} \lambda_i x_i = 0.$$

We re-index this into three disjoint sums, setting $\eta_i = \gamma_i - \lambda_i$:

$$\sum_{i \in I_1 \cap I_2} \eta_i x_i + \sum_{i \in I_1 \setminus I_2} \gamma_i x_i - \sum_{i \in I_2 \setminus I_1} \lambda_i x_i = 0 \tag{5}$$

There are two cases. If $\gamma_i = 0$ for all $i \in I_1 \setminus I_2$, then

$$y = \sum_{i \in I_1} \gamma_i x_i = \sum_{i \in I_1 \cap I_2} \gamma_i x_i + \sum_{i \in I_1 \setminus I_2} \gamma_i x_i = \sum_{i \in I_1 \cap I_2} \gamma_i x_i$$

and

$$1 = \sum_{i \in I_1} \gamma_i = \sum_{i \in I_1 \cap I_2} \gamma_i + \sum_{i \in I_1 \setminus I_2} \gamma_i = \sum_{i \in I_1 \cap I_2} \gamma_i,$$

and thus $y \in \operatorname{conv}(I_1 \cap I_2)$. Consider the other case, in which $\gamma_i \neq 0$ for some $i \in I_1 \setminus I_2$. Note that

$$\sum_{i \in I_1 \cap I_2} \eta_i + \sum_{i \in I_1 \setminus I_2} \gamma_i - \sum_{i \in I_2 \setminus I_1} \lambda_i = \sum_{i \in I_1} \gamma_i - \sum_{i \in I_2} \lambda_i = 1 - 1 = 0. \tag{6}$$

Disaster strikes! Now (2) and (3) give an affine dependence relation among the vertices $I_1 \cup I_2$—but Δ is by definition the convex hull of the affinely independent vertices $\{1, \ldots, n\}$, so the vertices $I_1 \cup I_2$ must be affinely independent. Contradiction! This completes the proof. $\square$

Note that Lemma 3 does not hold for a general polytope! Try it for a square. Lemma 3 is immediately extended by induction to any finite list of subsets $I_1, \ldots, I_k \subseteq I$.

We now return to proving KKM.

Proof. **(Theorem 2)** We will apply Sperner's lemma. For each integer $k \geq 0$, let T_k be the kth barycentric subdivision of Δ, where $T_0 = \Delta$. Let $V(T_k)$ be the set of all vertices of all simplices in T_k, and set $V = \bigcup_{k=0}^{\infty} V(T_k)$. For each $v \in V$, define

$$\mathcal{I}(v) = \{I \subseteq \{1, \ldots, n\} : v \in \operatorname{conv}(I)\}$$

and

$$I^*(v) = \bigcap_{I \in \mathcal{I}(v)} I.$$

Define a labeling $\ell : V \to \{1, \ldots, n\}$ by

$$\ell(v) = \min\{i \in I^*(v) : v \in C_i\}.$$

We need to prove that $\ell(v)$ always exists. Let $v \in V$. Note that $v \in \Delta = \operatorname{conv}(\{1, \ldots, n\})$, so $\mathcal{I}(v)$ is nonempty. Now we need $I^*(v)$ to be nonempty. We have

$$v \in \bigcap_{I \in \mathcal{I}(v)} \operatorname{conv}(I) =^* \operatorname{conv}\left(\bigcap_{I \in \mathcal{I}(v)} I\right) = \operatorname{conv}(I^*(v)).$$

*By Lemma 3. Because $\operatorname{conv}(I^*(v))$ is nonempty, it follows that $I^*(v)$ is nonempty. (Note that $\operatorname{carr}(v) := \operatorname{conv}(I^*(v))$ is the unique smallest dimensional face of Δ containing v, called the *carrier of* v.) Next, by the assumption (1) of $C_1, \ldots, C_n$,

$$v \in \operatorname{conv}(I^*(v)) \subseteq \bigcup_{i \in I^*(v)} C_i,$$

so there is an $i \in I^*(v)$ such that $v \in C_i$. (Note that we have now used the full power of the assumption (1).) Therefore, the set $\{i \in I^*(v) : v \in C_i\}$ is nonempty, so applying the well-ordering principle, $\ell(v)$ exists.

Now we show that for any $k \geq 0$, ℓ restricted to $V(T_k)$ is a Sperner labeling of T_k. Suppose that $v \in V(T_k)$ belongs to a face of Δ containing the vertices in $I \subseteq \{1, \ldots, n\}$; that is, $v \in \operatorname{conv}(I)$. Then by definition of ℓ, we have $\ell(v) \in I^*(v)$, and by definition of $I^*(v)$, we have $I^*(v) \subseteq I$. So $\ell(v) \in I$, which proves that ℓ is a Sperner labeling.

By Sperner's lemma, for each $k \geq 0$, there is a fully colored simplex $\sigma_k \in T_k$. We can write σ_k as the convex hull of its vertices:

$$\sigma_k = \operatorname{conv}(\{x_1^k, \ldots, x_n^k\}),$$

where, without loss of generality (by re-indexing the x_i^k's), $\ell(x_i^k) = i$ for each $i = 1, \ldots, n$. By definition of ℓ, we have $x_i^k \in C_i$ for every $k \geq 0$. Because the sequence $(x_1^k)_{k\geq 0}$ is contained in the compact set Δ, it has a convergent subsequence $(x_1^{k_j})_{j\geq 0}$ with limit $x \in \Delta$. Because $(x_1^{k_j})_{j\geq 0}$ is a sequence in C_1, and C_1 is closed, we have $x \in C_1$.

The final step is to show that every other subsequence $(x_i^{k_j})_{j\geq 0}$, $2 \leq i \leq n$ of vertices also converges to the same limit x. For any $k \geq 0$, recall that the mesh size of T_k is defined by

$$\mathrm{mesh}(T_k) = \max\{\mathrm{diam}(\sigma) : \sigma \in T_k\}$$

where for any $\sigma \in T_k$,

$$\mathrm{diam}(\sigma) = \max\{\|x - y\|_2 : x, y \in \sigma\}$$

and $\|\cdot\|_2$ is the Euclidean norm. (Note that the maximum in the mesh size is attained because T_k consists of finitely many simplices, and the maximum in the diameter is attained because $\|\cdot\|_2$ is continuous and σ is compact.) An important property of the barycentric subdivision (indeed, the reason for its existence) is that

$$\mathrm{mesh}(T_k) \leq \left(1 - \frac{1}{n}\right)^k \mathrm{mesh}(T_0)$$

for all $k \geq 0$ (see Munkres *Elements of Algebraic Topology*), where $\mathrm{mesh}(T_0) = \mathrm{diam}(\Delta)$. Let $i \in \{2, \ldots, n\}$. Then

$$\begin{aligned}
\|x_i^{k_j} - x\|_2 &\leq \|x_i^{k_j} - x_1^{k_j}\|_2 + \|x_1^{k_j} - x\|_2 \\
&\leq \mathrm{diam}(\sigma_{k_j}) + \|x_1^{k_j} - x\|_2 \\
&\leq \mathrm{mesh}(T_{k_j}) + \|x_1^{k_j} - x\|_2 \\
&\leq \left(1 - \frac{1}{n}\right)^{k_j} \mathrm{diam}(\Delta) + \|x_1^{k_j} - x\|_2 \to 0
\end{aligned}$$

as $j \to \infty$. Thus $x_i^{k_j} \to x$ as $j \to \infty$. Because $x_i^{k_j}$ is a sequence in C_i, and C_i is closed, it follows that $x \in C_i$. Therefore, $x \in C_i$ for all $i = 1, \ldots, n$, and $\bigcap_{i=1}^n C_i$ is nonempty. $\square$

We now return to Theorem 1, Brouwer's fixed point theorem. The theorem is stated for compact, convex sets, but we now show that it suffices to prove it for the standard $(n-1)$-dimensional simplex Δ^n, defined by

$$\Delta^n = \left\{ x \in \mathbb{R}^n : x \geq 0, \sum_{i=1}^n x_i = 1 \right\}.$$

Here, Δ^n is $(n-1)$-dimensional but "lives in" $\mathbb{R}^n$.

If ω is any other $(n-1)$-dimensional simplex, ω is homeomorphic to Δ^n. This is because every point of ω has unique barycentric coordinates. In other words, if $v_1, \ldots, v_n$ are the vertices of ω, then every point $p \in \omega$ can be written uniquely as a convex combination of the vertices:

$$p = \sum_{i=1}^{n} \beta_i v_i, \text{ with all } \beta_i \geq 0 \text{ and } \sum_{i=1}^{n} \beta_i = 1.$$

Then the map from ω to Δ^n given by $\sum_{i=1}^n \beta_i v_i \mapsto \sum_{i=1}^n \beta_i e_i$ is a homeomorphism, where e_i is the ith standard basis vector of $\mathbb{R}^n$. It follows that proving Brouwer for the standard simplex will prove it for any simplex, by considering the continuous function f composed with this homeomorphism.

In fact, if C is any nonempty, compact, convex subset of $\mathbb{R}^n$, then C is homeomorphic to a simplex. This is a fact that we will not prove, but which allows us to extend the following proof to its more general statement above.

Proof. **(Theorem 1)** We give a proof for the case that $C = \Delta^n$. Let $f : \Delta^n \to \Delta^n$ be continuous. We want to find $x \in \Delta^n$ so that $f(x) = x$. For each $j = 1, \ldots, n$, let

$$C_j := \{x \in \Delta^n : f(x)_j \leq x_j\}.$$

We claim that $C_1, \ldots, C_n$, satisfy the hypotheses of the KKM lemma.

1. Let $j \in \{1, \ldots, n\}$, and we show that C_j is closed. Let $(x^{(k)})_{k \geq 1}$ be any convergent sequence in C_j with limit x^*. So for every $k \geq 1$, we have $f(x^{(k)})_j \leq x_j^{(k)}$. By the continuity of f, $f(x^{(k)}) \to f(x^*)$ as $k \to \infty$. Then by convergence in the infinity norm on $\mathbb{R}^n$, we have $x_j^{(k)} \to x_j^*$ and $f(x^{(k)})_j \to f(x^*)_j$ as $k \to \infty$. Therefore,

$$f(x^*)_j = \lim_{k \to \infty} f(x^{(k)})_j \leq \lim_{k \to \infty} x_j^{(k)} = x_j^*.$$

   Thus $x^* \in C_j$, so C_j is closed.

2. To show: for any $I \subseteq \{1, \ldots, n\}$, $\operatorname{conv}(I) \subseteq \bigcup_{j \in I} C_j$.

   If $I = \emptyset$, then $\operatorname{conv}(I) = \emptyset$ and $\bigcup_{j \in I} C_j = \emptyset$, and $\emptyset \subseteq \emptyset$. So assume that $I \neq \emptyset$. Let $x \in \operatorname{conv}(I)$. Assume to the contrary that $f(x)_j > x_j$

for all $j \in I$. Then because $f(x) \in \Delta^n$,

$$1 = \sum_{j=1}^{n} f(x)_j \geq \sum_{j \in I} f(x)_j > \sum_{j \in I} x_j = 1,$$

a contradiction. So there is a $j \in I$ such that $f(x)_j \leq x_j$. Therefore, $x \in \bigcup_{j \in I} C_j$. So $\mathrm{conv}(I) \subseteq \bigcup_{j \in I} C_j$.

Since the sets $C_1, \ldots, C_n$ satisfy the hypotheses, KKM says that $\bigcap_{j=1}^{n} C_j \neq \emptyset$. So let $x \in \bigcap_{j=1}^{n} C_j$, and we claim that x is a fixed point. Because $x \in C_j$ for every j, $f(x)_j \leq x_j$ for all j. So

$$1 = \sum_{j=1}^{n} f(x)_j \leq \sum_{j=1}^{n} x_j = 1.$$

Since equality holds throughout, $f(x)_j = x_j$ for all j. Thus $f(x) = x$. $\square$

Brouwer's fixed point theorem has a lot of applications. Notably, John von Neumann used it to prove the Minimax Theorem in game theory, which is equivalent to strong duality in linear programming. John Nash proved the existence of equilibria in strategic games using Brouwer's fixed point theorem.

Another interesting result is a generalization of Sperner's lemma proved by our very own professor. Let P be a d-dimensional polytope, which is the convex hull of its n vertices $v_1, \ldots, v_n$. Let T be a triangulation of P, possibly using additional vertices. A *Sperner labeling* of T is a labeling of the vertices of T by $1, \ldots, n$ such that a vertex v of T can only be labeled by j if $v_j \in \mathrm{carr}(v)$.

2.3 lectures 4,5

Theorem 2.25 (Polytopal Sperner's Lemma (De Loera, Peterson, Su))**.** *Let P be a d-*
dimensional polytope with n vertices, and T a triangulation of P. For any Sperner labeling of T, there are at least $n - d$ d-dimensional simplices of T each with $d + 1$ different labels on its vertices ("fully colored").

For example, in the hexagon in Figure 2, we have $n = 6$ and $d = 2$. So by Theorem 4 there must be at least 4 fully colored simplices. In fact, there are 10 in this example. Notice that because we are using 6 colors, "fully colored" refers to any set of 3 distinct colors.

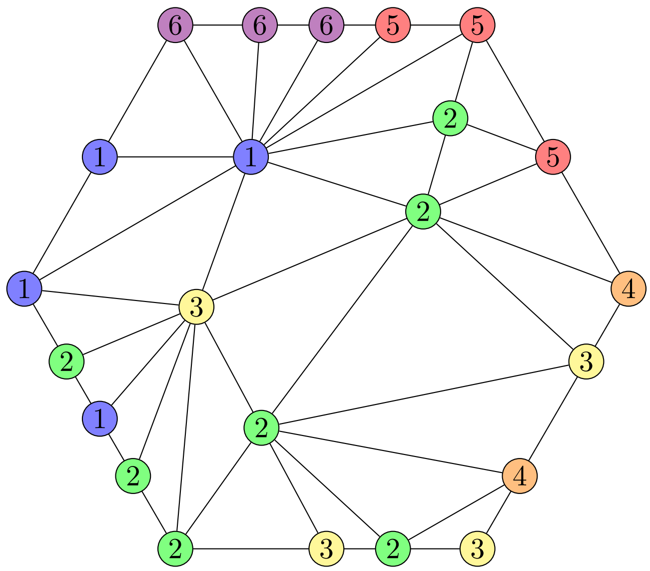


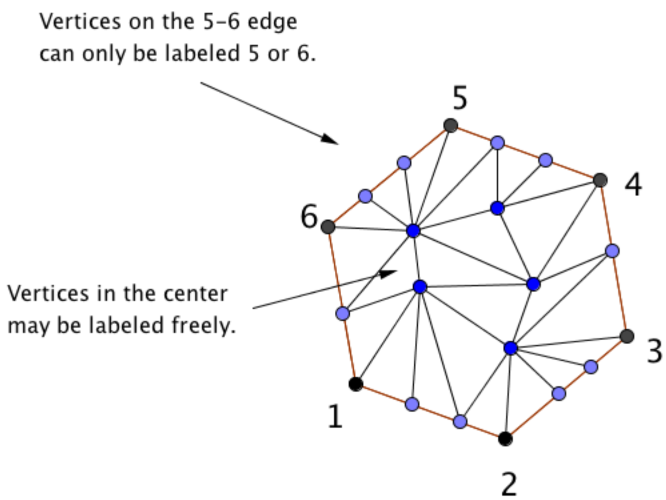
Figure 2: An illustration of the polytopal Sperner's lemma in two dimensions.

Using Brouwer's theorem we can prove this generalization of Sperner's lemma to labelings of vertices of triangulations of arbitrary convex polytopes (see Figure **??** for an example and [**?**, **?**]

Note that in particular this implies for P a d-simplex the original Sperner's lemma. Thus we have a proof that Brouwer and Sperner's lemma are in fact equivalent.

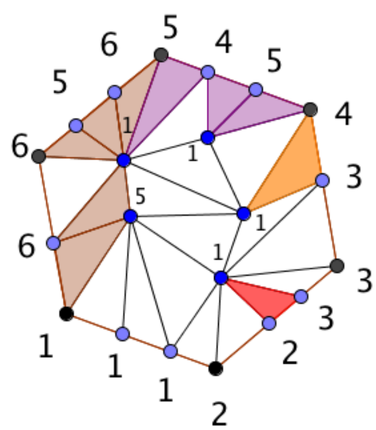
In what follows we must extend previous definitions to deal with polytopes.

Definition 2.26. *Given a polytope, P, with n vertices and a triangulation, T, let $v \in V(T)$. The* **carrier** *of v, denoted carrier(v) is the smallest face of P that contains v. Given a labeling $l : V(T) \to \{1, \cdots, n\}$ such that each original vertex in P has a distinct label, we extend the notion of a* **good** *labeling to mean that for any $v \in V(T)$, $l(v)$ can only take values l assigned to vertices in the carrier of v.*



Theorem 2.27 (Polytopal Sperner (De Loera, Peterson, Su))**.** *Let $P \subset \mathbb{R}^d$ be a convex polytope on n vertices, T be a triangulation on P and $l : V(T) \to \{1, \cdots, n\}$ a good labeling. Then there will be at least $n - d$ fully colored simplices.*

Here is a picture for the case when $n = 6$ and $d = 2$



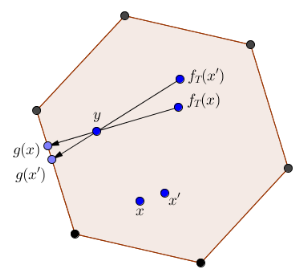
Proof. For convenience we will use $e_1, \cdots, e_n$ to denote the vertices of P and without loss of generality we assume $l(e_i) = i$. Define $f_T : P \to P$ be first defining $f_T(v) = e_{l(v)}$ for all vertices in T, then extend f_T linearly.

Claim 2.28. *This map f_T is surjective.*

Observation 2.29. *1. Since our labeling is good, if F is a face of P, we know $f_T(F) \subseteq F$.*

2. Since faces of a polytope are polytopes themselves, it suffices to show f_T is surjective on the interior of P.

Let $y \in int(P)$ and for a contradiction suppose y is not in the image of f_T. Define $g : P \to \partial P$ as follows, for $x \in P$, we look at the ray originating at $f_T(x)$ traveling in the direction of y. This ray hits the boundary of P at some point, we define $g(x)$ to be this point.



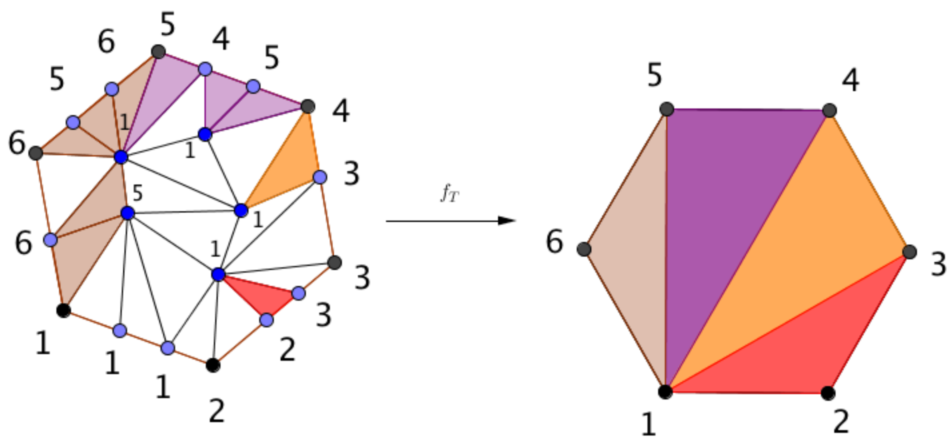
Note that because y is not in the image of f_T, g is defined on all of P. Also notice that g is continuous because f_T is continuous (as illustrated above[1]). Thus, by Brouwer's Fixed Point Theorem, there is a point $z \in P$ such that $z = g(z)$. Since g maps to the boundary, we must have that $z \in \partial P$. This of course means that z lies in some face, $F_z \subset P$. By our first observation above, this means $f_T(z) \in F_z$. Since $y \in int(P)$ the ray originating from $f_T(z)$ and going through y cannot possibly terminate on F_z. That is to say, we cannot have $g(z) = z$, a contradiction! Hence f_T is surjective on the interior of P and therefore on P.

It should be clear from how f_T is defined that the only points that map to the interior of P are those that lie in the interior of a fully colored simplex in T. For this reason.

Observation 2.30. *Let $FULL$ denote the set of fully colored simplices in T. The map f_T is still surjective when restricted to $FULL$.*

We illustrate this observation below:

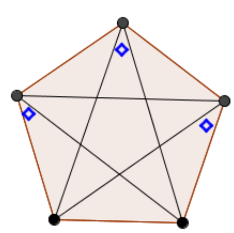
[1] If x' is close to x then $f_T(x)$ is close to $f_T(x')$ and so the rays terminate at points that are also close



So, $f_T(FULL)$ is a covering of P with simplices whose vertices come from the vertices of P. To address the issue of counting how many simplices must be used, we introduce pebble sets.

Definition 2.31. *A* pebble set *of a polytope P is a finite set of points (pebbles) in P such that any d-dimensional simplex with vertices in P contains at most one pebble.*

Here is an illustration of a pebble set for the pentagon. Note that pebble sets are not unique. Indeed, the rotational symmetry of the pentagon shows that there are at least four more pebble set, similar to the one below.

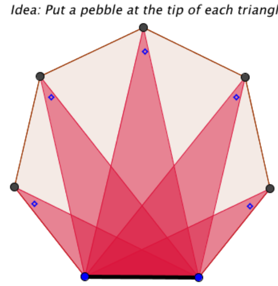


Let us call two simplices in $FULL$ distinct/different if their vertices are labeled differently. (Note this is why simplices with the same labels were shaded with the same color, while simplices with different labels were shaded with different colors in all the examples.)

Observation 2.32. *By the definition of a pebble set we know that distinct pebbles are must be contained in distinct simplices in $f_T(FULL)$. Therefore $|FULL| \geq |peb(P)|$ for all pebble sets, $peb(P)$.*

Claim 2.33. *Every polytope $P \subset \mathbb{R}^d$, with n vertices, has a pebble set of size at least $n - d$.*

The proof of this claim is rather messy and omitted, but the idea for two dimensions is demonstrated below:



This completes the proof. $\square$

Lemma 2.34 (Lebesgue's Lemma)**.** *Let Δ be a $(d+1)$ dimensional simplex cover of Δ by closed sets $\{M_1, M_2, \ldots, M_{d+2}\}$ such that $\Delta = conv(v_1, \ldots, v_{d+2})$. If M_i contains the facet $conv(v_1, \ldots, \hat{v}_i, \ldots, v_{d+2})$, then $\bigcap M_i \neq \emptyset$.*

Proof. [**Exercise 1 for Homework**] Use Sperner's Lemma by coloring (Hint: Use indices that are not containing a point) $\square$

Observation 2.35. *With Exercise* 1 *(Hint: Using linear map from Exercise* 1*) we can prove the following:*

[**Exercise 2 for Homework**] Let $\{a_1, a_2, \ldots, a_{d+2}\}$ be points in $\mathbb{R}^d$. $\Omega = conv(a_1, a_2, \ldots, a_{d+2})$. $K_1, K_2, \ldots, K_{d+2}$ are closed sets covering Ω such that $K_i \supseteq conv(a_i, \ldots, \hat{a}_i, \ldots, a_{d+2})$, then $\bigcap_{i=1}^{d+2} K_i \neq \emptyset$.

Theorem 2.36 (Helly's Theorem (Special Case))**.** *Let $K_1, K_2, \ldots, K_{d+2}$ be convex sets in $\mathbb{R}^d$. Suppose any subfamily of $d+1$ many intersects $\bigcap_{i \neq j} K_i \neq \emptyset$ for all $j, j = 1, \ldots, d+2$,*
then $\bigcap_{i=1}^{d+2} K_i \neq \emptyset$.

Proof. Let $a_i \in \bigcap_{j \neq i} K_j$, and $\Omega = conv(a_1, a_2, \ldots, a_{d+2})$.
Claim: $K_i \bigcap \Omega \supseteq conv(a_1, a_2, \ldots, \hat{a}_i, \ldots, a_{d+2})$. (Note that this satisfies Exercise 2). This implies that $\bigcap_{i \neq j} K_i \neq \emptyset$. $\square$

O. Musin generalized of Sperner and Tucker lemmas to large classes of manifolds with or without boundary;

One more fascinating consequence is S. Kakutani's 1941 theorem

Theorem (Kakutani's). *Let X be a compact, convex subset of Euclidean d-space. Let F be a continuous set-valued function on X; i.e., a mapping from X to the set of all subsets of X, If $T(x)$ is convex for all x belonging to X, then there exists a vector z such that $z \in T(z)$.*

To prove this theorem we will use Brouwer's fixed point theorem in combination of the barycentric triangulation we have used before.

Anecdote: In his game theory textbook, Ken Binmore recalls that Kakutani once asked him at a conference why so many economists had attended his talk. When Binmore told him that it was probably because of the Kakutani fixed point theorem, Kakutani was puzzled and replied, "What is the Kakutani fixed point theorem?"

First we need to generalize the classical concept of point-valued functions.

Definition 2.37. *A* set-valued *function is a function $F : X \longrightarrow 2^X$. In words, it is a function that send elements of a set X to subsets of X.*

We are interested in the membership problem, i.e. we want to find $x \in X$ such that $x \in F(X)$. This generalizes the notion of fixed point for the classical point-valued functions.

In the same way that most of the time we restrict our attention to *continuous* functions, we need to ask for a similar condition on set-valued functions.

Proposition 2.38. *A function $f : X \longrightarrow Y$ between two topological spaces is continuous if and only if its graph $\Gamma_f = \{(x, y) \in X \times Y : \quad y = f(x)\}$ is closed in $X \times Y$.*

Analogously we can define the graph for any set valued function F as

$$\Gamma_F = \{(x, x') \in X \times X : \quad x' \in F(x)\}$$

If additionally X is a topological space, it makes sense to talk about its closedness. More specifically we say that Γ_F is closed if given a sequence $\{(x_k, x'_k)\}$ with $x'_k \in F(x_k)$ and such that $x = \lim x_k$ and $x' = \lim x'_k$ exist, then $x' \in F(x)$.

Theorem 2.39. *Let $X \subset \mathbb{R}^n$ be a compact convex set. If $F : X \longrightarrow 2^X$ is a set valued function with closed graph and the property that $F(x)$ is nonempty and convex for all $x \in X$, then there exist a $x^* \in X$ such that $x^* \in F(x^*)$.*

It is enough to prove it for the standard simplices and that's what we'll do.

Proof. Let Δ the $d-$dimensional standard simplex. Consider the T_n the $n-$th barycentric subdivision. We construct a continuous function f_n as follows. For all vertices $v \in T_n$ we define $f_n(v)$ as an arbitrary point in $F(v)$. Having defined it on the vertices of T_n we extend linearly to all the simplices inside the barycentric subdivision. The resulting map f_n is continuous since it is piecewise linear.

Brouwer's theorem guarantess that for each n there exists a fixed point x^n. If any of the x^n is a vertex of its corresponding barycentric subdivision, then we are done since $x^n = f(x^n) \in F(x^n)$ by definition. From now on we assume that x^n is not a vertex of T_n, then it is in some simplex $\mathrm{conv}(x^{n_0}, \cdots, x^{n_d})$, hence we have a convex expression

$$x^n = \theta^{n_0} x^{n_0} + \cdots + \theta^{n_d} x^{n_d} \tag{7}$$

Since the function f is linear on simplices we can apply it to Equation 7 to get

$$x^n = \theta^{n_0} y^{n_0} + \cdots + \theta^{n_d} y^{n_d} \tag{8}$$

where $y^{n_i} = f(x^{n_i})$. Now we have $2d + 3$ sequences

$$\{x^n\}, \qquad \{\theta^{n_i}\} \quad \forall i, \qquad \{y^{n_j}\} \quad \forall j.$$

By repeatedly taking convergent subsequents, and relabeling, we can assume that each sequence convergences while Equation 8 remains true for all n.

Observation 2.40. *As n goes to infinity, the mesh in T_n goes to zero, so $x^* = \lim\limits_{n\to\infty} x^n = \lim\limits_{n\to\infty} x^{n_i}$ for all i.*

Taking limits on Equation 8 we get

$$x^* = \theta_0^* y_0^* + \cdots + \theta_d^* y_d^* \tag{9}$$

where $\lim\limits_{n\to\infty} \theta^{n_i} = \theta_i^*$. Here is where we use the closedness of the graph. Since $y^{n_j} = f_n(x^{n_j}) \in F(x^{n_j})$, in the limit we have

$$y_j^* = \lim_{n\to\infty} y^{n_j} \in F\left(\lim_{n\to\infty} x^{n_j}\right) = F(x^*)$$

This means that the right hand side of Equation 9 is a convex combination of points in $F(x^*)$. Since $F(x^*)$ is convex, this proves that $x^* \in F(x^*)$ as we wanted. $\square$

The following examples shows that we cannot drop some of the conditions in the statement.

Boundedness of X

Let $X = \mathbb{R}$ and consider the function $F : \mathbb{R} \to 2^{\mathbb{R}}$:

$$F(x) = \begin{cases} \{\frac{-1}{x}\} & \text{for } x < 0 \\ \{2\} & \text{for } x = 0 \\ \{\frac{1}{x}\} & \text{for } x > 0 \end{cases}$$

There is no fixed point. Kakutani's theorem doesn't apply since $\mathbb{R}$ is not compact.

Convexity of $F(x)$

Consider $X = [-1, 1] \subset \mathbb{R}$. Which is compact and convex. We define $F : X \longrightarrow 2^X$ as

$$F(x) = \begin{cases} \{\frac{1}{2}\} & \text{for } -1 \leq x < 0 \\ \{-\frac{1}{2}, \frac{1}{2}\} & \text{for } x = 0 \\ \{\frac{-1}{2}\} & \text{for } -1 \leq x > 0 \end{cases}$$

The graph is closed, but the conclusion fails. What fails is that $F(0)$ is not convex.

3 Nash Equilibria

A nice application of Kakutani's fixed point theorem appears in the proof of Nash Equilibrium (for finite games). We define games in the extensive form. Each game $\mathcal{G}$ has three main components: The set of players (P), sets of strategies for each player (S^p for $p \in P$) and the set of payoffs for each player given the strategies played (u^p). We'll star with a simple example with 3 players (results can be extended to any finite number of players and strategies). Let $\mathcal{G}$ be a game with 3 players: Abe (A), Bernie (B) and Charlie (C). The set of players is defined as $P = \{A, B, C\}$. In this game, each player has a finite set of pure-strategies defined, i.e., for Abe we define his strategy set as $S^A = \{1, 2, \ldots, i_0\}$, for Bernie $S^B = \{1, 2, \ldots, j_0\}$ and for Charlie $S^C = \{1, 2, \ldots, k_0\}$. Elements for each strategy set will be denoted by $s^p \in S^p$ with $p \in P$. A strategy profile is is a vector $s \in S$ where $S = \times_{p \in P} S^p$. Therefore, for each strategy profile $s \in S$, agents have payoffs given by $u^p(s) \in \mathbb{R}$. Thus, the game can summarized as follows $\mathcal{G} = \big(P, S^p, u^p(s),\big)_{p\in P}$.

To simplify notation, payoffs will be written as follows: For an strategy profile $s = (i, j, k) \in S$ (Abe plays $i \in S^A$, Bernie plays $j \in S^B$ and Charlie plays $k \in S^C$), payoffs will be a_{ijk}, b_{ijk} and c_{ijk}, for Abe, Bernie and Charlie, respectively.

We allow agents to play a given pure strategy $s^p \in S^P$ or to randomize of their set of strategies. For example, Abe is allowed to play $i \in S^A$ or to play $i \in S^A$ with probability p_i and all others with probability $1-p_i$ for $p_i \in [0, 1]$. Over the set of pure strategies, we define mixed strategies σ^p as a probability distribution over pure strategies. The assumption behind randomization is that mixed strategies for each player's are statistically independent of those of his opponents. Let the space of player's p mixed strategies be Σ^p, where $\sigma^p(s^p)$ is the probability that σ^p that player p assigns to s^p. The space of mixed strategy profiles is denoted $\Sigma = \times_{p\in P}\Sigma^p$, with element σ. Hence, if Abe plays a mixed strategy $\sigma^A = (p_1, \dots, p_{i_0})$, Bernie plays $\sigma^B = (q_1, \dots, q_{j_0})$ and Charlie plays $\sigma^C = (r_1, \dots, r_{i_0})$, the expected payoff to profile $\sigma = (\sigma^A, \sigma^B, \sigma^C)$ is given by

$$U^A(\sigma) = \sum_{i,j,k}(p_i q_j r_k)a_{ijk}, \quad U^B(\sigma) = \sum_{i,j,k}(p_i q_j r_k)b_{ijk}, \quad U^C(\sigma) = \sum_{i,j,k}(p_i q_j r_k)c_{ijk},$$

for Abe, Bernie and Charlie, respectively.

Since each mixed strategy σ^p is a probability distribution over pure strategies, we must have:

$$p_i \geq 0, \quad \sum_{i=1}^{i_0} p_i = 1; \qquad q_j \geq 0, \quad \sum_{j=1}^{j_0} q_j = 1; \qquad r_k \geq 0, \quad \sum_{k=1}^{k_0} r_k = 1.$$

To simplify notation, we will write the expected payoff for a mixed strategy $\sigma = (p, q, r)$ as $a(p, q, r)$, $b(p, q, r)$ and $c(p, q, r)$.

In this 3-person game, we can ask the following question: Which strategy (pure or mixed) each agent has to take given the rules of the game? Clearly we are missing one important assumption to answer this question: *selfishness*. In economics, is common to assume that agents are *selfish*, i.e., agents want to maximize their (expected) payoff for any given strategy of their opponents independent of the payoff of their opponents. For instance, when thinking about the optimal decision for Abe, for any mixed strategy $q \in \Sigma^B$ that Bernie plays and any mixed-strategy $r \in \Sigma^C$ that Charlie plays, Abe will choose a mixed strategy $\hat{p} \in \Sigma^A$ such that

$$a(\hat{p}, q, r) \geq a(p, q, r), \qquad \forall p \in \Sigma^A.$$

I.e., $\hat{p}$ is the mixed strategy that allows Abe to get the highest expected payoff. Clearly, for each strategy (pure or mixed) Abe can have different strategies that maximize his (expected) payoff.

In this scenario, we can *informally* define a Nash equilibrium as the profile of strategies such that each player's strategy is an optimal response to the other player's strategies. Formally:

Definition 3.1. *A mixed strategy profile* $(\hat{p}, \hat{q}, \hat{r}) \in \Sigma$ *is a Nash equilibrium if, for all players (Abe, Bernie and Charlie),*

$$a(\hat{p}, \hat{q}, \hat{r}) \geq a(p, \hat{q}, \hat{r}), \qquad \forall p \in \Sigma^A;$$

$$b(\hat{p}, \hat{q}, \hat{r}) \geq b(\hat{p}, q, \hat{r}), \qquad \forall q \in \Sigma^B;$$

$$c(\hat{p}, \hat{q}, \hat{r}) \geq c(\hat{p}, \hat{q}, r), \qquad \forall r \in \Sigma^C.$$

Finally, after all of this necessary notation, we are able to state the following theorem:

Theorem 3.2 (Nash (1950))**.** *A Nash equilibrium always exists on finite games.*

Before showing the proof, we present a classical example of Nash equilibrium in mixed strategies. Suppose two players, A and B, have one 1 coin each. In this game, each player has to decide if they show Tails (T) or Heads (H). Both players chose which side of the coin they will show and after the decision is made, they show simultaneously their decision to each other. If coins match, player A wins 1 dollar and player B looses one dollar and if they don't, player A looses one dollar and player B wins 1 dollar. The matrix of payoffs is shown in figure 14.

	H	T
H	1,-1	-1,1
T	-1,1	1,-1

Figure 14: matrix of payoffs

Since each agent is selfish, we look for strategies that maximize individual payoffs. This game doesn't have a pure strategy Nash equilibrium since agent A wins when coins match and player B wins when they don't. Then, from the Nash equilibrium theorem, since the game has at least one equilibrium, this equilibrium has to be in mixed strategies. Let p_B be the probability that agent B assigns to strategy H. If player A chooses H, the expected payoff is $U^A(H) = p_B(1)+(1-p_B)(-1) = 2p_B - 1$ and if A chooses T his payoff is $U^A(T) = p_B(-1) + (1 - p_B)(1) = 1 - 2p_B$. Then, if we call p_A the probability that player A assigns to playing H, the optimal decision will be

$$p_A(p_B) = \begin{cases} 1 & \text{if} \quad p_B > 1/2 \\ [0,1] & \text{if} \quad p_B = 1/2 \\ 0 & \text{if} \quad p_B < 1/2 \end{cases}$$

clearly p_A is a set function and by symmetry of payoffs, optimal probabilities for agent B given that player A plays H with probability p_A are given by

$$p_B(p_A) = \begin{cases} 0 & \text{if} \quad p_A > 1/2 \\ [0,1] & \text{if} \quad p_A = 1/2 \\ 1 & \text{if} \quad p_A < 1/2 \end{cases}$$

Hence, Nash equilibrium in the matching pennies game is given by $\hat{p}_A = 1/2$ and $\hat{p}_B = 1/2$.

From last episode: We presented a game with three players A, B, C, and denoted their payoffs by $A \to a_{ijk}$, $B \to b_{ijk}$ and $C \to c_{ijk}$. Therefore, expected payoffs for each agent where written as:

$$A: \qquad a(p,q,r) = \sum_{i,j,k} a_{ijk} p_i q_j r_k,$$

$$B: \qquad b(p,q,r) = \sum_{i,j,k} b_{ijk} p_i q_j r_k,$$

$$C: \qquad c(p,q,r) = \sum_{i,j,k} b_{ijk} p_i q_j r_k.$$

Finally, we gave the definition of Nash equilibrium for this game:

Definition 3.3. *A Nash equilibrium is a triple of probability vectors* $(\hat{p}, \hat{q}, \hat{r})$ *such that:*

$$a(\hat{p},\hat{q},\hat{r}) \geq a(p,\hat{q},\hat{r}), \qquad \forall p \quad \textit{probability vector},$$

$$b(\hat{p},\hat{q},\hat{r}) \geq c(\hat{p},q,\hat{r}), \qquad \forall q \quad \textit{probability vector},$$

$$c(\hat{p},\hat{q},\hat{r}) \geq c(\hat{p},\hat{q},r), \qquad \forall r \quad \textit{probability vector},$$

the above inequalities are called Nash inequalities.

Commercial break: Simplices are useful! Probability vectors for each player are just points inside the standard simplex in the dimension defined by the number of pure strategies for that player. For instance, for player A we have

$$\triangle_{i_0-1} = \{(p_1, p_2, ..., p_{i_0}) | \sum_i p_i = 1, p_i \geq 0\}.$$

Note: The space $\triangle$ of mixed strategies (p, q, r) will be denoted by

$$\triangle = \triangle_{i_0-1} \times \triangle_{j_0-1} \times \triangle_{k_0-1}.$$

Observation 3.4. *If 3 players have two strategies (such as choosing heads or tails) then the space of mixed strategies (p, q, r) is the 3-D cube. See Figure 15.*

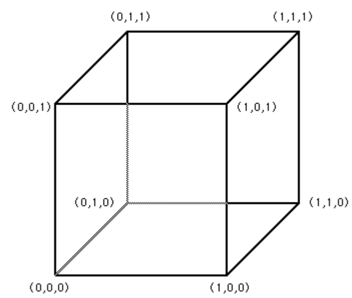


Figure 15: Space of strategies Δ when each player has 2 strategies

An important observation is that the Nash inequalities are linear. Thus, the problem of satisfying each inequality is a linear program (LP). For example, for agent A Nash inequality:

$$\sum_{i,j,k} a_{ijk}\hat{p}_i\hat{q}_j\hat{r}_k \geq \sum_{i,j,k} a_{ijk}\hat{q}_j\hat{r}_k p_i \text{ for all } p,$$

implies solving:

$$\max \sum_{i,j,k} a_{ijk}\hat{q}_j\hat{r}_k p_i \text{ s.t. } p \in \triangle_{i_0-1}.$$

where clearly the objective function is linear in p.
Example: this simple example can help you meditate:

$$\max 2p_1 + 4p_2 + 7p_3 + 7p_4 + 7p_5, \text{ s.t. } \quad p \in \triangle_4.$$

Any probability vector that puts a positive probability to strategies 1 and 2 is not optimal because you can increase the objective function giving more probability to strategies 3, 4 and 5. Hence, at the optimum we must have $\hat{p}_1 = \hat{p}_2 = 0$.

The maximum on the first Nash inequality occurs on the face of $\triangle_{i_0-1}$ defined by the vertices selected by the largest coefficients $a_{ijk}\hat{q}_j\hat{r}_k$:

$$P(\hat{q},\hat{r}) := \left\{ p : p_i \geq 0, \sum_i p_i = 1, p_i = 0 \text{ when } a_{ijk}\hat{q}_j\hat{r}_k < \max\left\{a_{ijk}\hat{q}_j\hat{r}_k\right\} \right\},$$

which is a face of $\triangle_{i_0-1}$.

Similarly, on the second Nash inequality, again the maximum exists only within the face

$$Q(\hat{p},\hat{r}) := \left\{ q : q_j \geq 0, \sum_j q_j = 1, q_j = 0 \text{ when } b_{ijk}\hat{p}_i\hat{r}_k < \max\left\{b_{ijk}\hat{p}_i\hat{r}_k\right\} \right\},$$

which is a face of $\triangle_{j_0-1}$.

Finally, for the third Nash inequality, the maximum exists within the face

$$R(\hat{p},\hat{q}) := \left\{ r : r_k \geq 0, \sum_k r_k = 1, r_k = 0 \text{ when } c_{ijk}\hat{p}_i\hat{q}_j < \max\left\{c_{ijk}\hat{p}_i\hat{q}_j\right\} \right\},$$

which is a face of $\triangle_{k_0-1}$.

Now, as Nash realized, equilibria exist if and only if there exists a triple $(\hat{p},\hat{q},\hat{r})$ with $\hat{p} \in P(\hat{q},\hat{r})$, $\hat{q} \in Q(\hat{p},\hat{r})$, and $\hat{r} \in R(\hat{p},\hat{q})$.

Theorem 3.5. *(1950 Nash) A Nash equilibrium of mixed strategies $(\hat{p},\hat{q},\hat{r})$ exists.*

Proof. We will use Kakutani's theorem to prove this. Define $X := \triangle_{i_0-1} \times \triangle_{j_0-1} \times \triangle_{k_0-1}$ to be the space of mixed strategies. Define $F : X \to 2^X$ as the set-valued function that maps

$$\begin{pmatrix} \tilde{p} \\ \tilde{q} \\ \tilde{r} \end{pmatrix} \mapsto \begin{pmatrix} P(\tilde{q},\tilde{r}) \\ Q(\tilde{p},\tilde{r}) \\ R(\tilde{p},\tilde{q}) \end{pmatrix}.$$

Now, clearly, since X is the product of simplices, it is a compact, convex set. Additionally, $F(x)$ is convex since it is the product of faces of simplices. Now, we need only show that F satisfies the last hypothesis of Kakutani's theorem, that the graph of F is closed, i.e., $\{(x, y) : y \in F(x)\}$ is a closed set.

Take a convergent sequence $\{x_n\} \subset X$. Say that $\lim_{n\to\infty} x_n =: x_0$. Take another sequence $\{y_n\}$ with $y_n \in F(x_n)$ and say that $\lim_{n\to\infty} y_n = y_0$. Our goal is to show that $y_0 \in F(x_0)$.

Now, each x_n and y_n is a triple of mixed strategies, call them

$$x_n = \begin{pmatrix} p_n \\ q_n \\ r_n \end{pmatrix}, y_n = \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}.$$

By definition, $u_n \in P(q_n, r_n), v_n \in Q(p_n, r_n)$ and $w_n \in R(p_n, q_n)$ and so the three Nash inequalities are satisfied:

$$\begin{aligned} a(u_n, q_n, r_n) &\geq a(p', q_n, r_n), \quad \forall p' \in \Delta_{i_0-1}, \\ b(p_n, v_n, r_n) &\geq b(p_n, q', r_n), \quad \forall q' \in \Delta_{j_0-1}, \\ c(p_n, q_n, w_n) &\geq c(p_n, q_n, r'), \quad \forall r' \in \Delta_{k_0-1}, \end{aligned}$$

Note that as n goes to infinity the vector (p_n, q_n, r_n) converges to x_0 and (u_n, v_n, w_n) converges to y_0. Since the inequalities are linear and thus continuous, they are satisfied in the limit and so they hold true at x_0, y_0 so we have that $y_0 \in F(x_0)$.

Thus, Kakutani's theorem gives us that there exists some $x^* \in X$ with $x^* \in F(x^*)$. Thus, this point provides a Nash equilibrium

$$x^* = \begin{pmatrix} p^* \\ q^* \\ r^* \end{pmatrix} \text{ with } p^* \in P(q^*, r^*),\ q^* \in Q(p^*, r^*) \text{ and } r^* \in R(p^*, q^*).$$

□

We move to the final application of Sperner's lemma. We will provide proof of the 1982 theorem of Imre Bárány the colorful Carathéodory's theorem. Recall the usual form first:

Let $A \in \mathbb{Z}^{d\times n}$ be an integer matrix.

Theorem 3.6. *(Carathéodory) If $Ax = b, x \geq 0$ has a solution then there exists a solution with no more than d non-zero entries in x.*

The nonnegative entries in x are called the support of x, i.e., $supp(x) = \{i : x_i \neq 0\}$.

Theorem 3.7. *(Colorful Carathéodory) Suppose $B_1, B_2, ..., B_d$ are d pairwise disjoint subsets of indices of columns, each with d-columns, of the matrix A. If the d systems $A_{B_i}x = b, x \geq 0$ all have a solution for all $i = 1, \dots, d$ then there exists a set of indices B such that*

$$A_B x = b, x \geq 0$$

has solution too and, in addition, $|B \cap B_i| \leq 1$ for all i.

One can prove this theorem using Sperner's theorem. It has some remarkable consequences, e.g., Tverberg's theorem can be proved from it.

3.1 Lecture Borsuk-Ulam and Tucker's Lemma

In this lecture, we introduce the Borsuk–Ulam theorem and some of its applications. The Borsuk–Ulam theorem comes in many different forms, and in this lecture we will prove that its different statements are indeed equivalent. The proof of the theorem itself will come later using combinatorial geometry again.

Let $\|\cdot\|_2$ be the Euclidean norm on a Euclidean space. Let $\S^n$ be the n-dimensional sphere, that is,

$$\S^n = \left\{x \in \mathbb{R}^{n+1} : \|x\|_2 = 1\right\}.$$

A function $f : \S^n \to \mathbb{R}^n$ is **antipodal** if $f(-x) = -f(x)$ for all $x \in \S^n$. That is, f is an odd function.

Theorem 3.8 (Borsuk–Ulam)**.**

(1) If a function $f : \S^n \to \S^m$ is continuous and antipodal, then $m \geq n$.

(2) If a function $f : \S^n \to \mathbb{R}^n$ is continuous and antipodal, there is an $x \in \S^n$ with $f(x) = 0$.

(3) If a function $f : \S^n \to \mathbb{R}^n$ is continuous, there is an $x \in \S^n$ with $f(x) = f(-x)$.

(4) If $\S^n = \bigcup_{i=0}^n S_i$ and each S_i is closed, there is an $x \in \S^n$ and an $i \in \{0, 1, \dots, n\}$ with $\{x, -x\} \subseteq S_i$.

Proof. We postpone the proof until later. □

Claim 3.9. *The statements (1)–(4) in Theorem 3.8 are equivalent.*

Proof. We prove the implications (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2): For a contradiction, assume there is a function $f : \S^n \to \mathbb{R}^n$ that is continuous and antipodal, but $f(x) \neq 0$ for all $x \in \S^n$. Then we can consider the function $F : \S^n \to \S^{n-1}$ defined by

$$F(x) = \frac{f(x)}{\|f(x)\|_2}.$$

For any $x \in \S^n$,

$$F(-x) = \frac{f(-x)}{\|f(-x)\|_2} = \frac{-f(x)}{\|-f(x)\|_2} = -F(x),$$

so F is antipodal. Because f is continuous, F is also continuous. This produces a contradiction to (1).

(2) $\Rightarrow$ (3): Let $f : \S^n \to \mathbb{R}^n$ be a continuous function. Consider $F : \S^n \to \mathbb{R}^n$ defined by $F(x) = f(x) - f(-x)$. Then $F(-x) = -F(x)$, so F is antipodal, and F is continuous. By (2), there is an $x^* \in \S^n$ such that $F(x^*) = 0$, i.e. $f(x^*) = f(-x^*)$.

(3) $\Rightarrow$ (4): Assume that $\S^n = \bigcup_{i=0}^n S_i$ and each S_i is closed. Define $f : \S^n \to \mathbb{R}^n$ by

$$f(x) = \big(d(x, S_1), \dots, d(x, S_n)\big),$$

where $d(x, S_i)$ is the Euclidean distance of x from S_i, which because S_i is compact is given by

$$d(x, S_i) = \min\{\|x - y\|_2 : y \in S_i\}.$$

Because f is continuous, (3) gives an $x^* \in \S^n$ with $f(x^*) = f(-x^*)$.

Case (a): There is some coordinate $f(x^*)_i = 0 = f(-x^*)_i$. Then $x^* \in S_i$ and $-x^* \in S_i$.

Case (b): $f(x^*)_i > 0$ for every $i = 1, \dots, n$. So $x^* \notin S_i$ for every $i = 1, \dots, d$. But $\S^n = \bigcup_{i=0}^n S_i$, so $x^* \in S_0$, and similarly for $-x^*$.

(4) $\Rightarrow$ (1): For a contradiction, assume that $f : \S^n \to \S^m$ is continuous and antipodal, but $m < n$. Let $\varepsilon > 0$ and build a cover $\S^n = \bigcup_{i=0}^n S_i$ by defining

$$R_i = \{(x_0, \dots, x_m) \in \S^m : x_i \geq \varepsilon\}$$

for $i = 0, 1, \dots, m$, and

$$R_{m+1} = \left\{(x_0, \dots, x_m) \in \S^m : \sum_{i=0}^m x_i \leq -\varepsilon\right\}.$$

For sufficiently small $\varepsilon > 0$, we have $\S^m = \bigcup_{i=0}^{m+1} R_i$. Moreover, no R_i contains a pair $\{x, -x\}$. Set $S_i = f^{-1}(R_i)$, so each S_i is closed because f is continuous. Then $\S^n = \bigcup_{i=0}^{m+1} S_i$. There are $m+2$ of the sets R_i, and since $n+1 \geq m+2$, there are not too many S_i's. (If there are too few S_i's to apply (4), we can just add empty sets to the list.) By (4), there is an $x \in \S^n$ and an $i \in \{0, 1, \ldots, n\}$ with $\{x, -x\} \subseteq S_i$. Because f is antipodal, we have $R_i \supseteq \{f(x), f(-x)\} = \{f(x), -f(x)\}$, a contradiction. $\square$

The Borsuk–Ulam theorem has a whole host of applications, of which we will share two. The first is a memorably-named theorem in measure theory.

Theorem 3.10 (Ham Sandwich theorem)**.** *If $X_1, \ldots, X_n$ are compact subsets of $\mathbb{R}^n$, there is a hyperplane*

$$h(u) = \{v \in \mathbb{R}^n : u_1 v_1 + \cdots + u_n v_n = u_0\},$$

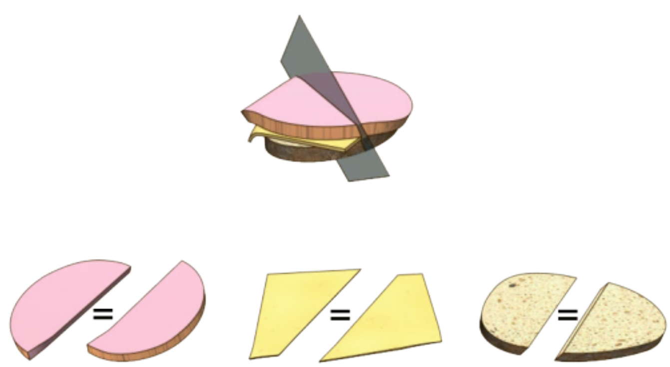
where $u = (u_0, u_1, \ldots, u_n) \in \S^n$, with corresponding half-spaces

$$h_+(u) = \{v \in \mathbb{R}^n : u_1 v_1 + \cdots + u_n v_n \geq u_0\}$$

$$h_-(u) = \{v \in \mathbb{R}^n : u_1 v_1 + \cdots + u_n v_n \leq u_0\}$$

such that $\mathrm{vol}(X_i \cap h_+(u)) = \mathrm{vol}(X_i \cap h_-(u))$ for every $i = 1, \ldots, n$.

The Ham Sandwich theorem gets its name from the case $n = 3$ in which the three sets X_1, X_2, X_3 are thought of as three layers of a ham sandwich. Then the theorem says that we can cut all three layers of the sandwich perfectly in half (by volume) with a single cut, i.e. a hyperplane!



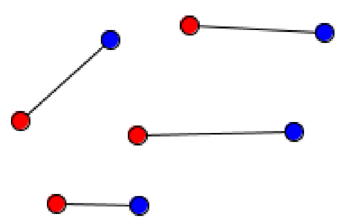
A practical interpretation of the Ham Sandwich theorem.

Proof. We will apply statement (2) of Theorem 3.8. Define $f : \S^n \to \mathbb{R}^n$ by

$$f(u) = \Big(\mathrm{vol}\big(X_1 \cap h_+(u)\big) - \mathrm{vol}\big(X_1 \cap h_-(u)\big), \ldots, \mathrm{vol}\big(X_n \cap h_+(u)\big) - \mathrm{vol}\big(X_n \cap h_-(u)\big)\Big).$$

Note that $h_+(-u) = h_-(u)$ for any $u \in \S^n$. Thus $f(-u) = -f(u)$, so f is antipodal. It can be shown that f is also continuous. Then applying statement (2) of Theorem 3.8, there is a $u^* \in \S^n$ such that $f(u^*) = 0$. Then $h(u^*)$ is the desired hyperplane. $\square$

Theorem 3.11 (Akiyama, Alon)**.** *Suppose that $A_1, \ldots, A_d$ are general-position finite sets in $\mathbb{R}^d$ with $|A_i| = m \geq 1$ for all $i = 1, \ldots, d$. Then there is a partition of $\bigcup_{i=1}^d A_i$ into m d-element sets $B_1, \ldots, B_m$ such that $|A_i \cap B_j| = 1$ for every $i = 1, \ldots, d$ and $j = 1, \ldots, m$, and $B_1, \ldots, B_m$ have pairwise disjoint convex hulls.*



An illustration of Theorem 3.11 in the case $d = 2$ and $m = 4$.
The red points belong to A_1, and the blue points to A_2.

Proof. (Sketch of idea only) Apply induction on m. If m is odd, then by the Ham Sandwich theorem, there is a hyperplane intersecting each A_i in one point with $(m-1)/2$ of each color on each side. Now take one B_j to be the intersection of the hyperplane with $\bigcup_{i=1}^d A_i$, and continue with $A_i \cap h_+$ and $A_i \cap h_-$. $\square$

The following is another version of Borsuk-Ulam theorem (we leave it as an exercise):

Theorem 3.12. *There is no continuous map $f : B^n \to S^{n-1} = \partial B^n$ that is antipodal ($f(-x) = -f(x)$) for all $x \in \partial B^n$.*

Lemma 3.13. *Borsuk-Ulam in the presentation of Theorem **??** implies Brouwer's fixed point theorem.*

Proof. Suppose there exists a continuous map $f : B^n \to B^n$ with no fixed points. That is, such that $f(x) \neq x$ for all $x \in B^n$.

Let us consider the map $g : B^n \to S^{n-1}$, where $g(x)$ is defined to be the unique intersection point between S^{n-1} and the open ray that starts at $f(x)$ and passes through x. That is,

$$S^{n-1} \cap \{f(x) + \lambda(x - f(x)) : \lambda > 0\} = \{g(x)\}$$

Notice that g is well defined because f has no fixed points and the ray is considered open. Also, since f is continuous, so is g.

Let $x \in \partial B^n$, by definition $g(x) = x$. Hence, if we consider antipodal points $\{x, -x\} \subset \partial B^n$ we have that $g(-x) = -x = -g(x)$. That is, g is antipodal in ∂B^n, which contradicts (5). $\square$

We are now going to see that Borsuk-Ulam theorem can be proved from Tucker's Lemma. Let us first introduce some notions:

Definition 3.14. *Let T be a triangulation of B^n, we say that it is* **antipodal symmetric** *if:*

1. *The set of simplices of T that are contained in $S^{n-1} = \partial B^n$ are a triangulation of S^{n-1}.*
2. *If $\sigma \in T$, $\sigma \subset S^{n-1}$, then $-\sigma$ is also in T.*

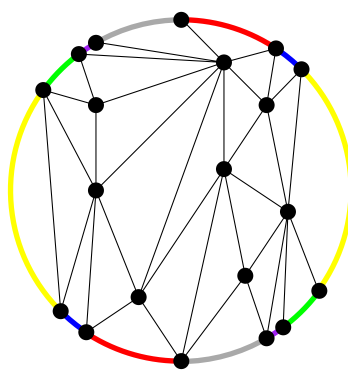


Figure 16: An antipodal symmetric triangulation of the 2-dimensional ball. Antipodal simplices in the boundary are painted with the same color.

We present to equivalent versions of Tucker's lemma (see exercises)

Lemma 3.15 (Tucker's Lemma, 1946- Version A)**.**

Let T be an antipodal symmetric triangulation of B^n and let $\lambda : V(T) \to \{\pm 1, \ldots, \pm n\}$ be a labeling on the vertices of T that satisfies $\lambda(-v) = -\lambda(v)$, $\forall v \in V(T) \cap \partial B^n$.

Then there exists and edge $(v, v') \in E(T)$ such that $\lambda(v) + \lambda(v') = 0$, that is, there exists a **complete edge** *in T.*

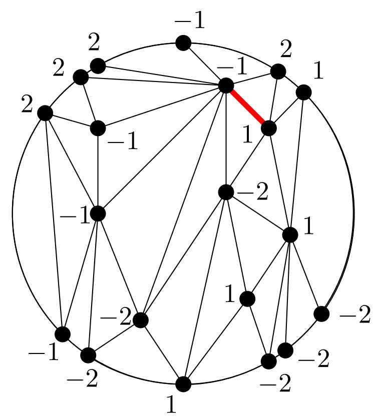


Figure 17: An antipodal symmetric triangulation of B^2 with a labeling in its vertices that is antipodal in the boundary. The red edge is a complete edge.

Lemma 3.16 (Tucker's Lemma, 1946- Version B)**.**

Let T be an antipodal symmetric triangulation of B^n. Then there is no simplicial map from T into $\Diamond^{n-1}$ that is antipodal in ∂B^n.

In the previous lemma, $\Diamond^{n-1}$ is the $(n-1)$-dimensional boundary of the n-crosspolytope.

Remember that a simplicial map between simplicial complexes sends vertices to vertices, and the image of the rest of the points is defined by affinely extending the images of the vertices. Notice that a simplicial map maps simplices to simplices.

Theorem 3.17. *Borsuk-Ulam in Theorem 3.12 is equivalent to Tucker's lemma.*

Proof. <u>Borsuk-Ulam (5) $\Longrightarrow$ Tucker (B)</u>

Suppose there exists a simplicial map g from T to $\Diamond^{n-1}$ that is antipodal in the boundary of T. That is, we have a continuous function $g : B^n \to \Diamond^{n-1}$

that is antipodal in ∂B^n. Consider now the function $f : B^n \to S^{n-1}$ defined by $f(x) = g(x)/||g(x)||$. f is, by construction, continuous and antipodal on the boundary, which is a contradiction with Borsuk-Ulam (5).

Tucker (A) $\Longrightarrow$ Borsuk-Ulam (5)

Suppose $f : B^n \to S^{n-1}$ is a continuous function, antipodal in ∂B^n. Since f is continuous and B^n is compact, f is uniformly continuous. Given $\epsilon = n^{-1/2}$ there exists a universal constant $\delta > 0$ such that, if $x, y \in B^n$ are such that $||x - y||_2 < \delta$ (norm 2 or ∞???), then $||f(x) - f(y)||_\infty < 2\epsilon$.

Let us now choose an antipodal symmetric triangulation T of diameter $\leq \delta$. Observe that such a triangulation can be found, since any antipodal symmetric triangulation can be made finer and finer by barycentric subdivision, being careful that antipodality at the boundary is preserved.

Let $k : V(T) \to \{1, \dots, n\}$ be a map defined as

$$k(v) := \min\{i \in \{1, \dots, n\} : |f(v)_i| \geq \epsilon\}, \quad v \in V(T).$$

Let us see that k well defined. Let $y = f(v) \in S^{n-1}$, and suppose that $|y_i| < \epsilon$, for all $i \in \{1, \dots, n\}$. Then $||y||_2^2 := y_1^2 + \cdots + y_n^2 < n\epsilon^2 = n(n^{-1/2})^2 = 1$, which is a contradiction with the fact that $y \in S^{n-1}$. Hence there exists at least one index i such that $|f(v)_i| \geq \epsilon$.

Let us now define the following labeling:

$$\lambda(v) := \begin{cases} +k(v) & \text{if } f(v)_{k(v)} > 0 \\ -k(v) & \text{if } f(v)_{k(v)} < 0 \end{cases}$$

Suppose now there exists a complete edge, that is there exist $v, v' \in V(T)$ such that $(v, v') \in E(T)$, $\lambda(v) = i$ and $\lambda(v') = -i$, for some $i \in \{1, \dots, n\}$. On one hand, since (v, v') is an edge, we have that $||v - v'||_2 \leq \mathrm{diam}(T) \leq \delta$. On the other, we have $f(v)_i \geq \epsilon$ and $-f(v')_i \geq \epsilon$, which implies that $||f(v) - f(v')||_\infty \geq |f(v)_i - f(v')_i| \geq 2\epsilon$. But this is a contradiction with the choice of δ. Hence no such complete edge exists, which contradicts Tucker (A). $\square$

3.2 Exercises:

1. Take a regular tetrahedron. Can you triangulate it, with the help of extra interior points, in such a way that only regular tetrahedra appear inside?

2. (**) Prove that it is always possible to label the vertices of the k-th barycentric subdivision of a $(d-1)$-dimensional simplex with labels

$1, 2, \ldots, d$ such that each simplex present obtains all d labels in its vertices.

3. How many vertices are there in the k-th barycenter subdivision of a simplex?

4. Provide an inductive proof of Sperner's lemma for arbitrary dimension. This is another existence proof, but can you think of an algorithmic way to find the fully colored simplices?

5. Prove Sperner's lemma (for simplices) using Brouwer's theorem directly (show they are equivalent!).

6. (**) Prove that it is enough to prove Brouwer's theorem for simplices to derive it for convex sets (or even topological balls).

7. Perron's theorem say that $n \times n$ positive matrix must have a positive eigenvalue with an eigenvector all of whose entries are positive. Use Brouwer's fixed point theorem to prove this fact.

8. (**) Here is another polytopal Sperner lemma (a la Budapest): Let P be a simple d-dimensional convex polytope (i.e., every vertex is in exactly d-facets). You color the facets with d-colors (no way or rule in particular). Suppose that at the end of your coloring, one of the vertices has its touching facets of different colors. Show that there most be at least another vertex of P with all the colors surrounding it. HINT: How can you reduce this to the usual Sperner's lemma?

9. Let X be a compact set in $\mathbb{R}^d$. Let $f : X \to \mathbb{R}^d$ be a usual function. Prove that f is a continuous function if and only if its graph $\{(x, y) : y = f(x)\}$ is a closed set.

10. Let X be a compact-convex set and define $f : X \to X$, with f a continuous function. Prove that f has a fixed point using Kakutani's fixed point theorem, i.e., prove that Kakutani fixed point theorem implies Brouwer fixed point theorem (under the above assumptions).

11. Prove that it is enough to prove Kakutani's fixed point theorem for simplexes to derive it for compact-convex sets.

12. (**) Write a careful proof of the existence of Nash equilibria for N players.

13. Let $X = [-1, 1]$ and $F(x) = \{$set of numbers y satisfying $x^2 + y^2 \geq 1/4\}$. Do all of the hypotheses of Kakutani's theorem hold for this example?

14. Consider the game with three players A, B, C in which each player writes a number from 1 to 10. If A writes i, B writes j and C writes k, then they have the playoffs

$$a_{ijk} = |i - j|,\ b_{ijk} = |j - k|,\ c_{ijk} = |k - i|.$$

Find at least one Nash equilibria. (BONUS: Can you find them all?).

15. Prove that the four statements (1), (2), (3) and (4) in Borsuk-Ulam's theorem Theorem 3.8 are in fact equivalent to:

There is no continuous map $f : B^n \to S^{n-1} = \partial B^n$ that is antipodal $(f(-x) = -f(x))$ for all $x \in \partial B^n$.

————–FORTHCOMING TOPICS ————–

4 Combinatorial Geometry Tools: Midterm 2

4.1 Lectures 8,9,10 Carathéodory-type theorems

4.2 Lectures 11,12,13 Helly-type theorems

5 Applications: Midterm 3

5.1 Lecture 14,15,16 Game Theory and Fair-Division

5.2 Lecture 17,18,19,20 Data Analysis and Optimization

References