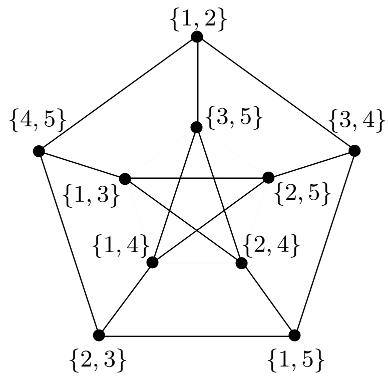
The chromatic number of Kneser graphs

Chapter 38

In 1955 the number theorist Martin Kneser posed a seemingly innocuous problem that became one of the great challenges in graph theory until a brilliant and totally unexpected solution, using the "Borsuk–Ulam theorem" from topology, was found by László Lovász twenty-three years later.

It happens often in mathematics that once a proof for a long-standing problem is found, a shorter one quickly follows, and so it was in this case. Within weeks Imre Bárány showed how to combine the Borsuk–Ulam theorem with another known result to elegantly settle Kneser's conjecture. Then in 2002 Joshua Greene, an undergraduate student, simplified Bárány's argument even further, and it is his version of the proof that we present here.



The Kneser graph $K(5,2)$ is the famous *Petersen graph*.

But let us start at the beginning. Consider the following graph $K(n,k)$, now called *Kneser graph*, for integers $n \geq k \geq 1$. The vertex-set $V(n,k)$ is the family of k-subsets of $\{1,\ldots,n\}$, thus $|V(n,k)| = \binom{n}{k}$. Two such k-sets A and B are adjacent if they are disjoint, $A \cap B = \varnothing$.

If $n < 2k$, then any two k-sets intersect, resulting in the uninteresting case where $K(n,k)$ has no edges. So we assume from now on that $n \geq 2k$.

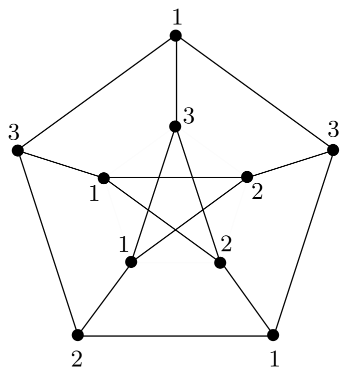
Kneser graphs provide an interesting link between graph theory and finite sets. Consider, e.g., the *independence number* $\alpha(K(n,k))$, that is, we ask how large a family of pairwise intersecting k-sets can be. The answer is given by the Erdős–Ko–Rado theorem of Chapter 27: $\alpha(K(n,k)) = \binom{n-1}{k-1}$.

This implies that
$\chi(K(n,k)) \geq \frac{|V|}{\alpha} = \frac{\binom{n}{k}}{\binom{n-1}{k-1}} = \frac{n}{k}$.

We can similarly study other interesting parameters of this graph family, and Kneser picked out the most challenging one: the *chromatic number* $\chi(K(n,k))$. We recall from previous chapters that a (vertex) coloring of a graph G is a mapping $c : V \to \{1,\ldots,m\}$ such that adjacent vertices are colored differently. The chromatic number $\chi(G)$ is then the minimum number of colors that is sufficient for a coloring of V. In other words, we want to present the vertex set V as a disjoint union of as few *color classes* as possible, $V = V_1 \,\dot\cup \cdots \dot\cup\, V_{\chi(G)}$, such that each set V_i is edgeless.

For the graphs $K(n,k)$ this asks for a partition $V(n,k) = V_1 \,\dot\cup \cdots \dot\cup\, V_\chi$, where every V_i is an *intersecting* family of k-sets. Since we assume that $n \geq 2k$, we write from now on $n = 2k+d$, $k \geq 1$, $d \geq 0$.

Here is a simple coloring of $K(n,k)$ that uses $d+2$ colors: For $i = 1, 2, \ldots, d+1$, let V_i consist of all k-sets that have i as smallest element. The remaining k-sets are contained in the set $\{d+2, d+3, \ldots, 2k+d\}$, which has only $2k-1$ elements. Hence they all intersect, and we can use color $d+2$ for all of them.



The 3-coloring of the Petersen graph.

M. Aigner, G.M. Ziegler, *Proofs from THE BOOK*,
DOI 10.1007/978-3-642-00856-6_38,

So we have $\chi(K(2k+d,k)) \leq d+2$, and Kneser's challenge was to show that this is the right number.

> **Kneser's conjecture:** We have
>
> $$\chi(K(2k+d,k)) \;=\; d+2.$$

Probably anybody's first crack at the proof would be to try induction on k and d. Indeed, the starting cases $k=1$ and $d=0,1$ are easy, but the induction step from k to $k+1$ (or d to $d+1$) does not seem to work. So let us instead reformulate the conjecture as an existence problem:

For $d=0$, $K(2k,k)$ consists of disjoint edges, one for every pair of complementary k-sets. Hence $\chi(K(2k,k))=2$, in accordance with the conjecture.

If the family of k-sets of $\{1,2,\ldots,2k+d\}$ is partitioned into $d+1$ classes, $V(n,k) = V_1 \,\dot\cup \cdots \dot\cup\, V_{d+1}$, then for some i, V_i contains a pair A,B of disjoint k-sets.

Lovász' brilliant insight was that at the (topological) heart of the problem lies a famous theorem about the d-dimensional unit sphere S^d in $\mathbb{R}^{d+1}$, $S^d = \{x \in \mathbb{R}^{d+1} : |x| = 1\}$.

> **The Borsuk–Ulam theorem.**
> *For every continuous map $f : S^d \to \mathbb{R}^d$ from d-sphere to d-space, there are antipodal points $x^*, -x^*$ that are mapped to the same point $f(x^*) = f(-x^*)$.*

This result is one of the cornerstones of topology; it first appears in Borsuk's famous 1933 paper. We sketch a proof in the appendix; for the full proof we refer to Section 2.2 in Matoušek's wonderful book "Using the Borsuk–Ulam theorem", whose very title demonstrates the power and range of the result. Indeed, there are many equivalent formulations, which underline the central position of the theorem. We will employ a version that can be traced back to a book by Lyusternik–Shnirel'man from 1930, which even predates Borsuk.

Theorem. *If the d-sphere S^d is covered by $d+1$ sets,*

$$S^d \;=\; U_1 \cup \;\cdots\; \cup U_d \cup U_{d+1},$$

such that each of the first d sets $U_1,\ldots,U_d$ is either open or closed, then one of the $d+1$ sets contains a pair of antipodal points $x^*, -x^*$.

The case when all $d+1$ sets are closed is due to Lyusternik and Shnirel'man. The case when all $d+1$ sets are open is equally common, and also called the Lyusternik–Shnirel'man theorem. Greene's insight was that the theorem is also true if each of the $d+1$ sets is *either open or closed*. As you will see, we don't even need that: No such assumption is needed for U_{d+1}. For the proof of Kneser's conjecture, we only need the case when $U_1,\ldots,U_d$ are open.

■ **Proof of the Lyusternik–Shnirel'man theorem using Borsuk–Ulam.** Let a covering $S^d = U_1 \cup \cdots \cup U_d \cup U_{d+1}$ be given as specified, and assume that there are no antipodal points in any of the sets U_i. We define a map $f : S^d \to \mathbb{R}^d$ by

$$f(x) := \big(\delta(x, U_1), \delta(x, U_2), \ldots, \delta(x, U_d)\big).$$

Here $\delta(x, U_i)$ denotes the distance of x from U_i. Since this is a continuous function in x, the map f is continuous. Thus the Borsuk–Ulam theorem tells us that there are antipodal points $x^*, -x^*$ with $f(x^*) = f(-x^*)$. Since U_{d+1} does not contain antipodes, we get that at least one of x^* and $-x^*$ must be contained in one of the sets U_i, say in U_k ($k \le d$). After exchanging x^* with $-x^*$ if necessary, we may assume that $x^* \in U_k$. In particular this yields $\delta(x^*, U_k) = 0$, and from $f(x^*) = f(-x^*)$ we get that $\delta(-x^*, U_k) = 0$ as well.

If U_k is closed, then $\delta(-x^*, U_k) = 0$ implies that $-x^* \in U_k$, and we arrive at the contradiction that U_k contains a pair of antipodal points.

If U_k is open, then $\delta(-x^*, U_k) = 0$ implies that $-x^*$ lies in $\overline{U_k}$, the closure of U_k. The set $\overline{U_k}$, in turn, is contained in $S^d \backslash (-U_k)$, since this is a closed subset of S^d that contains U_k. But this means that $-x^*$ lies in $S^d \backslash (-U_k)$, so it cannot lie in $-U_k$, and x^* cannot lie in U_k, a contradiction. □

The closure of U_k is the smallest closed set that contains U_k (that is, the intersection of all closed sets containing U_k).

As the second ingredient for his proof, Imre Bárány used another existence result about the sphere S^d.

Gale's Theorem. *There is an arrangement of $2k + d$ points on S^d such that every open hemisphere contains at least k of these points.*

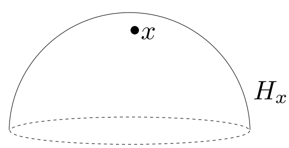
David Gale discovered his theorem in 1956 in the context of polytopes with many faces. He presented a complicated induction proof, but today, with hindsight, we can quite easily exhibit such a set and verify its properties.

Armed with these results it is just a short step to settle Kneser's problem, but as Greene showed we can do even better: We don't even need Gale's result. It suffices to take any arrangement of $2k + d$ points on S^{d+1} in *general position*, meaning that no $d + 2$ of the points lie on a hyperplane through the center of the sphere. Clearly, for $d \ge 0$ this can be done.

■ **Proof of the Kneser conjecture.** For our ground set let us take $2k + d$ points in general position on the sphere S^{d+1}. Suppose the set $V(n, k)$ of all k-subsets of this set is partitioned into $d + 1$ classes, $V(n, k) = V_1 \,\dot\cup \cdots \dot\cup\, V_{d+1}$. We have to find a pair of disjoint k-sets A and B that belong to the same class V_i.

For $i = 1, \ldots, d + 1$ we set

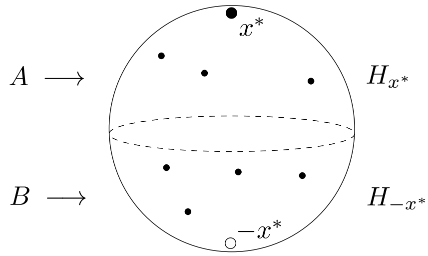
$$O_i = \{x \in S^{d+1} : \text{the open hemisphere } H_x \text{ with pole } x \text{ contains a } k\text{-set from } V_i\}.$$



An open hemisphere in S^2

Clearly, each O_i is an open set. Together, the open sets O_i and the closed set $C = S^{d+1} \setminus (O_1 \cup \cdots \cup O_{d+1})$ cover S^{d+1}. Invoking Lyusternik–Shnirel'man we know that one of these sets contains antipodal points x^*

and $-x^*$. This set cannot be C! Indeed, if $x^*, -x^*$ are in C, then by the definition of the O_i's, the hemispheres H_{x^*} and H_{-x^*} would contain fewer than k points. This means that at least $d+2$ points would be on the equator $\overline{H}_{x^*} \cap \overline{H}_{-x^*}$ with respect to the north pole x^*, that is, on a hyperplane through the origin. But this cannot be since the points are in general position. Hence some O_i contains a pair $x^*, -x^*$, so there exist k-sets A and B both in class V_i, with $A \subseteq H_{x^*}$ and $B \subseteq H_{-x^*}$.



But since we are talking about *open* hemispheres, H_{x^*} and H_{-x^*} are disjoint, hence A and B are disjoint, and this is the whole proof. $\square$

The reader may wonder whether sophisticated results such as the theorem of Borsuk–Ulam are really necessary to prove a statement about finite sets. Indeed, a beautiful combinatorial argument has recently been found by Jiří Matoušek — but on closer inspection it has a distinct, albeit discrete, topological flavor.

Appendix: A proof sketch for the Borsuk–Ulam theorem

For any *generic* map (also known as *general position* map) from a compact d-dimensional space to a d-dimensional space, any point in the image has only a finite number of pre-images. For a generic map from a $(d+1)$-dimensional space to a d-dimensional space, we expect every point in the image to have a 1-dimensional pre-image, that is, a collection of curves. Both in the case of smooth maps, and in the setting of piecewise-linear maps, one quite easily proves one can deform any map to a nearby generic map.

For the Borsuk–Ulam theorem, the idea is to show that every generic map $S^d \to \mathbb{R}^d$ identifies an odd (in particular, finite and nonzero) number of antipodal pairs. If f did not identify any antipodal pair, then it would be arbitrarily close to a generic map $\widetilde{f}$ without any such identification.

Now consider the projection $\pi : S^d \to \mathbb{R}^d$ that just deletes the last coordinate; this map identifies the "north pole" e_{d+1} of the d-sphere with the "south pole" $-e_{d+1}$. For any given map $f : S^d \to \mathbb{R}^d$ we construct a continuous deformation from π to f, that is, we interpolate between these two maps (linearly, for example), to obtain a continuous map

$$F : S^d \times [0,1] \;\longrightarrow\; \mathbb{R}^d,$$

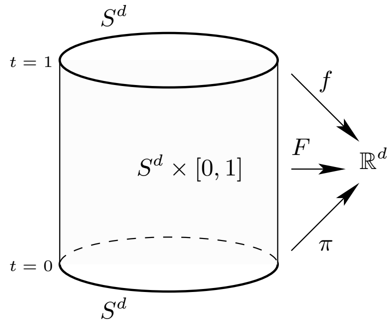
with $F(x,0) = \pi(x)$ and $F(x,1) = f(x)$ for all $x \in S^d$. (Such a map is known as a *homotopy*.)

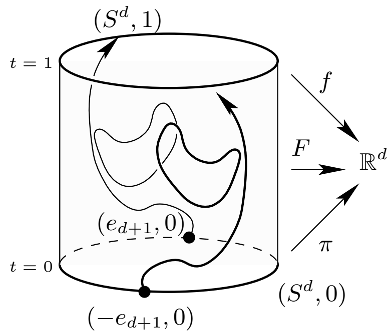
Now we perturb F carefully into a generic map $\widetilde{F} : S^d \times [0,1] \to \mathbb{R}^d$, which again we may assume to be smooth, or piecewise-linear on a fine triangulation of $S^d \times [0,1]$. If this perturbation is "small enough" and performed carefully, then the perturbed version of the projection $\widetilde{\pi}(x) := \widetilde{F}(x,0)$ should still identify the two antipodal points $\pm e_{d+1}$ and no others. If $\widetilde{F}$ is sufficiently generic, then the points in $S^d \times [0,1]$ given by

$$M \;:=\; \big\{(x,t) \in S^d \times [0,1] : \widetilde{F}(-x,t) = \widetilde{F}(x,t)\big\}$$

according to the implicit function theorem (smooth or piecewise-linear version) form a collection of paths and of closed curves. Clearly this collection is *symmetric*, that is, $(-x,t) \in M$ if and only if $(x,t) \in M$.

The paths in M can have endpoints only at the boundary of $S^d \times [0,1]$, that is, at $t = 0$ and at $t = 1$. The only ends at $t = 0$, however, are at $(\pm e_{d+1}, 0)$, and the two paths that start at these two points are symmetric copies of each other, so they are disjoint, and they can end only at $t = 1$. This proves that there are solutions for $\widetilde{F}(-x,t) = \widetilde{F}(x,t)$ at $t = 1$, and hence for $f(-x) = f(x)$. □





References

[1] I. BÁRÁNY: *A short proof of Kneser's conjecture,* J. Combinatorial Theory, Ser. B **25** (1978), 325-326.

[2] K. BORSUK: *Drei Sätze über die n-dimensionale Sphäre,* Fundamenta Math. **20** (1933), 177-190.

[3] D. GALE: *Neighboring vertices on a convex polyhedron,* in: "Linear Inequalities and Related Systems" (H. W. Kuhn, A. W. Tucker, eds.), Princeton University Press, Princeton 1956, 255-263.

[4] J. E. GREENE: *A new short proof of Kneser's conjecture,* American Math. Monthly **109** (2002), 918-920.

[5] M. KNESER: *Aufgabe 360,* Jahresbericht der Deutschen Mathematiker-Vereinigung **58** (1955), 27.

[6] L. LOVÁSZ: *Kneser's conjecture, chromatic number, and homotopy,* J. Combinatorial Theory, Ser. B **25** (1978), 319-324.

[7] L. LYUSTERNIK & S. SHNIREL'MAN: *Topological Methods in Variational Problems (in Russian),* Issledowatelskiĭ Institute Matematiki i Mechaniki pri O. M. G. U., Moscow, 1930.

[8] J. MATOUŠEK: *Using the Borsuk–Ulam Theorem. Lectures on Topological Methods in Combinatorics and Geometry,* Universitext, Springer-Verlag, Berlin 2003.

[9] J. MATOUŠEK: *A combinatorial proof of Kneser's conjecture,* Combinatorica **24** (2004), 163-170.