Thirty-three Miniatures: Mathematical and Algorithmic Applications of Linear Algebra

Jiřì Matoušek

This is a preliminary version of the book ***Thirty-three Miniatures: Mathematical and Algorithmic Applications of Linear Algebra*** published by the American Mathematical Society (AMS). This preliminary version is made available with the permission of the AMS and may not be changed, edited, or reposted at any other website without explicit written permission from the author and the AMS.

1991 *Mathematics Subject Classification.* 05C50, 68Wxx, 15-01

Contents

Introduction vii

Notation 1

Miniature 1. Fibonacci Numbers, Quickly 3

Miniature 2. Fibonacci Numbers, the Formula 5

Miniature 3. The Clubs of Oddtown 7

Miniature 4. Same-Size Intersections 9

Miniature 5. Error-Correcting Codes 11

Miniature 6. Odd Distances 17

Miniature 7. Are These Distances Euclidean? 19

Miniature 8. Packing Complete Bipartite Graphs 23

Miniature 9. Equiangular Lines 27

Miniature 10. Where is the Triangle? 31

Miniature 11. Checking Matrix Multiplication 35

Miniature 12. Tiling a Rectangle by Squares 39

Miniature 13. Three Petersens Are Not Enough 41

Miniature 14. Petersen, Hoffman–Singleton, and Maybe 57 43

Miniature 15. Only Two Distances 49

Miniature 16. Covering a Cube Minus One Vertex 53

Miniature 17. Medium-Size Intersection Is Hard To Avoid 55

Miniature 18. On the Difficulty of Reducing the Diameter 59

Miniature 19. The End of the Small Coins 65

Miniature 20. Walking in the Yard 69

Miniature 21. Counting Spanning Trees 75

Miniature 22. In How Many Ways Can a Man Tile a Board? 83

Miniature 23. More Bricks—More Walls? 95

Miniature 24. Perfect Matchings and Determinants 105

Miniature 25. Turning a Ladder Over a Finite Field 111

Miniature 26. Counting Compositions 117

Miniature 27. Is It Associative? 123

Miniature 28. The Secret Agent and the Umbrella 129

Miniature 29. Shannon Capacity of the Union: A Tale of Two Fields 137

Miniature 30. Equilateral Sets 145

Miniature 31. Cutting Cheaply Using Eigenvectors 151

Miniature 32. Rotating the Cube 161

Miniature 33. Set Pairs and Exterior Products 169

Index 177

Introduction

Some years ago I started gathering nice applications of linear algebra, and here is the resulting collection. The applications belong mostly to the main fields of my mathematical interests—combinatorics, geometry, and computer science. Most of them are mathematical, in proving theorems, and some include clever ways of computing things, i.e., algorithms. The appearance of linear-algebraic methods is often unexpected.

At some point I started to call the items in the collection "miniatures". Then I decided that in order to qualify for a miniature, a complete exposition of a result, with background and everything, shouldn't exceed four typeset pages (A4 format). This rule is absolutely arbitrary, as rules often are, but it has some rational core—namely, this extent can usually be covered conveniently in a 90 minute lecture, the standard length at the universities where I happened to teach. Then, of course, there are some exceptions to the rule, six-page miniatures that I just couldn't bring myself to omit.

The collection could obviously be extended indefinitely, but I thought thirty three was a nice enough number and a good point to stop.

The exposition is intended mainly for lecturers (I've taught almost all of the pieces at various occasions) and also for students interested in nice mathematical ideas even when they require some thinking. The material is hopefully class-ready, where all details left to the reader should indeed be devil-free.

I assume background of basic linear algebra, a bit of familiarity with polynomials, and some graph-theoretical and geometric terminology. The sections have varying levels of difficulty and generally I have ordered them from what I personally regard as the most accessible to the more demanding.

I wanted each section to be essentially self-contained. With a good undergraduate background you can as well start reading at Section 24. This is kind of opposite to a typical mathematical textbook, where material is developed gradually, and if one wants to make sense of something on page 123, one usually has to understand the previous 122 pages, or with some luck, suitable 38 pages.

Of course, the anti-textbook structure leads to some boring repetitions and, perhaps more seriously, it puts a limit on the degree of achievable sophistication. On the other hand, I believe there are advantages as well: I gave up reading several textbooks well before page 123, after I realized that between the usually short reading sessions I couldn't remember the key definitions (people with small children will know what I'm talking about).

After several sections the reader may spot certain common patterns in the presented proofs, which could be discussed at great length, but I have decided to leave out any general accounts on linear-algebraic methods.

Nothing in this text is original, and some of the examples are rather well known and appear in many publications (including, in few cases, other books of mine). Several general reference books are listed below. I've also added references to the original sources where I could find them. However, I've kept the historical notes at a minimum and I've put only a limited effort into tracing the origins of the ideas (many apologies to authors whose work is quoted badly or not at all—I will be glad to hear about such cases).

I would appreciate to learn about mistakes and suggestions of how to improve the exposition.

Further reading. An excellent textbook is

> L. Babai and P. Frankl, *Linear Algebra Methods in Combinatorics (Preliminary version 2)*, Department of Computer Science, The University of Chicago, 1992.

Unfortunately, it has never been published officially and it can be obtained, with some effort, as lecture notes of the University of Chicago. It contains several of the topics discussed here, a lot of other material in a similar spirit, and a very nice exposition of some parts of linear algebra.

Algebraic graph theory is treated, e.g., in the books

> N. Biggs, *Algebraic Graph Theory*, 2nd edition, Cambridge Univ. Press, Cambridge, 1993

and

> C. Godsil and G. Royle, *Algebraic Graph Theory*, Springer, New York, NY, 2001.

Probabilistic algorithms in the spirit of Sections 11 and 24 are well explained in the book

> R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, Cambridge, 1995.

Acknowledgments. For valuable comments on preliminary versions of this booklet I would like to thank Otfried Cheong, Esther Ezra, Nati Linial, Jana Maxová, Helena Nyklová, Yoshio Okamoto, Pavel Paták, Oleg Pikhurko, and Zuzana Safernová, as well as all other people whom I may have forgotten to include in this list. Thanks also to David Wilson for permission to use his picture of a random lozenge tiling in Miniature 22. Finally, I'm grateful to many people at the Department of Applied Mathematics of the Charles University in Prague and at the Institute of Theoretical Computer Science of the ETH Zurich for excellent working environments.

Notation

Most of the notation is recalled in each section where it is used. Here are several general items that may not be completely unified in the literature.

The integers are denoted by $\mathbb{Z}$, the rationals by $\mathbb{Q}$, the reals by $\mathbb{R}$, and $\mathbb{F}_q$ stands for the q-element finite field.

The transpose of a matrix A is written as A^T. The elements of that matrix are denoted by a_{ij}, and similarly for all other Latin letters. Vectors are typeset in boldface: $\mathbf{v}, \mathbf{x}, \mathbf{y}$, and so on. If $\mathbf{x}$ is a vector in $\mathbb{K}^n$, where $\mathbb{K}$ is some field, x_i stands for the ith component, so $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

We write $\langle \mathbf{x}, \mathbf{y} \rangle$ for the standard scalar (or inner) product of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$: $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y + x_2 y_2 + \cdots + x_n y_n$. We also interpret such $\mathbf{x}, \mathbf{y}$ as $n \times 1$ (single-column) matrices, and thus $\langle \mathbf{x}, \mathbf{y} \rangle$ could also be written as $\mathbf{x}^T \mathbf{y}$. Further, for $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}$ is the Euclidean norm (length) of the vector $\mathbf{x}$.

Graphs are simple and undirected unless stated otherwise; i.e., a graph G is regarded as a pair (V, E), where V is the vertex set and E is the edge set, which is a set of unordered pairs of elements of V. For a graph G, we sometimes write $V(G)$ for the vertex set and $E(G)$ for the edge set.

Miniature 1

Fibonacci Numbers, Quickly

The **Fibonacci numbers** $F_0, F_1, F_2, \ldots$ are defined by the relations $F_0 = 0$, $F_1 = 1$, and $F_{n+2} = F_{n+1} + F_n$ for $n = 0, 1, 2, \ldots$. Obviously, F_n can be calculated using roughly n arithmetic operations.

By the following trick we can compute it faster, using only about $\log n$ arithmetic operations. We set up the 2×2 matrix

$$M := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$\begin{pmatrix} F_{n+2} \\ F_{n+1} \end{pmatrix} = M \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix},$$

and therefore,

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = M^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

(we use the associativity of matrix multiplication).

For $n = 2^k$ we can compute M^n by repeated squaring, with k multiplications of 2×2 matrices. For n arbitrary, we write n in binary as $n = 2^{k_1} + 2^{k_2} + \cdots + 2^{k_t}$, $k_1 < k_2 < \cdots < k_t$, and then we calculate the power M^n as $M^n = M^{2^{k_1}} M^{2^{k_2}} \cdots M^{2^{k_t}}$. This needs at most $2k_t \le 2\log_2 n$ multiplications of 2×2 matrices.

Remarks. A similar trick can be used for any sequence $(y_0, y_1, y_2, \ldots)$ defined by a recurrence $y_{n+k} = a_{k-1}y_{n+k-1} + \cdots + a_0 y_n$, where k and $a_0, a_1, \ldots, a_{k-1}$ are constants.

If we want to compute the Fibonacci numbers by this method, we have to be careful, since the F_n grow very fast. From a formula in Miniature 2 below, one can see that the number of decimal digits of F_n is of order n. Thus we must use multiple precision arithmetic, and so the arithmetic operations will be relatively slow.

Sources. This trick is well known but so far I haven't encountered any reference to its origin.

Miniature 2

Fibonacci Numbers, the Formula

We derive a formula for the nth Fibonacci number F_n. Let us consider the vector space of all infinite sequences $(u_0, u_1, u_2, \ldots)$ of real numbers, with coordinate-wise addition and multiplication by real numbers. In this space we define a subspace W of all sequences satisfying the equation $u_{n+2} = u_{n+1} + u_n$ for all $n = 0, 1, \ldots$. Each choice of the first two members u_0 and u_1 uniquely determines a sequence from W, and therefore, $\dim(W) = 2$. (In more detail, the two sequences beginning with $(0, 1, 1, 2, 3, \ldots)$ and with $(1, 0, 1, 1, 2 \ldots)$ constitute a basis of W.)

Now we find another basis of W: two sequences whose terms are defined by a simple formula. Here we need an "inspiration": We should look for sequences $\mathbf{u} \in W$ in the form $u_n = \tau^n$ for a suitable real number τ.

Finding the right values of τ leads to the quadratic equation $\tau^2 = \tau + 1$, which has two distinct roots $\tau_{1,2} = (1 \pm \sqrt{5})/2$.

The sequences $\mathbf{u} := (\tau_1^0, \tau_1^1, \tau_1^2, \ldots)$ and $\mathbf{v} := (\tau_2^0, \tau_2^1, \tau_2^2, \ldots)$ both belong to W, and it is easy to verify that they are linearly independent (this can be checked by considering the first two terms). Hence they form a basis of W.

We express the sequence $\mathbf{F} := (F_0, F_1, \ldots)$ of the Fibonacci numbers in this basis: $\mathbf{F} = \alpha\mathbf{u} + \beta\mathbf{v}$. The coefficients α, β are calculated by considering the first two terms of the sequences; that is, we need to solve the linear system $\alpha\tau_1^0 + \beta\tau_2^0 = F_0$, $\alpha\tau_1^1 + \beta\tau_2^1 = F_1$.

The resulting formula is

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

It is amazing that this formula full of irrationals yields an integer for every n.

A similar technique works for other recurrences in the form $y_{n+k} = a_{k-1}y_{n+k-1} + \cdots + a_0 y_n$, but additional complications appear in some cases. For example, for $y_{n+2} = 2y_{n+1} - y_n$, one has to find a different kind of basis, which we won't do here.

Sources. The above formula for F_n is sometimes called *Binet's formula*, but it was known to Daniel Bernoulli, Euler, and de Moivre in the 18th century before Binet's work.

A more natural way of deriving the formula is using generating functions, but doing this properly and from scratch takes more work.

Miniature 3

The Clubs of Oddtown

There are n citizens living in Oddtown. Their main occupation was forming various clubs, which at some point started threatening the very survival of the city. In order to limit the number of clubs, the city council decreed the following innocent-looking rules:

- Each club has to have an *odd* number of members.
- Every two clubs must have an *even* number of members in common.

Theorem. *Under these rules, it is impossible to form more clubs than n, the number of citizens.*

Proof. Let us call the citizens $1, 2, \ldots, n$ and the clubs $C_1, C_2, \ldots, C_m$. We define an $m \times n$ matrix A by

$$a_{ij} = \begin{cases} 1 & \text{if } j \in C_i, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

(Thus clubs correspond to rows and citizens to columns.)

Let us consider the matrix A over the two-element field $\mathbb{F}_2$. Clearly, the rank of A is at most n.

Next, we look at the product AA^T. This is an $m \times m$ matrix whose entry at position (i,k) equals $\sum_{j=1}^{n} a_{ij}a_{kj}$, and so it counts the number of citizens in $C_i \cap C_k$. More precisely, since we now work over $\mathbb{F}_2$, the entry is 1 if $|C_i \cap C_k|$ is odd, and it is 0 for $|C_i \cap C_k|$ even.

Therefore, the rules of the city council imply that $AA^T = I_m$, where I_m denotes the identity matrix. So the rank of AA^T is at least m. Since the rank of a matrix product is no larger that the minimum of the ranks of the factors, we have $\mathrm{rank}(A) \geq m$ as well, and so $m \leq n$. The theorem is proved. □

Sources. This is the opening example in the book of Babai and Frankl cited in the introduction. I am not sure if it appears earlier in this "pure form", but certainly it is a special case of other results, such as the Frankl–Wilson inequality (see Miniature 17).

Miniature 4

Same-Size Intersections

The result and proof of this section are similar to those in Miniature 3.

Theorem (Generalized Fisher inequality)**.** *If $C_1, C_2, \ldots, C_m$ are distinct and nonempty subsets of an n-element set such that all the intersections $C_i \cap C_j$, $i \neq j$, have the same size, then $n \geq m$.*

Proof. Let $|C_i \cap C_j| = t$ for all $i \neq j$.

First we need to deal separately with the situation that some C_i, say C_1, has size t. Then $t \geq 1$ and C_1 is contained in every other C_j. Thus $C_i \cap C_j = C_1$ for all $i, j \geq 2$, $i \neq j$. Then the sets $C_i \setminus C_1$, $i \geq 2$, are all disjoint and nonempty, and so their number is at most $n - |C_1| \leq n - 1$. Together with C_1 these are at most n sets.

Now we assume that $d_i := |C_i| > t$ for all i. As in Miniature 3, we set up the $m \times n$ matrix A with

$$a_{ij} = \begin{cases} 1 & \text{if } j \in C_i, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Now we consider A as a matrix with real entries, and we let $B := AA^T$. Then

$$B = \begin{pmatrix} d_1 & t & t & \ldots & t \\ t & d_2 & t & \ldots & t \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t & t & t & \ldots & d_m \end{pmatrix},$$

$t \geq 0$, $d_1, d_2, \ldots, d_m > t$. It remains to verify that B is nonsingular; then we will have $m = \text{rank}(B) \leq \text{rank}(A) \leq n$ and we will be done.

The nonsingularity of B can be checked in a pedestrian way, by bringing B to a triangular form by a suitably organized Gaussian elimination.

Here is another way. We will show that B is **positive definite**; that is, B is symmetric and $\mathbf{x}^T B\mathbf{x} > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^m$.

We can write $B = tJ_n + D$, where J_n is the all 1's matrix and D is the diagonal matrix with $d_1 - t, d_2 - t, \ldots, d_n - t$ on the diagonal.

Let $\mathbf{x}$ be an arbitrary nonzero vector in $\mathbb{R}^n$. Clearly, D is positive definite, since $\mathbf{x}^T D\mathbf{x} = \sum_{i=1}^n (d_i - t)x_i^2 > 0$. For J_n, we have $\mathbf{x}^T J_n \mathbf{x} = \sum_{i,j=1}^n x_i x_j = \left(\sum_{i=1}^n x_i\right)^2 \geq 0$, so J_n is *positive semidefinite*. Finally, $\mathbf{x}^T B\mathbf{x} = \mathbf{x}^T(tJ_n + D)\mathbf{x} = t\mathbf{x}^T J_n\mathbf{x} + \mathbf{x}^T D\mathbf{x} > 0$, an instance of a general fact that the sum of a positive definite matrix and a positive semidefinite one is positive definite.

So B is positive definite. It remains to see (or know) that all positive definite matrices are nonsingular. Indeed, if $B\mathbf{x} = \mathbf{0}$, then $\mathbf{x}^T B\mathbf{x} = \mathbf{x}^T\mathbf{0} = 0$, and hence $\mathbf{x} = \mathbf{0}$. □

Sources. A somewhat special case of the inequality comes from

> R. A. Fisher, *An examination of the different possible solutions of a problem in incomplete blocks*, Ann. Eugenics **10** (1940), 52–75.

A linear-algebraic proof of a "uniform" version of Fisher's inequality is due to

> R. C. Bose, *A note on Fisher's inequality for balanced incomplete block designs*, Ann. Math. Statistics **20,4** (1949), 619–620.

The nonuniform version as above was noted in

> K. N. Majumdar, *On some theorems in combinatorics relating to incomplete block designs*, Ann. Math. Statistics **24** (1953), 377–389

and rediscovered in

> J. R. Isbell, *An inequality for incidence matrices*, Proc. Amer. Math. Soc. **10** (1959), 216–218.

Miniature 5

Error-Correcting Codes

We want to transmit (or write and read) some data, say a string $\mathbf{v}$ of 0's and 1's. The transmission channel is not completely reliable, and so some errors may occur—some 0's may be received as 1's and vice versa. We assume that the probability of error is small, and that the probability of k errors in the message is substantially smaller than the probability of $k-1$ or fewer errors.

The main idea of error-correcting codes is to send, instead of the original message $\mathbf{v}$, a somewhat longer message $\mathbf{w}$. This longer string $\mathbf{w}$ is constructed so that we can correct a small number of errors incurred in the transmission.

Today error-correcting codes are used in many kinds of devices, ranging from CD players to spacecrafts, and the construction of error-correcting codes constitutes an extensive area of research. Here we introduce the basic definitions and we present an elegant construction of an error-correcting code based on linear algebra.

Let us consider the following specific problem: We want to send arbitrary 4-bit strings $\mathbf{v}$ of the form $abcd$, where $a, b, c, d \in \{0, 1\}$. We assume that the probability of two or more errors in the transmission is negligible, but a single error occurs with a non-negligible probability, and we would like to correct it.

One way of correcting a single error is to triple every bit and send $\mathbf{w} = aaabbbcccddd$ (12 bits). For example, instead of $\mathbf{v} = 1011$, we send $\mathbf{w} = 111000111111$. If, say, 110000111111 is received at the other end of the channel, we know that there was an error in the third bit and the correct string was 111000111111 (unless, of course, there were two or more errors after all).

That is a rather wasteful way of coding. We will see that one can correct an error in any single bit using a code that transforms a 4-bit message into a 7-bit string. So the message is expanded not three times, but only by 75 %.

Example: The Hamming code. This is probably the first known non-trivial error-correcting code and it was discovered in the 1950s. Instead of a given 4-bit string $\mathbf{v} = abcd$, we send the 7-bit string $\mathbf{w} = abcdefg$, where $e := a+b+c$ (addition modulo 2), $f := a+b+d$ and $g := a+c+d$. For example, for $\mathbf{v} = 1011$, we have $\mathbf{w} = 1011001$. This encoding also allows us to correct any single-bit error, as we will prove using linear algebra.

Before we get to that, we introduce some general definitions from coding theory.

Let S be a finite set, called the **alphabet**; for example, we can have $S = \{0, 1\}$ or $S = \{\mathrm{a}, \mathrm{b}, \mathrm{c}, \mathrm{d}, \ldots, \mathrm{z}\}$. We write $S^n = \{\mathbf{w} = a_1a_2 \ldots a_n : a_1, \ldots, a_n \in S\}$ for the set of all possible words of length n (here a **word** means any arbitrary finite sequence of letters of the alphabet).

Definition. A **code** of length n over an alphabet S is an arbitrary subset $C \subseteq S^n$.

For example, for the Hamming code, we have $S = \{0, 1\}$, $n = 7$, and C is the set of all 7-bit words that can arise by the encoding procedure described above from all the $2^4 = 16$ possible 4-bit words. That is, $C = \{$0000000, 0001011, 0010101, 0011110, 0100110, 0101101, 0110011, 0111000, 1000111, 1001100, 1010010, 1011001, 1100001, 1101010, 1110100, 1111111$\}$.

The essential property of this code is that every two of its words differ in at least 3 bits. We could check this directly, but laboriously, by comparing every pair of words in C. Soon we will prove it differently and almost effortlessly.

We introduce the following terminology:

- The **Hamming distance** of two words $\mathbf{u}, \mathbf{v} \in S^n$ is
$$d(\mathbf{u}, \mathbf{v}) := |\{i : u_i \neq v_i, i = 1, 2, \ldots, n\}|,$$
where u_i is the ith letter of the word $\mathbf{u}$. It means that we can get $\mathbf{v}$ by making $d(\mathbf{u}, \mathbf{v})$ "errors" in $\mathbf{u}$.
- A code C **corrects t errors** if for every $\mathbf{u} \in S^n$ there is at most one $\mathbf{v} \in C$ with $d(\mathbf{u}, \mathbf{v}) \leq t$.
- The **minimum distance** of a code C is defined as $d(C) := \min\{d(\mathbf{u}, \mathbf{v}) : \mathbf{u}, \mathbf{v} \in C, \mathbf{u} \neq \mathbf{v}\}$.

It is easy to check that the last two notions are related as follows: *A code C corrects t errors if and only if $d(C) \geq 2t + 1$.* So for showing that the Hamming code corrects one error we need to prove that $d(C) \geq 3$.

Encoding and decoding. The above definition of a code may look strange, since in everyday usage, a "code" refers to a method of encoding messages. Indeed, in order to actually use a code C as in the above definition, we also need an injective mapping $c\colon \Sigma^k \to C$, where Σ is the alphabet of the original message and k is its length (or the length of a block used for transmission).

For a given message $\mathbf{v} \in \Sigma^k$, we compute the code word $\mathbf{w} = c(\mathbf{v}) \in C$ and we send it. Then, having received a word $\mathbf{w}' \in S^n$, we find a word $\mathbf{w}'' \in C$ minimizing $d(\mathbf{w}', \mathbf{w}'')$, and we calculate $\mathbf{v}' = c^{-1}(\mathbf{w}'') \in \Sigma^k$ for this $\mathbf{w}''$. If at most t errors occurred during the transmission and C corrects t errors, then $\mathbf{w}'' = \mathbf{w}$, and thus $\mathbf{v}' = \mathbf{v}$. In other words, we recover the original message.

One of the main problems of coding theory is to find, for given S, t, and n, a code C of length n over the alphabet S with $d(C) \geq t$ and with as many words as possible (since the larger $|C|$, the more information can be transmitted).

We also need to compare the quality of codes with different $|S|, t, n$. Such things are studied by *Shannon's information theory*, which we will not pursue here.

When constructing a code, other aspects besides its size need also be taken into account, e.g., the speed of encoding and decoding.

Linear codes. Linear codes are codes of a special type, and the Hamming code is one of them. In this case, the alphabet S is a finite field (the most important example is $S = \mathbb{F}_2$), and thus S^n is a vector space over S. Every linear subspace of S^n is called a **linear code**.

Observation. *For every linear code C, we have*

$$d(C) = \min\{d(\mathbf{0}, \mathbf{w}) : \mathbf{w} \in C, \mathbf{w} \neq \mathbf{0}\}.$$

□

A linear code need not be given as a list of codewords. Linear algebra offers us two basic ways of specifying a linear subspace. Here is the first one.

(1) (By a basis.) We can specify C by a **generating matrix** G, which is a $k \times n$ matrix, $k := \dim(C)$, whose rows are vectors of some basis of C.

A generating matrix is very useful for encoding. When we need to transmit a vector $\mathbf{v} \in S^k$, we send the vector $\mathbf{w} := \mathbf{v}^T G \in C$.

We can always get a generating matrix in the form $G = (I_k \mid A)$ by choosing a suitable basis of the subspace C. Then the vector $\mathbf{w}$ agrees with $\mathbf{v}$ on the first k coordinates. It means that the encoding procedure adds $n - k$ extra symbols to the original message. (These are sometimes called *parity check bits*; which makes sense for the case $S = \mathbb{F}_2$—each such bit is a linear combination of some of the bits in the original message, and thus it “checks the parity” of these bits.) It is important to realize that the transmission channel makes no distinction between the original message and the parity check bits; errors can occur anywhere including the parity check bits.

The Hamming code is a linear code of length 7 over $\mathbb{F}_2$ and with a generating matrix

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Here is another way of specifying a linear code.

(2) (By linear equations) A linear code C can also be given as the set of all solutions of a system of linear equation of the form $P\mathbf{w} = \mathbf{0}$, where P is called a **parity check matrix** of the code C.

This way of presenting C is particularly useful for decoding, as we will see. If the generating matrix of C is $G = (I_k \mid A)$, then it is easy to check that $P := (-A^T \mid I_{n-k})$ is a parity check matrix of C.

Example: The generalized Hamming code. The Hamming code has a parity check matrix

$$P = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

The columns are exactly all possible non-zero vectors from $\mathbb{F}_2^3$. This construction can be generalized: We choose a parameter $\ell \geq 2$ and define a **generalized Hamming code** as the linear code over $\mathbb{F}_2$ of length $n := 2^\ell - 1$ whose parity check matrix P has ℓ rows, n columns and the columns are all non-zero vectors from $\mathbb{F}_2^\ell$.

Proposition. *The generalized Hamming code C has $d(C) = 3$, and thus it corrects 1 error.*

Proof. For showing that $d(C) \geq 3$, it suffices to verify that every nonzero $\mathbf{w} \in C$ has at least 3 nonzero entries. We thus need that $P\mathbf{w} = \mathbf{0}$ holds for no $\mathbf{w} \in \mathbb{F}_2^n$ with one or two 1's. For $\mathbf{w}$ with one 1 it would mean that P has a zero column, and for $\mathbf{w}$ with two 1's we would get an equality between two columns of P. Thus none of these possibilities occur. $\square$

Let us remark that the (generalized) Hamming code is optimal in the following sense: There exists no code $C \subseteq \mathbb{F}_2^{2^\ell - 1}$ with $d(C) \geq 3$ and with more words than the generalized Hamming code. We leave the proof as a (nontrivial) exercise.

Decoding a generalized Hamming code. We send a vector $\mathbf{w}$ of the generalized Hamming code and receive $\mathbf{w}'$. If at most one error has occurred, we have $\mathbf{w}' = \mathbf{w}$, or $\mathbf{w}' = \mathbf{w} + \mathbf{e}_i$ for some $i \in \{1, 2, \ldots, n\}$, where $\mathbf{e}_i$ has 1 at position i and 0's elsewhere.

Looking at the product $P\mathbf{w}'$, for $\mathbf{w}' = \mathbf{w}$ we have $P\mathbf{w}' = \mathbf{0}$, while for $\mathbf{w}' = \mathbf{w} + \mathbf{e}_i$ we get $P\mathbf{w}' = P\mathbf{w} + P\mathbf{e}_i = P\mathbf{e}_i$, which is the ith column of the matrix P. Hence, assuming that there was at most one error, we can immediately tell whether an error has occurred, and if it has, we can identify the position of the incorrect letter.

Sources. R. W. Hamming, *Error detecting and error correcting codes*, Bell System Tech. J. **29** (1950), 147–160.

As was mentioned above, error-correcting codes form a major area with numerous textbooks. A good starting point, although not to all tastes, can be

M. Sudan, *Coding theory: Tutorial & survey*, in Proc. 42nd Annual Symposium on Foundations of Computer Science (FOCS), 2001, 36–53, `http://people.csail.mit.edu/madhu/papers/focs01-tut.ps`.

Miniature 6

Odd Distances

Theorem. *There are no* 4 *points in the plane such that the distance between each pair is an odd integer.*

Proof. Let us suppose for contradiction that there exist 4 points with all the distances odd. We can assume that one of them is $\mathbf{0}$, and we call the three remaining ones $\mathbf{a}, \mathbf{b}, \mathbf{c}$. Then $\|\mathbf{a}\|$, $\|\mathbf{b}\|$, $\|\mathbf{c}\|$, $\|\mathbf{a}-\mathbf{b}\|$, $\|\mathbf{b}-\mathbf{c}\|$, and $\|\mathbf{c}-\mathbf{a}\|$ are odd integers, where $\|\mathbf{x}\|$ is the Euclidean length of a vector $\mathbf{x}$.

We observe that if m is an odd integer, then $m^2 \equiv 1 \pmod 8$ (here $\equiv$ denotes congruence; $x \equiv y \pmod k$ means that k divides $x-y$). Hence the squares of all the considered distances are congruent to 1 modulo 8. From the cosine theorem we also have $2\langle \mathbf{a}, \mathbf{b}\rangle = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - \|\mathbf{a}-\mathbf{b}\|^2 \equiv 1 \pmod 8$, and the same holds for $2\langle \mathbf{a}, \mathbf{c}\rangle$ and $2\langle \mathbf{b}, \mathbf{c}\rangle$. If B is the matrix

$$\begin{pmatrix} \langle \mathbf{a}, \mathbf{a}\rangle & \langle \mathbf{a}, \mathbf{b}\rangle & \langle \mathbf{a}, \mathbf{c}\rangle \\ \langle \mathbf{b}, \mathbf{a}\rangle & \langle \mathbf{b}, \mathbf{b}\rangle & \langle \mathbf{b}, \mathbf{c}\rangle \\ \langle \mathbf{c}, \mathbf{a}\rangle & \langle \mathbf{c}, \mathbf{b}\rangle & \langle \mathbf{c}, \mathbf{c}\rangle \end{pmatrix},$$

then $2B$ is congruent to the matrix

$$R := \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

modulo 8. Since $\det(R) = 4$, we have $\det(2B) \equiv 4 \pmod 8$. (To see this, we consider the expansion of both determinants according to the definition, and we note that the corresponding terms for $\det(2B)$ and for $\det(R)$ are congruent modulo 8.) Thus $\det(2B) \neq 0$, and so $\det(B) \neq 0$. Hence, $\operatorname{rank}(B) = 3$.

On the other hand, $B = A^T A$, where

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}.$$

We have $\operatorname{rank}(A) \leq 2$ and, as it is well known, the rank of a product of matrices is no larger than the minimum of the ranks of the factors. Thus, $\operatorname{rank}(B) \leq 2$, and this contradiction concludes the proof. □

Sources. The result is from

R. L. Graham, B. L. Rothschild, and E. G. Straus, *Are there $n + 2$ points in E^n with pairwise odd integral distances?*, Amer. Math. Monthly **81** (1974), 21–25.

I've heard the proof above from Moshe Rosenfeld.

Miniature 7

Are These Distances Euclidean?

Can we find three points $\mathbf{p}, \mathbf{q}, \mathbf{r}$ in the plane whose mutual Euclidean distances are all 1's? Of course we can—the vertices of an equilateral triangle.

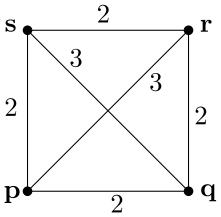
Can we find $\mathbf{p}, \mathbf{q}, \mathbf{r}$ with $\|\mathbf{p}-\mathbf{q}\| = \|\mathbf{q}-\mathbf{r}\| = 1$ and $\|\mathbf{p}-\mathbf{r}\| = 3$? No, since the direct path from $\mathbf{p}$ to $\mathbf{r}$ can't be longer than the path via $\mathbf{q}$; these distances violate the **triangle inequality**, which in the Euclidean case tells us that

$$\|\mathbf{p}-\mathbf{r}\| \leq \|\mathbf{p}-\mathbf{q}\| + \|\mathbf{q}-\mathbf{r}\|$$

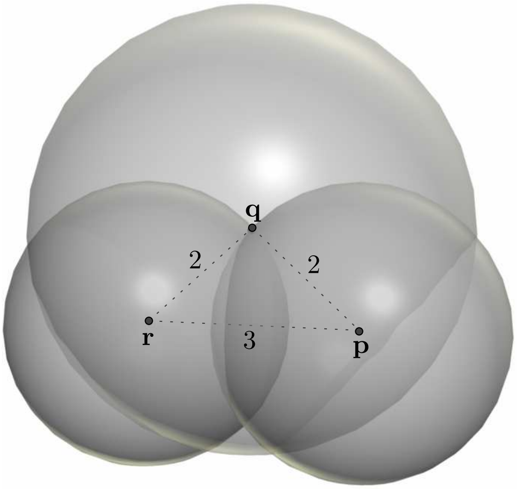
for every three points $\mathbf{p}, \mathbf{q}, \mathbf{r}$ (in any Euclidean space).

It turns out that the triangle inequality is the *only* obstacle for three points: Whenever nonnegative real numbers x, y, z satisfy $x \leq y + z$, $y \leq x + z$, and $z \leq x + y$, then there are $\mathbf{p}, \mathbf{q}, \mathbf{r} \in \mathbb{R}^2$ such that $\|\mathbf{p}-\mathbf{q}\| = x$, $\|\mathbf{q}-\mathbf{r}\| = y$, and $\|\mathbf{p}-\mathbf{r}\| = z$. These are well known conditions for the existence of a triangle with given side lengths.

What about prescribing distances for four points? We need to look for points in $\mathbb{R}^3$; that is, we ask for a tetrahedron with given side lengths. Here the triangle inequality is a necessary, but not sufficient condition. For example, if we require the distances as indicated in the picture,



then there is no violation of a triangle inequality, yet there are no corresponding $\mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s} \in \mathbb{R}^3$. Geometrically, if we construct the triangle $\mathbf{pqr}$, then the spheres around $\mathbf{p}, \mathbf{q}, \mathbf{r}$ that would have to contain $\mathbf{s}$ have no common intersection:



This is a rather ad-hoc argument. Linear algebra provides a very elegant characterization of the systems of numbers that can appear as Euclidean distances, using the notion of a *positive semidefinite matrix*. The characterization works for any number of points; if there are $n+1$ points, we want them to live in $\mathbb{R}^n$. The formulation becomes more convenient to state if we number the desired points starting from 0:

Theorem. *Let m_{ij}, $i, j = 0, 1, \ldots, n$, be nonnegative real numbers with $m_{ij} = m_{ji}$ for all i, j and $m_{ii} = 0$ for all i. Then points $\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{R}^n$ with $\|\mathbf{p}_i - \mathbf{p}_j\| = m_{ij}$ for all i, j exist if and* *only if the* $n \times n$ *matrix* G *with*

$$g_{ij} = \frac{1}{2}\left(m_{0i}^2 + m_{0j}^2 - m_{ij}^2\right)$$

is positive semidefinite.

Let us note that the triangle inequality doesn't appear explicitly in the theorem—it is hidden in the condition of positive semidefiniteness (you may want to check this for the case $n = 2$).

The proof of the theorem relies on the following characterization of positive semidefinite matrices.

Fact. *An real symmetric* $n \times n$ *matrix* A *is positive semidefinite if and only if there exists an* $n \times n$ *real matrix* X *such that* $A = X^T X$.

Reminder of a proof. If $A = X^T X$, then for every $\mathbf{x} \in \mathbb{R}^n$ we have $\mathbf{x}^T A\mathbf{x} = (X\mathbf{x})^T(X\mathbf{x}) = \|X\mathbf{x}\|^2 \geq 0$, and so A is positive semidefinite.

Conversely, every real symmetric square matrix A is **diagonalizable**, i.e., it can be written as $A = T^{-1}DT$ for a nonsingular $n \times n$ matrix T and a diagonal matrix D (with the eigenvalues of A on the diagonal). Moreover, an inductive proof of this fact even yields T **orthogonal**, i.e., such that $T^{-1} = T^T$ and thus $A = T^T DT$. Then we can set $X := RT$, where $R = \sqrt{D}$ is the diagonal matrix having the square roots of the eigenvalues of A on the diagonal; here we use the fact that A, being positive semidefinite, has all eigenvalues nonnegative.

It turns out that one can even require X to be upper triangular, and in such case one speaks about a *Cholesky factorization* of A. □

Proof of the theorem. First we check necessity: If $\mathbf{p}_0, \ldots, \mathbf{p}_n$ are given points in $\mathbb{R}^n$ and $m_{ij} := \|\mathbf{p}_i - \mathbf{p}_j\|$, then G as in the theorem is positive semidefinite.

For this, we need the *cosine theorem*, with tells us that $\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle$ for any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Thus, if we define $\mathbf{x}_i := \mathbf{p}_i - \mathbf{p}_0$, $i = 1, 2, \ldots, n$, we get that $\langle \mathbf{x}_i, \mathbf{x}_j\rangle = \frac{1}{2}(\|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2 - \|\mathbf{x}_i - \mathbf{x}_j\|^2) = g_{ij}$. So G is the **Gram matrix** of the vectors $\mathbf{x}_i$, we can write $G = X^T X$, and hence G is positive semidefinite.

Conversely, if G is positive semidefinite, we can decompose it as $G = X^T X$ for some $n \times n$ real matrix X. Then we let $\mathbf{p}_i \in \mathbb{R}^n$ be the ith column of X for $i = 1, 2, \ldots, n$, while $\mathbf{p}_0 := \mathbf{0}$. Reversing the above calculation, we arrive at $\|\mathbf{p}_i - \mathbf{p}_j\| = m_{ij}$, and the proof is finished. □

The theorem solves the question for points living in $\mathbb{R}^n$, the largest dimension one may ever need for $n+1$ points. One can also ask when the desired points can live in $\mathbb{R}^d$ with some given d, say $d = 2$. An extension of the above argument shows that the answer is positive if and only if $G = X^T X$ for some matrix X of rank at most d.

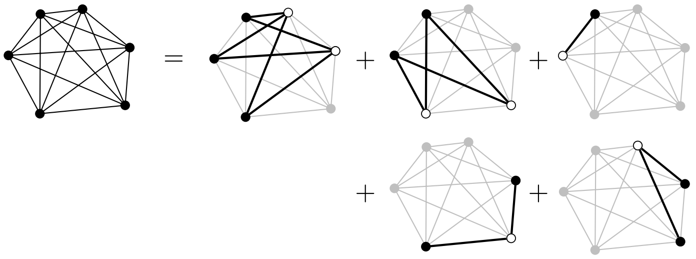
Source. I. J. Schoenberg: *Remarks to Maurice Fréchet's Article "Sur La Definition Axiomatique D'Une Classe D'Espace Distances Vectoriellement Applicable Sur L'Espace De Hilbert"*, Ann. Math. 2nd Ser. **36** (1935), 724–732.

Miniature 8

Packing Complete Bipartite Graphs

We want to decompose the edge set a complete graph, say K_6, into edge sets of complete bipartite subgraphs, and use as few subgraphs as possible. We recall that a graph G is **complete bipartite** if its vertex set $V(G)$ can be partitioned into two disjoint subsets A, B so that $E(G) = \{\{a, b\} : a \in A, b \in B\}$. Such A and B are called the **color classes**of G.

Here is one such decomposition, using 5 complete bipartite subgraphs:



There are several other possible decompositions using 5 complete bipartite subgraphs, and in general, it is not hard to find a decomposition of K_n using $n-1$ complete bipartite subgraphs. But can one do better?

This problem was motivated by a question in telecommunications. We present a neat linear-algebraic proof of the following:

Theorem. *If the set $E(K_n)$, i.e., the set of the edges of a complete graph on n vertices, is expressed as a disjoint union of edge sets of m complete bipartite graphs, then $m \geq n-1$.*

Proof. Suppose that complete bipartite graphs $H_1, H_2, \ldots, H_m$ disjointly cover all edges of K_n. Let X_k and Y_k be the color classes of H_k. (The set $V(H_k) = X_k \cup Y_k$ is not necessarily all of $V(K_n)$.)

We assign an $n \times n$ matrix A_k to each graph H_k. The entry of A_k in the ith row and jth column is

$$a_{ij}^{(k)} = \begin{cases} 1 & \text{if } i \in X_k \text{ and } j \in Y_k \\ 0 & \text{otherwise.} \end{cases}$$

We claim that each of the matrices A_k has rank 1. This is because all the nonzero rows of A_k are equal to the same vector, namely, the vector with 1's at positions whose indices belong to Y_k and with 0's elsewhere.

Let us now consider the matrix $A = A_1 + A_2 + \cdots + A_m$. The rank of a sum of two matrices is never larger than the sum of their ranks (why?), and thus the rank of A is at most m. It is enough to prove that this rank is also at least $n-1$.

Each edge $\{i, j\}$ belongs to exactly one of the graphs H_k, and hence for each $i \neq j$, we have either $a_{ij} = 1$ and $a_{ji} = 0$, or $a_{ij} = 0$ and $a_{ji} = 1$, where a_{ij} is the entry of the matrix A at position (i, j). We also have $a_{ii} = 0$. From this we get $A + A^T = J_n - I_n$, where I_n is the identity matrix and J_n denotes the $n \times n$ matrix having 1's everywhere.

For contradiction, let us assume that $\text{rank}(A) \leq n-2$. If we add an extra row consisting of all 1's to A, the resulting $(n+1) \times n$ matrix still has rank at most $n-1$, and hence there exists a nontrivial linear combination of its columns equal to $\mathbf{0}$. In other words, there exists a (column) vector $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$, such that $A\mathbf{x} = \mathbf{0}$ and $\sum_{i=1}^n x_i = 0$.

From the last equality we get $J_n\mathbf{x} = \mathbf{0}$. We calculate

$$\begin{aligned}\mathbf{x}^T\left(A + A^T\right)\mathbf{x} &= \mathbf{x}^T(J_n - I_n)\mathbf{x} = \mathbf{x}^T(J_n\mathbf{x}) - \mathbf{x}^T(I_n\mathbf{x}) = \\ &= 0 - \mathbf{x}^T\mathbf{x} = -\sum_{i=1}^{n} x_i^2 < 0.\end{aligned}$$

On the other hand, we have

$$\mathbf{x}^T\left(A^T + A\right)\mathbf{x} = \left(\mathbf{x}^T A^T\right)\mathbf{x} + \mathbf{x}^T(A\mathbf{x}) = \mathbf{0}^T\mathbf{x} + \mathbf{x}^T\mathbf{0} = 0,$$

and this is a contradiction. □

Sources. The result is due to

R. L. Graham and H. O. Pollak, *On the addressing problem for loop switching*, Bell System Tech. J. **50** (1971), 2495–2519.

The proof is essentially that of

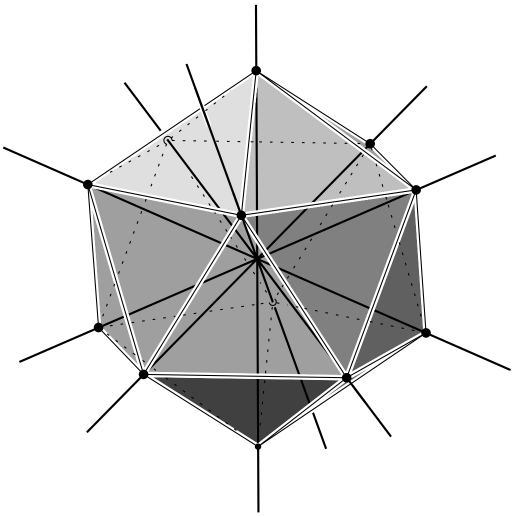
H. Tverberg, *On the decomposition of K_n into complete bipartite graphs*, J. Graph Theory **6,4** (1982), 493–494.

Miniature 9

Equiangular Lines

What is the largest number of lines in $\mathbb{R}^3$ such that the angle between every two of them is the same?

Everybody knows that in $\mathbb{R}^3$ there cannot be more than three mutually orthogonal lines, but the situation for angles other than 90 degrees is more complicated. For example, the six longest diagonals of the regular icosahedron (connecting pairs of opposite vertices) are equiangular:



As we will prove, this is the largest number one can get.

Theorem. *The largest number of equiangular lines in* $\mathbb{R}^3$ *is* 6*, and in general, there cannot be more that* $\binom{d+1}{2}$ *equiangular lines in* $\mathbb{R}^d$*.*

Proof. Let us consider a configuration of n lines, where each pair has the same angle $\vartheta \in (0, \frac{\pi}{2}]$. Let $\mathbf{v}_i$ be a unit vector in the direction of the ith line (we choose one of the two possible orientations of $\mathbf{v}_i$ arbitrarily). The condition of equal angles is equivalent to

$$|\langle \mathbf{v}_i, \mathbf{v}_j \rangle| = \cos\vartheta, \qquad \text{for all } i \neq j.$$

Let us regard $\mathbf{v}_i$ as a column vector, or a $d\times 1$ matrix. Then $\mathbf{v}_i^T \mathbf{v}_j$ is the scalar product $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$, or more precisely, the 1×1 matrix whose only entry is $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$. On the other hand, $\mathbf{v}_i \mathbf{v}_j^T$ is a $d\times d$ matrix.

We show that the matrices $\mathbf{v}_i \mathbf{v}_i^T$, $i = 1, 2, \ldots, n$, are linearly independent. Since they are the elements of the vector space of all real symmetric $d\times d$ matrices, and the dimension of this space is $\binom{d+1}{2}$, we get $n \leq \binom{d+1}{2}$, just as we wanted.

To check linear independence, we consider a linear combination

$$\sum_{i=1}^{n} a_i \mathbf{v}_i \mathbf{v}_i^T = 0,$$

where $a_1, a_2, \ldots, a_n$ are some coefficients. We multiply both sides of this equality by $\mathbf{v}_j^T$ from the left and by $\mathbf{v}_j$ from the right. Using the associativity of matrix multiplication, we obtain

$$0 = \sum_{i=1}^{n} a_i \mathbf{v}_j^T (\mathbf{v}_i \mathbf{v}_i^T) \mathbf{v}_j = \sum_{i=1}^{n} a_i \langle \mathbf{v}_i, \mathbf{v}_j \rangle^2 = a_j + \sum_{i \neq j} a_i \cos^2\vartheta$$

for all i, j. In other words, we have deduced that $M\mathbf{a} = \mathbf{0}$, where $\mathbf{a} = (a_1, \ldots, a_n)$ and $M = (1 - \cos^2\vartheta) I_n + (\cos^2\vartheta) J_n$. Here I_n is the unit matrix and J_n is the matrix of all 1's. It is easy to check that the matrix M is nonsingular (using $\cos\vartheta \neq 1$); for example, as in Miniature 4, we can show that M is positive definide. Therefore, $\mathbf{a} = \mathbf{0}$, the matrices $\mathbf{v}_i \mathbf{v}_i^T$ are linearly independent, and the theorem is proved. $\square$

Remark. While the upper bound of this theorem is tight for $d = 3$, for some larger values of d it can be improved by other methods. The best possible value is not known in general. The best known lower bound (from the year 2000) is $\frac{2}{9}(d+1)^2$, holding for all numbers d of the form $3 \cdot 2^{2t-1} - 1$, where t is a natural number.

Sources. The theorem is stated in

P. W. H. Lehmmens and J. J. Seidel, *Equiangular lines*, J. of Algebra **24** (1973), 494–512.

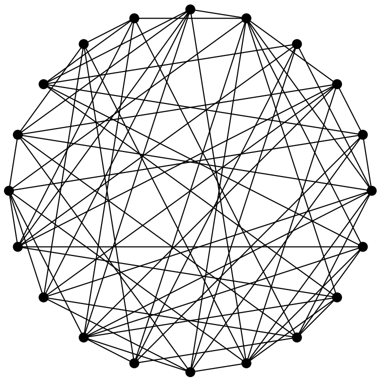
and attributed to Gerzon (private communication). The best upper bound mentioned above is from

D. de Caen, *Large equiangular sets of lines in Euclidean space*, Electr. J. Comb. **7** (2000), R55.

Miniature 10

Where is the Triangle?

Does a given graph contain a **triangle**, i.e., three vertices u, v, w, every two of them connected by an edge? This question is not entirely easy to answer for graphs with many vertices and edges. For example, where is a triangle in this graph?



An obvious algorithm for finding a triangle inspects every triple of vertices, and thus it needs roughly n^3 operations for an n-vertex graph (there are $\binom{n}{3}$ triples to look at, and $\binom{n}{3}$ is approximately $n^3/6$ for large n). Is there a significantly faster method?

There is, but surprisingly, the only known approach for breaking the n^3 barrier is algebraic, based on fast matrix multiplication.

To explain it, we assume for notational convenience that the vertex set of the given graph G is $\{1, 2, \ldots, n\}$, and we define the **adjacency matrix** of G as the $n \times n$ matrix A with

$$a_{ij} = \begin{cases} 1 & \text{if } i \neq j \text{ and } \{i, j\} \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

The key insight is to understand the square $B := A^2$. By the definition of matrix multiplication we have $b_{ij} = \sum_{k=1}^{n} a_{ik} a_{kj}$, and

$$a_{ik} a_{kj} = \begin{cases} 1 & \text{if the vertex } k \text{ is adjacent to both } i \text{ and } j, \\ 0 & \text{otherwise.} \end{cases}$$

So b_{ij} counts the number of common neighbors of i and j.

Finding a triangle is equivalent to finding two adjacent vertices i, j with a common neighbor k. So we look for two indices i, j such that both $a_{ij} \neq 0$ and $b_{ij} \neq 0$.

To do this, we need to compute the matrix $B = A^2$. If we perform the matrix multiplication according to the definition, we need about n^3 arithmetic operations and thus we save nothing compared to the naive method of inspecting all triples of vertices.

However, ingenious algorithms are known that multiply $n \times n$ matrices asymptotically faster. The oldest one, due to Strassen, needs roughly $n^{2.807}$ arithmetic operations. It is based on a simple but very clever trick—if you haven't seen it, it is worth looking it up (Wikipedia?).

The **exponent of matrix multiplication** is defined as the infimum of numbers ω for which there exists an algorithm that multiplies two square matrices using $O(n^\omega)$ operations. Its value is unknown (the common belief is that it equals 2); the current best upper bound is roughly 2.376.

Many computational problems are known where fast matrix multiplication brings asymptotic speedup. Finding triangles is among the simplest of them, we will meet and several other, more sophisticated algorithms of this kind appear later.

Remarks. The described method for finding triangles is the fastest known for *dense* graphs, i.e., graphs that have relatively many edges compared to the number of vertices. Another nice algorithm, which we won't discuss here, can detect a triangle in time $O(m^{2\omega/(\omega+1)})$, where m is the number of edges.

One can try to use similar methods for detecting subgraphs other than the triangle; there is an extensive literature concerning this problem. For example, a cycle of length 4 can be detected in time $O(n^2)$, much faster than a triangle!

Sources. A. Itai and M. Rodeh, *Finding a minimum circuit in a graph*, SIAM J. Comput., **7,4** (1978), 413–423.

Among the numerous papers dealing with fast detection of a fixed subgraph in a given graph, we mention

T. Kloks, D. Kratsch, and H. Müller, *Finding and counting small induced subgraphs efficiently*, Inform. Process. Lett. **74,3–4** (2000), 115–121,

which can be used as a starting point for further explorations of the topic.

The first "fast" matrix multiplication algorithm is due to

V. Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969), 354–356.

The asymptotically fastest known matrix multiplication algorithm is from

D. Coppersmith and S. Winograd, *Matrix multiplication via arithmetic progressions*, J. Symbolic Computation **9** (1990), 251–280.

An interesting new method, which provides similarly fast algorithms in a different way, appeared in

H. Cohn, R. Kleinberg, B. Szegedy, and C. Umans, *Group-theoretic algorithms for matrix multiplication*, in Proc. 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS), 2005, 379–388.

Miniature 11

Checking Matrix Multiplication

Multiplying two $n \times n$ matrices is a very important operation. A straightforward algorithm requires about n^3 arithmetic operations, but as was mentioned in Miniature 10, ingenious algorithms have been discovered that are asymptotically much faster. The current record is an $O(n^{2.376})$ algorithm. However, the constant of proportionality is so astronomically large that the algorithm is interesting only theoretically. Indeed, matrices for which it would prevail over the straightforward algorithm can't fit into any existing or future computer.

But progress cannot be stopped and soon a software company may start selling a program called MATRIX WIZARD that, supposedly, multiplies matrices real fast. Since wrong results could be disastrous for you, you would like to have a simple *checking program* appended to MATRIX WIZARD that would always check whether the resulting matrix C is really the product of the input matrices A and B.

Of course, a checking program that actually multiplies A and B and compares the result with C makes little sense, since you do not know how to multiply matrices as fast as MATRIX WIZARD. But it turns out that if we allow for some slight probability of error in the checking, there is a very simple and efficient checker for matrix multiplication.

We assume that the considered matrices consist of rational numbers, although everything works without change for matrices over any field. The checking algorithm receives $n \times n$ matrices A, B, C as the input. Using a random number generator, it picks a random n-component vector $\mathbf{x}$ of zeros and ones. More precisely, each vector in $\{0,1\}^n$ appears with the same probability, equal to 2^{-n}. The algorithm computes the products $C\mathbf{x}$ (using $O(n^2)$ operations) and $AB\mathbf{x}$ (again with $O(n^2)$ operations; the right parenthesizing is, of course, $A(B\mathbf{x})$). If the results agree, the algorithm answers YES, and otherwise, it answers NO.

If $C = AB$, the algorithm always answers YES, which is correct. But if $C \neq AB$, it can answer both YES and NO. We claim that the wrong answer YES has probability at most $\frac{1}{2}$, and thus the algorithm detects a wrong matrix multiplication with probability at least $\frac{1}{2}$.

Let us set $D := C - AB$. It suffices to show that if D is any nonzero $n \times n$ matrix and $\mathbf{x} \in \{0,1\}^n$ is random, then the vector $\mathbf{y} := D\mathbf{x}$ is zero with probability at most $\frac{1}{2}$.

Let us fix indices i, j such that $d_{ij} \neq 0$. We will derive that then the probability of $y_i = 0$ is at most $\frac{1}{2}$.

We have

$$y_i = d_{i1}x_1 + d_{i2}x_2 + \cdots + d_{in}x_n = d_{ij}x_j + S,$$

where

$$S = \sum_{\substack{k=1,2,\ldots,n \\ k \neq j}} d_{ik}x_k.$$

Imagine that we choose the values of the entries of $\mathbf{x}$ according to successive coin tosses and that the toss deciding the value of x_j is made as the last one (since the tosses are independent it doesn't matter).

Before this last toss, the quantity S has already been fixed, because it doesn't depend on x_j. After the last toss, we have $x_j = 0$ with probability $\frac{1}{2}$ and $x_j = 1$ with probability $\frac{1}{2}$. In the first case, we have $y_i = S$, while in the second case, $y_i = S + d_{ij} \neq S$. Therefore, $y_i \neq 0$ in at least one of these two cases, and so $D\mathbf{x} \neq \mathbf{0}$ has probability at least $\frac{1}{2}$ as claimed.

The described checking algorithm is fast but not very reliable: It may fail to detect an error with probability as high as $\frac{1}{2}$. But if we repeat it, say, fifty times for a single input A, B, C, it fails to detect an error with probability at most $2^{-50} < 10^{-15}$, and this probability is totally negligible for practical purposes.

Remark. The idea of *probabilistic checking* of computations, which we have presented here in a simple form, turned out to be very fruitful. The so called PCP theorem from the theory of computational complexity shows that for any effectively solvable computational problem, it is possible to check the solution probabilistically in a very short time. A slow personal computer can, in principle, check the work of the most powerful supercomputers. Furthermore, a surprising connections of these results to approximation algorithms have been discovered.

Sources. R. Freivalds, *Probabilistic machines can use less running time*, in Information processing 77, IFIP Congr. Ser. **7**, North-Holland, Amsterdam, 1977, 839–842.

For introduction to PCP and computational complexity see, e.g.,

O. Goldreich, *Computational complexity: A conceptual perspective*, Cambridge University Press, Cambridge, 2008.

Miniature 12

Tiling a Rectangle by Squares

Theorem. *A rectangle R with side lengths 1 and x, where x is irrational, cannot be "tiled" by finitely many squares (so that the squares have disjoint interiors and cover all of R).*

Proof. For contradiction, let us assume that a tiling exists, consisting of squares $Q_1, Q_2, \ldots, Q_n$, and let s_i be the side length of Q_i.

We need to consider the set $\mathbb{R}$ of all real numbers as a vector space over the field $\mathbb{Q}$ of rationals. This is a rather strange, infinite-dimensional vector space, but a very useful one.

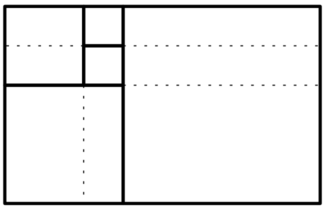
Let $V \subseteq \mathbb{R}$ be the linear subspace generated by the numbers x and $s_1, s_2, \ldots, s_n$, in other words, the set of all rational linear combinations of these numbers.

We define a linear mapping $f\colon V \to \mathbb{R}$ such that $f(1) = 1$ and $f(x) = -1$ (and otherwise arbitrarily). This is possible, because 1 and x are linearly independent over $\mathbb{Q}$. Indeed, there is a basis $(b_1, b_2, \ldots, b_k)$ of V with $b_1 = 1$ and $b_2 = x$, and we can set, e.g., $f(b_1) = 1$, $f(b_2) = -1$, $f(b_3) = \cdots = f(b_k) = 0$, and extend f linearly on V.

For each rectangle A with edges a and b, where $a, b \in V$, we define a number $v(A) := f(a)f(b)$.

We claim that if the $1 \times x$ reclangle R is tiled by the squares $Q_1, Q_2, \ldots, Q_n$, then $v(R) = \sum_{i=1}^{n} v(Q_i)$. This leads to a contradiction, since $v(R) = f(1)f(x) = -1$, while $v(Q_i) = f(s_i)^2 \geq 0$ for all i.

To check the claim just made, we extend the edges of all squares Q_i of the hypothetical tiling across the whole of R, as is indicated in the picture:



This partitions R into small rectangles, and using the linearity of f, it is easy to see that $v(R)$ equals to the sum of $v(B)$ over all these small rectangles B. Similarly $v(Q_i)$ equals the sum of $v(B)$ over all the small rectangles lying inside Q_i. Thus, $v(R) = \sum_{i=1}^{n} v(Q_i)$. $\square$

Remark. It turns out that a rectangle can be tiled by squares if and only if the ratio of its sides is rational. Various other theorems about the impossibility of tilings can be proved by similar methods. For example, it is impossible to dissect the cube into finitely many convex pieces that can be rearranged so that they tile a regular tetrahedron.

Sources. The theorem is a special case of a result from

> M. Dehn, *Über Zerlegung von Rechtecken in Rechtecke*, Math. Ann. **57,3** (1903), 314–332.

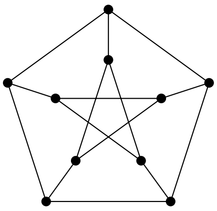
Unfortunately, so far I haven't found the source of the above proof. Another very beautiful proof follows from a remarkable connection of square tilings to planar electrical networks:

> R. L. Brooks, C. A. B. Smith, A. H. Stone, and W. T. Tutte, *The dissection of rectangles into squares*, Duke Math. J. **7** (1940), 312–340.

Miniature 13

Three Petersens Are Not Enough

The famous **Petersen graph**



has 10 vertices of degree 3. The complete graph K_{10} has 10 vertices of degree 9. Yet it is not possible to cover all edges of K_{10} by three copies of the Petersen graph.

Theorem. *There are no three subgraphs of* K_{10}*, each isomorphic to the Petersen graph, that together cover all edges of* K_{10}*.*

The theorem can obviously be proved by an extensive case analysis. The following elegant proof is a little sample of a part of graph theory dealing with properties of the eigenvalues of the adjacency matrix of a graph.

Proof. We recall that the **adjacency matrix** of a graph G on the vertex set $\{1, 2, \ldots, n\}$ is the $n \times n$ matrix A with

$$a_{ij} = \begin{cases} 1 & \text{if } i \neq j \text{ and } \{i, j\} \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

It means that the adjacency matrix of the graph K_{10} is $J_{10} - I_{10}$, where J_n is the $n \times n$ matrix of all 1's and I_n is the identity matrix.

Let us assume that the edges of K_{10} are covered by subgraphs P, Q and R, each of them isomorphic to the Petersen graph. If A_P is the adjacency matrix of P, and similarly for A_Q and A_R, then $A_P + A_Q + A_R = J_{10} - I_{10}$.

It is easy to check that the adjacency matrices of two isomorphic graphs have the same set of eigenvalues, and also the same dimensions of the corresponding eigenspaces.

We can use the Gaussian elimination to calculate that for the adjacency matrix of Petersen graph, the eigenspace corresponding to the eigenvalue 1 has dimension 5; i.e., the matrix $A_P - I_{10}$ has a 5-dimensional kernel.

Moreover, this matrix has exactly three 1's and one -1 in every column. So if we sum all the equations of the system $(A_P - I_{10})\mathbf{x} = \mathbf{0}$, we get $2x_1 + 2x_2 + \cdots + 2x_{10} = 0$. In other words, the kernel of $A_P - I_{10}$ is contained in the 9-dimensional orthogonal complement of the vector $\mathbf{1} = (1, 1, \ldots, 1)$.

The same is true for the kernel of $A_Q - I_{10}$, and therefore, the two kernels have a common non-zero vector $\mathbf{x}$. We know that $J_{10}\mathbf{x} = \mathbf{0}$ (since $\mathbf{x}$ is orthogonal to $\mathbf{1}$), and we calculate

$$\begin{aligned} A_R\mathbf{x} &= (J_{10} - I_{10} - A_P - A_Q)\mathbf{x} \\ &= J_{10}\mathbf{x} - I_{10}\mathbf{x} - (A_P - I_{10})\mathbf{x} - (A_Q - I_{10})\mathbf{x} - 2I_{10}\mathbf{x} \\ &= \mathbf{0} - \mathbf{x} - \mathbf{0} - \mathbf{0} - 2\mathbf{x} = -3\mathbf{x}. \end{aligned}$$

It means that -3 must be an eigenvalue of A_R, but it is not an eigenvalue of the adjacency matrix of the Petersen graph—a contradiction. □

Source. O. P. Lossers and A. J. Schwenk, *Solution of advanced problem 6434*, Am. Math. Monthly **94** (1987), 885–887.

Miniature 14

Petersen, Hoffman–Singleton, and Maybe 57

This is a classical piece from the 1960s, reproduced many times, but still one of the most beautiful applications of graph eigenvalues I've seen. Moreover, the proof nicely illustrates the general flavor of algebraic nonexistence proofs for various "highly regular" structures.

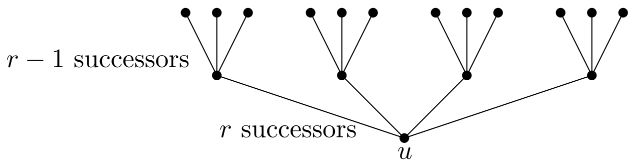
Let G be a graph of girth $g \geq 4$ and minimum degree $r \geq 3$, where the **girth** of G is the length of its shortest cycle, and **minimum degree** r means that every vertex has at least r neighbors. It is not obvious that such graphs exist for all r and g, but it is known that they do.

Let $n(r, g)$ denote the smallest possible number of vertices of such a G. Determining this quantity, at least approximately, belongs to the most fascinating problems in graph theory, whose solution would probably have numerous interesting consequences.

A lower bound. A lower bound for $n(r, g)$ is obtained by a simple "branching" argument (linear algebra comes later). First let us assume that $g = 2k + 1$ is odd.

Let G be a graph of girth g and minimum degree r. Let us fix a vertex u in G and consider two paths of length k in G starting at u.

For some time they may run together, then they branch off, and they never meet again past the branching point—otherwise, they would close a cycle of length at most $2k$. Thus, G has a subgraph as in the following picture:



(the picture is for $r = 4$ and $k = 2$). It is a tree T of height k, with branching degree r at the root and $r - 1$ at the other inner vertices. (In G, we may have additional edges connecting some of the leaves at the topmost level, and of course, G may have more vertices than T.)

It is easy to count that the number of vertices of T equals

$$1 + r + r(r-1) + r(r-1)^2 + \cdots + r(r-1)^{k-1}, \tag{1}$$

and this is the promised lower bound for $n(r, 2k + 1)$. For $g = 2k$ even, a similar but slightly more complicated argument, which we omit here, yields the lower bound

$$1 + r + r(r-1) + \cdots + r(r-1)^{k-2} + (r-1)^{k-1} \tag{2}$$

for $n(r, 2k)$.

Upper bounds. For large r and g, the state of knowledge about $n(r, g)$ is unsatisfactory: The best known upper bounds are roughly the $\frac{3}{2}$-th *power* of the lower bounds (1), (2), and so there is uncertainty already in the exponent.

Still, (1), (2) remain essentially the best known lower bounds for $n(r, g)$, and a considerable attention has been paid to graphs for which these bounds are attained, since they are highly regular and usually have many remarkable properties. For historical reasons they are called **Moore graphs** for odd g and **generalized polygons**[1] for even g.

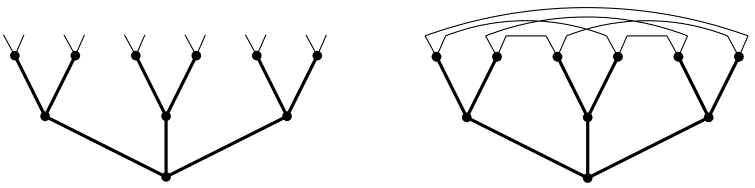
---

[1] In some sources, though, the term Moore graph is used for both the odd-girth and even-girth cases.

Moore graphs. Here we will consider only Moore graphs (for information on generalized polygons and the known exact values of $n(r, g)$ we refer, e.g., to G. Royle's web page `http://people.csse.uwa.edu.au/gordon/cages/`). Explicitly, a Moore graph is a graph of girth $2k+1$, minimum degree r, and with $1+r+r(r-1)+\cdots+r(r-1)^{k-1}$ vertices.

To avoid trivial cases we assume $r \geq 3$ and $k \geq 2$. We also note that *every* vertex in a Moore graph must have degree exactly r, for if there were a vertex of larger degree, we could take it as u in the lower bound argument and show that the number of vertices is larger than (1).

The question of whether a Moore graph exists for given k and r can be cast as a kind of "connecting puzzle." The vertex set must coincide with the vertex set of the tree T in the lower-bound argument, and the additional edges besides those of T may connect only the leaves of T. Thus, we draw T, add $r-1$ "paws" to each leaf, and then we want to connect the paws by edges so that no cycle shorter than $2k+1$ arises. The picture illustrates this for girth $2k+1 = 5$ and $r = 3$:



In this case the puzzle has a solution depicted on the right. The solution, which can be shown to be unique up to isomorphism, is the famous **Petersen graph**, whose more usual picture was shown in Miniature 13.

The only other known Moore graph has 50 vertices, girth 5, and degree $r = 7$. It is obtained by gluing together many copies of the Petersen graph in a highly symmetric fashion, and it is called the **Hoffman–Singleton graph**. Surprisingly, it is known that this very short list exhausts all Moore graphs, with a single possible exception:

The existence of a Moore graph of girth 5 and degree 57 has been neither proved nor disproved.

Here we give the proof that Moore graphs of girth 5 can't have degree other than $3, 7, 57$. The nonexistence of Moore graphs of higher girth is proved by somewhat similar methods.

Theorem. *If a graph G of girth 5 with minimum degree $r \geq 3$ and with $n = 1 + r + (r-1)r = r^2 + 1$ vertices exists, then $r \in \{3, 7, 57\}$.*

We begin the proof of the theorem by a graph-theoretic argument, which is a simple consequence of the argument used above for deriving (1), specialized to $k = 2$.

Lemma. *If G is a graph as in the theorem, then every two non-adjacent vertices have exactly one common neighbor.*

Proof. If u, v are two arbitrary non-adjacent vertices, we let u play the u in the argument leading to (1). The tree T has height 2 in our case, and so v is necessarily a leaf of T and there is a unique path of length 2 connecting it with u. □

Proof of the theorem. We recall the notion of **adjacency matrix** A of G, already used in Miniatures 10 and 13. Assuming that the vertex set of G is $\{1, 2, \ldots, n\}$, A is the $n \times n$ matrix with entries given by

$$a_{ij} = \begin{cases} 1 & \text{for } i \neq j \text{ and } \{i,j\} \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

The key step in the proof is to consider $B := A^2$. As was already mentioned in Miniature 10, from the definition of matrix multiplication one can easily see that for an arbitrary graph G, b_{ij} is the number of vertices k adjacent to both of the vertices i and j. So for $i \neq j$, b_{ij} is the number of common neighbors of i and j, while b_{ii} is simply the degree of i.

Specializing these general facts to a G as in the theorem, we obtain

$$b_{ij} = \begin{cases} r & \text{for } i = j, \\ 0 & \text{for } i \neq j \text{ and } \{i,j\} \in E(G), \\ 1 & \text{for } i \neq j \text{ and } \{i,j\} \notin E(G). \end{cases} \tag{3}$$

Indeed, the first case states that all degrees are r, the second one that two adjacent vertices in G have no common neighbor (since G has girth 5 and thus contains no triangles), and the third case restates the assertion of the lemma that every two non-adjacent vertices have exactly one common neighbor.

Next, we rewrite (3) in a matrix form:

$$A^2 = rI_n + J_n - I_n - A, \tag{4}$$

where I_n is the identity matrix and J_n is the matrix of all 1's.

Now we enter the business of graph eigenvalues. The usual opening move in this area is to recall the following from linear algebra: Every symmetric real $n \times n$ matrix A has n mutually orthogonal eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, and the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all real (and not necessarily distinct).

If A is the adjacency matrix of a graph with all degrees r, then $A\mathbf{1} = r\mathbf{1}$, with $\mathbf{1}$ standing for the vector of all 1's. Hence r is an eigenvalue with eigenvector $\mathbf{1}$, and we can thus assume $\lambda_1 = r$, $\mathbf{v}_1 = \mathbf{1}$. Then by the orthogonality of the eigenvectors, for all $i \neq 1$ we have $\mathbf{1}^T\mathbf{v}_i = 0$, and thus also $J_n\mathbf{v}_i = \mathbf{0}$.

Armed with these facts, let us see what happens if we multiply (4) by some $\mathbf{v}_i$, $i \neq 1$, from the right. The left-hand side becomes $A^2\mathbf{v}_i = A\lambda_i\mathbf{v}_i = \lambda_i^2\mathbf{v}_i$, while the right-hand side yields $r\mathbf{v}_i - \mathbf{v}_i - \lambda_i\mathbf{v}_i$. Both sides are scalar multiples of the nonzero vector $\mathbf{v}_i$, and so the scalar multipliers must be the same, which leads to

$$\lambda_i^2 + \lambda_i - (r-1) = 0.$$

Thus, each λ_i, $i \neq 1$ equals one of the roots ρ_1, ρ_2 of the quadratic equation $\lambda^2 + \lambda - (r-1) = 0$, which gives

$$\rho_{1,2} := (-1 \pm \sqrt{D})/2, \quad \text{with } D := 4r - 3.$$

Hence A has only 3 distinct eigenvalues: r, ρ_1, and ρ_2. Let us suppose that ρ_1 occurs m_1 times among the λ_i and ρ_2 occurs m_2 times; since r occurs once, we have $m_1 + m_2 = n - 1$.

The last linear algebra fact we need is that the sum of all eigenvalues of A equals its **trace**, i.e., the sum of all diagonal elements, which in our case is 0. Hence

$$r + m_1\rho_1 + m_2\rho_2 = 0. \tag{5}$$

The rest of the proof is pure calculation plus a simple divisibility consideration (a bit of number theory if we wanted to sound fancy). After substituting in (5) for ρ_1 and ρ_2, multiplying by 2, and using $m_1 + m_2 = n - 1 = r^2$ (the last equality is one of the assumptions of the theorem), we arrive at

$$(m_1 - m_2)\sqrt{D} = r^2 - 2r. \tag{6}$$

If D is not a square of a natural number, then $\sqrt{D}$ is irrational, and (6) can hold only if $m_1 = m_2$. But then $r^2 - 2r = 0$, which cannot happen for $r \geq 3$. Therefore, we have $D = 4r - 3 = s^2$ for a natural number s. Expressing $r = (s^2 + 3)/4$, substituting into (6), and simplifying leads to

$$s^4 - 2s^2 - 16(m_1 - m_2)s = s\left(s^3 - 2s - 16(m_1 - m_2)\right) = 15.$$

Hence 15 is a multiple of the positive integer s, and so $s \in \{1, 3, 5, 15\}$, corresponding to $r \in \{1, 3, 7, 57\}$. The theorem is proved. $\square$

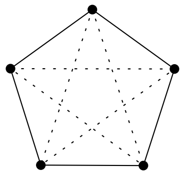
Source. A. J. Hoffman and R. R. Singleton, *On Moore graphs with diameters* 2 *and* 3, IBM J. Res. Develop. **4** (1960), 497–504.

Miniature 15

Only Two Distances

What is the largest number of points in the plane such that every two of them have the same distance? If we have at least three, they must form the vertices of an equilateral triangle, and there is no way of adding a fourth point.

How many points in the plane can we have if we allow the distances to attain *two* different values? We can easily find a 4-point configuration with only two distances, for example, the vertices of a square. A bit of thinking reveals even a 5-point configuration:



But how can one prove that there are no larger configurations?

We can ask a similar question in a higher dimension, that is, in the space $\mathbb{R}^d$, $d \geq 3$: What is the maximum number $n = n(d)$ such that there are n points in $\mathbb{R}^d$ with only two distances? The following elegant method gives a rather good upper bound for $n(d)$, even though the result for the plane is not really breathtaking (we get an upper bound of 9 instead of the correct number 5).

Theorem. $n(d) \leq \frac{1}{2}(d^2 + 5d + 4)$.

Proof. Let $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ be points in $\mathbb{R}^d$. Let $\|\mathbf{p}_i - \mathbf{p}_j\|$ be the Euclidean distance of $\mathbf{p}_i$ from $\mathbf{p}_j$. We have

$$\|\mathbf{p}_i - \mathbf{p}_j\|^2 = (p_{i1} - p_{j1})^2 + (p_{i2} - p_{j2})^2 + \cdots + (p_{id} - p_{jd})^2,$$

where p_{ij} is the jth coordinate of the point $\mathbf{p}_i$. We suppose that $\|\mathbf{p}_i - \mathbf{p}_j\| \in \{a, b\}$ for every $i \neq j$.

With each point $\mathbf{p}_i$ we associate a carefully chosen function

$$f_i \colon \mathbb{R}^d \to \mathbb{R},$$

defined by

$$f_i(\mathbf{x}) := \left(\|\mathbf{x} - \mathbf{p}_i\|^2 - a^2\right)\left(\|\mathbf{x} - \mathbf{p}_i\|^2 - b^2\right),$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$.

The key property of these functions is

$$f_i(\mathbf{p}_j) = \begin{cases} 0 & \text{for } i \neq j, \\ a^2b^2 \neq 0 & \text{for } i = j, \end{cases} \tag{7}$$

which follows immediately from the two-distance assumption.

Let us consider the vector space of all real functions $\mathbb{R}^d \to \mathbb{R}$, and the linear subspace V spanned by the functions $f_1, f_2, \ldots, f_n$. First, we claim that $f_1, f_2, \ldots, f_n$ are linearly independent. Let us assume that a linear combination $f = \alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_n f_n$ equals 0; i.e., it is the zero function $\mathbb{R}^d \to \mathbb{R}$. In particular, it is 0 at each $\mathbf{p}_i$. According to (7) we get $0 = f(\mathbf{p}_i) = \alpha_i a^2 b^2$, and therefore, $\alpha_i = 0$ for every i. Thus $\dim V = n$.

Now we want to find a (preferably small) system G of functions $\mathbb{R}^d \to \mathbb{R}$, not necessarily belonging to V, that generates V; that is, every $f \in V$ is a linear combination of functions in G. Then we will have the bound $|G| \geq \dim V = n$.

Each of the f_i is a polynomial in the variables $x_1, x_2, \ldots, x_d$ of degree at most 4, and so it is a a linear combination of monomials in $x_1, x_2, \ldots, x_d$ of degree at most 4. It is easy to count that there are $\binom{d+4}{4}$ such monomials, and this gives a generating system G with $|G| = \binom{d+4}{4}$.

Next, proceeding more carefully, we will find a still smaller G. We express $\|\mathbf{x}-\mathbf{p}_i\|^2 = \sum_{j=1}^d (x_j - p_{ij})^2 = X - \sum_{j=1}^d 2x_j p_{ij} + P_i$, where $X := \sum_{j=1}^d x_j^2$ and $P_i := \sum_{j=1}^d p_{ij}^2$. Then we have

$$\begin{aligned} f_i(\mathbf{x}) &= \left(\|\mathbf{x}-\mathbf{p}_i\|^2 - a^2\right)\left(\|\mathbf{x}-\mathbf{p}_i\|^2 - b^2\right) = \\ &= \left(X - \sum_{j=1}^d 2x_j p_{ij} + A_i\right)\left(X - \sum_{j=1}^d 2x_j p_{ij} + B_i\right), \end{aligned}$$

where $A_i := P_i - a^2$ and $B_i := P_i - b^2$. By another rearrangement we get

$$\begin{aligned} f_i(\mathbf{x}) &= X^2 - 4X\sum_{j=1}^d p_{ij}x_j + \left(\sum_{j=1}^d 2p_{ij}x_j\right)^2 + \\ &\qquad + (A_i + B_i)\left(X - \sum_{j=1}^d 2p_{ij}x_j\right) + A_iB_i. \end{aligned}$$

From this we can see that each f_i is a linear combination of functions in the following system G:

$$\begin{array}{ll} X^2, & \\ x_jX, & j = 1,2,\ldots,d, \\ x_j^2, & j = 1,2,\ldots,d, \\ x_ix_j, & 1 \le i < j \le d, \\ x_j, & j = 1,2,\ldots,d, \text{ and} \\ 1. & \end{array}$$

(Let us remark that X itself is a linear combination of the x_j^2.) We have $|G| = 1 + d + d + \binom{d}{2} + d + 1 = \frac{1}{2}(d^2 + 5d + 4)$, and so $n \le \frac{1}{2}(d^2 + 5d + 4)$. The theorem is proved. $\square$

Remark. The upper bound in the theorem can be improved to $\binom{d+2}{2}$ by additional tricks, which we don't consider here. However, it has the right order of magnitude, as the following example shows. We start with the $\binom{d}{2}$ points in $\mathbb{R}^d$ with two coordinates equal to 1 and the remaining 0. This is a two-distance set, and it lies in the hyperplane $\sum_{i=1}^d x_i = 2$. Thus we can place it into $\mathbb{R}^{d-1}$ as well, and that gives a lower bound $n(d) \ge \binom{d+1}{2} = \frac{1}{2}(d^2 + d)$.

Sources. D. G. Larman, C. A. Rogers, and J. J. Seidel, *On two-distance sets in Euclidean space*, Bull. London Math. Soc. **9,3** (1977), 261–267.

According to Babai and Frankl, a similar trick first appears in

T. H. Koornwinder, *A note on the absolute bound for systems of lines*, Indag. Math. **38,2** (1976), 152–153.

The improved upper bound of $\binom{d+2}{2}$ is due to

A. Blokhuis, *A new upper bound for the cardinality of 2-distance sets in Euclidean space*, Ann. Discrete Math. **20** (1984), 65–66.

Miniature 16

Covering a Cube Minus One Vertex

We consider the set $\{0,1\}^d \subset \mathbb{R}^d$ of the vertices of the d-dimensional unit cube, and we want to cover all of these vertices except for one, say $\mathbf{0} = (0, 0, \ldots, 0)$, by hyperplanes. (We recall that a **hyperplane** in $\mathbb{R}^d$ is a set of the form $\{\mathbf{x} \in \mathbb{R}^d : a_1x_1 + \cdots + a_dx_d = b\}$ for some coefficients $a_1, \ldots, a_d, b \in \mathbb{R}$ with at least one a_i nonzero.)

Of course, we can cover all the vertices using only two hyperplanes, say $\{x_1 = 0\}$ and $\{x_1 = 1\}$, but the problem becomes interesting if none of the covering hyperplanes may contain the point $\mathbf{0}$. What is the smallest possible number of hyperplanes under these conditions?

One easily finds (at least) two different ways of covering with d hyperplanes. We can use the hyperplanes $\{x_i = 1\}$, $i = 1, 2, \ldots, d$, or the hyperplanes $\{x_1 + x_2 + \cdots + x_d = k\}$, $k = 1, 2, \ldots, d$. As we will see, d is the smallest possible number.

Theorem. *Let $h_1, \ldots, h_m$ be hyperplanes in $\mathbb{R}^d$ not passing through the origin that cover all points of the set $\{0,1\}^d$ except for the origin. Then $m \geq d$.*

Proof. Let h_i be defined by the equation $a_{i1}x_1+a_{i2}x_2+\cdots+a_{id}x_d = b$. We consider the following, cleverly chosen, polynomial

$$f(x_1,x_2,\ldots,x_d) = \prod_{i=1}^{m}\Bigl(1-\sum_{j=1}^{d} a_{ij}x_j\Bigr) - \prod_{j=1}^{d}(1-x_j).$$

It is constructed so that $f(\mathbf{x}) = 0$ for all $\mathbf{x} = (x_1,\ldots,x_n) \in \{0,1\}^n$ (to check this, one needs to distinguish the cases $\mathbf{x} = \mathbf{0}$ and $\mathbf{x} \neq \mathbf{0}$, and to use the assumptions of the theorem).

For contradiction, let us suppose that $m < d$. Then the degree of f is d, and the only monomial of degree d with a non-zero coefficient is $x_1x_2\cdots x_d$. It remains to show that this monomial is not a linear combination of monomials of lower degrees, when we consider these monomials as real functions on $\{0,1\}^d$.

First we observe that on $\{0,1\}^d$ we have $x_i = x_i^2$, and therefore, every polynomial is equivalent to a linear combination of multilinear monomials $x_I = \prod_{i\in I} x_i$, where $I \subseteq \{1,2,\ldots,d\}$. So it suffices to prove that such x_I are linearly independent.

Let us consider a linear combination

$$\sum_{I\subseteq\{1,2,\ldots,d\}} \alpha_I x_I = 0. \tag{8}$$

Assuming that there is an $\alpha_I \neq 0$, we take a minimum such I, minimum in the sense that $\alpha_J = 0$ for every $J \subset I$. Let us substitute $x_i = 1$ for $i \in I$ and $x_i = 0$ for $i \notin I$ into (8). Then we get $\alpha_I = 0$—a contradiction. ☐

Source. N. Alon and Z. Füredi, *Covering the cube by affine hyperplanes*, European J. Combin. **14,2** (1993), 79–83.

Miniature 17

Medium-Size Intersection Is Hard To Avoid

An extensive branch of combinatorics, *extremal set theory*, investigates problems of the following kind: Suppose that $\mathcal{F}$ is a system of subsets of an n-element set, and suppose that certain simply described configuration of sets doesn't occur in $\mathcal{F}$. What is the maximum possible number of sets in $\mathcal{F}$?

Here is a short list of famous examples:

- The **Sperner lemma** (one of the Sperner lemmas, that is): If there are no two distinct sets $A, B \in \mathcal{F}$ with $A \subset B$, then $|\mathcal{F}| \le \binom{n}{\lfloor n/2 \rfloor}$.
- The **Erdős–Ko–Rado "sunflower theorem"**: If $k \le n/2$, each $A \in \mathcal{F}$ has exactly k elements, and $A \cap B \ne \emptyset$ for every two $A, B \in \mathcal{F}$, then $|\mathcal{F}| \le \binom{n-1}{k-1}$.
- The Oddtown theorem from Miniature 3: If each $A \in \mathcal{F}$ has an odd number of elements and $|A \cap B|$ is even for every two distinct $A, B \in \mathcal{F}$, then $|\mathcal{F}| \le n$.

Such a list of theorems could be extended over many pages, and linear algebra methods constitute one of the main tools for proving them.

Here we present a strong and perhaps surprising result of this kind. It has a beautiful geometric application, which we explain in Miniature 18 below.

Theorem. *Let p be a prime number and let $\mathcal{F}$ be a system of $(2p-1)$-element subsets of an n-element set X such that no two sets in $\mathcal{F}$ intersect in precisely $p-1$ elements. Then the number of sets in $\mathcal{F}$ is at most*

$$|\mathcal{F}| \le \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{p-1}.$$

There are $\binom{n}{2p-1}$ subsets of an n-element set altogether. The theorem tells us that if we forbid one single intersection size, namely $p-1$, we can have only much fewer sets. The following is a possible quantification of "much fewer."

Corollary. *Let $\mathcal{F}$ be as in the theorem with $n = 4p$. Then*

$$\frac{\binom{4p}{2p-1}}{|\mathcal{F}|} \ge 1.1^n.$$

Proof of the corollary. First of all, $\binom{n}{k-1} = \frac{k}{n-k}\binom{n}{k}$, and so for $n \ge 4k$ we have $\binom{n}{k-1} \le \frac{1}{3}\binom{n}{k}$. So

$$\binom{4p}{p-1}+\binom{4p}{p-2}+\cdots+\binom{4p}{0} \le \binom{4p}{p}\Big(\frac{1}{3}+\frac{1}{3^2}+\frac{1}{3^3}+\cdots\Big) \le \frac{1}{2}\binom{4p}{p}.$$

Then

$$\begin{aligned}\frac{\binom{4p}{2p-1}}{|\mathcal{F}|} &\ge 2\cdot\frac{\binom{4p}{2p-1}}{\binom{4p}{p}} = 2\cdot\frac{(3p)(3p-1)\cdots(2p+2)}{(2p-1)(2p-2)\cdots(p+1)}\\ &\ge 2\left(\frac{3}{2}\right)^{p-1} > \left(\frac{3}{2}\right)^{p} > 1.1^n.\end{aligned}$$

(There are many other ways of doing this kind of calculation, and one can use well-known estimates such as $(n/k)^k \le \binom{n}{k} \le (en/k)^k$ or Stirling's formula, but we didn't want to assume those here.) □

Proof of the theorem. The proof combines tricks we have already met in Miniatures 15 and 16.

To each set $A \in \mathcal{F}$ we assign two things:

- A vector $\mathbf{c}_A \in \{0,1\}^n$. This is simply the **characteristic vector** of A, whose ith component is 1 if $i \in A$ and 0 otherwise.
- A function $f_A\colon \{0,1\}^n \to \mathbb{F}_p$, given by

$$f_A(\mathbf{x}) := \prod_{s=0}^{p-2}\Big(\Big(\sum_{i\in A} x_i\Big) - s\Big).$$

All the arithmetic operations in the definition of f_A are in the finite field $\mathbb{F}_p$, i.e., modulo p (and thus 0 and 1 are also treated as elements of $\mathbb{F}_p$). For example, for $p = 3$, $n = 5$, and $A = \{2,3\}$ we have $\mathbf{c}_A = (0,1,1,0,0)$ and $f_A(\mathbf{x}) = (x_2 + x_3)(x_2 + x_3 - 1)$.

We consider the set of all functions from $\{0,1\}^d$ to $\mathbb{F}_p$ as a vector space over $\mathbb{F}_p$ in the usual way, and we let $V_{\mathcal{F}}$ be the subspace spanned in it by the functions f_A, $A \in \mathcal{F}$.

First we check that the f_A's are linearly independent, and hence $\dim(V_{\mathcal{F}}) = |\mathcal{F}|$.

We have $f_A(\mathbf{c}_B) = \prod_{s=0}^{p-2}(|A\cap B| - s) \pmod p$, and this product is nonzero exactly if $|A \cap B| \equiv p-1 \pmod p$. For $A = B$ we have $|A \cap A| = 2p-1 \equiv p-1 \pmod p$, and so $f_A(\mathbf{c}_A) \neq 0$. For $A \neq B$ we have $|A\cap B| \le 2p-2$ and $|A \cap B| \neq p-1$ by the "omitted intersection" assumption, thus $|A\cap B| \not\equiv p-1 \pmod p$, and consequently $f_A(\mathbf{c}_B) = 0$.

Now a standard argument shows the independence of the f_A: Assuming $\sum_{A\in\mathcal{F}} \alpha_A f_A = 0$ for some coefficients $\alpha_A \in \mathbb{F}_p$, we substitute $\mathbf{c}_B$ into the left-hand side. All terms $f_A(\mathbf{c}_B)$ with $A \neq B$ vanish, and we are left with $\alpha_B f_B(\mathbf{c}_B) = 0$, which yields $\alpha_B = 0$ in view of $f_B(\mathbf{c}_B) \neq 0$. Since B was arbitrary, the f_A are linearly independent as claimed.

We proceed to bound $\dim(V_{\mathcal{F}})$ from above. In our concrete example above, we had $f_A(\mathbf{x}) = (x_2+x_3)(x_2+x_3-1)$, and multiplying out the parentheses we get $f_A(\mathbf{x}) = x_2^2 + x_3^2 + 2x_2x_3 - x_2 - x_3$. In general, each f_A is a polynomial in $x_1, x_2, \ldots, x_n$ of degree at most $p-1$, and hence it is a linear combination of monomials of the form $x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$, $i_1 + i_2 + \cdots + i_n \le p-1$.

We can still get rid of the monomials with some exponent i_j larger than 1, because x_j^2 and x_j represent the *same* function $\{0,1\}^n \to \mathbb{F}_p$ (we substitute only 0's and 1's for the variables). So it suffices to count the monomials with $i_j \in \{0,1\}$, and their number is the same as the number of all subsets of $\{1,2,\ldots,n\}$ of size at most $p-1$. Thus $\dim(V_{\mathcal{F}}) \leq \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{p-1}$, and the theorem follows. $\Box$

Sources. The theorem is a special case of the Frankl–Wilson inequality from

P. Frankl and R. M. Wilson, *Intersection theorems with geometric consequences*, Combinatorica **1,4** (1981), 357–368.

The proof follows

N. Alon, L. Babai, and H. Suzuki, *Multilinear polynomials and Frankl–Ray-Chaudhuri–Wilson type intersection theorems*, J. Combin. Theory Ser. A **58,2** (1991), 165–180.

More general "omitted intersection" theorems were proved by different methods in

P. Frankl and V. Rödl, *Forbidden intersections*, Trans. Amer. Math. Soc. **300,1** (1987), 259–286.

Miniature 18

On the Difficulty of Reducing the Diameter

Exceptionally in this collection, the next result relies on a theorem proved earlier, in Miniature 17.

The **diameter** of a set $X \subseteq \mathbb{R}^d$ is defined as

$$\operatorname{diam}(X) := \sup\{\|\mathbf{x}-\mathbf{y}\| : \mathbf{x}, \mathbf{y} \in X\},$$

where $\|\mathbf{x}-\mathbf{y}\|$ stands for the Euclidean distance of $\mathbf{x}$ and $\mathbf{y}$. If X is finite, or more generally, closed and bounded, the supremum is always attained and we can simply say that the diameter is the largest distance among two points of X.

The question. The following was asked by Karol Borsuk in 1933:

> *Can every set $X \subset \mathbb{R}^d$ of finite diameter be partitioned into $d{+}1$ subsets $X_1, X_2, \ldots, X_{d+1}$ so that each X_i has diameter strictly smaller than X?*

Let us call a partition of a set $X \subset \mathbb{R}^d$ into subsets $X_1, X_2, \ldots, X_k$ with $\operatorname{diam}(X_i) < \operatorname{diam}(X)$ for all i a **diameter-reducing partition** of X into k parts.

It is easily seen that there are sets in $\mathbb{R}^d$ with no diameter-reducing partition into d parts. For example, let X consist of $d+1$ points, every two with distance 1 (in other words, the vertex set of a regular d-dimensional simplex—an explicit construction of such a set is sketched in Miniature 30). If we partition X into d parts, one of the parts contains at least two points and thus it has diameter 1, same as X. Borsuk in his 1933 paper proved, among others, that the d-dimensional ball has a diameter-reducing partition into $d+1$ parts (this is easy) but it has none into d parts (this isn't).

Up until 1993, it was widely believed that *every* $X \subset \mathbb{R}^d$ of finite diameter should have a diameter-reducing partition into $d+1$ parts, so widely that people started speaking about *Borsuk's conjecture* (although Borsuk didn't express any such belief in his paper).

Borsuk's question was often reformulated with the additional assumption that X is convex. It is easy to see that involves no loss of generality, since the diameter of a set is the same as the diameter of its convex hull.

Several partial results supporting (so-called) Borsuk's conjecture were proved over the years. It was proved for arbitrary sets X in dimensions 2 and 3, for all *smooth* convex sets in all dimensions (where smooth, roughly speaking, means "no sharp corners or edges"), and for some other special classes of convex sets.

The answer. As the reader might have already guessed or known, Borsuk's question has eventually been answered *negatively*. Let us begin with preliminary considerations, which are not really necessary for the proof but may be helpful for understanding it.

The first thing to understand is that the additional assumption of convexity, which seemingly simplifies the problem, happens to be a smoke-screen: The essence of Borsuk's question lies in finite sets.

A useful class of finite sets is obtained from finite set systems. For a system $\mathcal{F}$ of subsets of $\{1,2,\ldots,n\}$ let $X_{\mathcal{F}} \subset \mathbb{R}^n$ be the set of all characteristic vectors of sets in $\mathcal{F}$: $X_{\mathcal{F}} := \{\mathbf{c}_A : A \in \mathcal{F}\}$, where the ith component of $\mathbf{c}_A$ is 1 if $i \in A$ and 0 otherwise.

We will translate the result of Miniature 17 into the language of characteristic vectors and distances. We recall the corollary of the theorem in that miniature: If p is a prime, $n = 4p$, and $\mathcal{F}$ is a system of $(2p-1)$-element subsets of $\{1,2,\ldots,n\}$ such that $|A \cap B| \neq p-1$ for every $A, B \in \mathcal{F}$, then $\binom{n}{2p-1}/|\mathcal{F}| \geq 1.1^n$.

Let $\mathcal{A}$ be the system of *all* $(2p-1)$-element subsets of $\{1, 2, \ldots, n\}$ (so $|\mathcal{A}| = \binom{n}{2p-1}$). The statement in the previous paragraph implies the following:

(9) *If we partition the sets in $\mathcal{A}$ into fewer than 1.1^n classes, then at least one of the classes contains two sets with intersection of size exactly $p-1$.*

We observe that since all sets in $\mathcal{A}$ have the same size, the Euclidean distance of two characteristic vectors $\mathbf{c}_A, \mathbf{c}_B \in X_{\mathcal{A}}$ is determined by the size of the intersection $|A \cap B|$. Indeed, $\|\mathbf{c}_A - \mathbf{c}_B\|^2 = |A \setminus B| + |B \setminus A| = |A| + |B| - 2|A \cap B| = 2(2p-1) - 2|A \cap B|$. In particular, if $|A \cap B| = p - 1$, then $\|\mathbf{c}_A - \mathbf{c}_B\| = \sqrt{2p}$. So whenever the point set $X_{\mathcal{A}}$ is partitioned into fewer than 1.1^n subsets, one of the subsets contains two points $\mathbf{c}_A, \mathbf{c}_B$ with distance $\sqrt{2p}$.

This already sounds similar to Borsuk's question: It tells us that we can't get rid of the distance $\sqrt{2p}$ by partitioning $X_{\mathcal{A}}$ into fewer than exponentially many parts. The only problem is that $\sqrt{2p}$ is not the diameter of $X_{\mathcal{A}}$ but rather some smaller distance. We thus want to transform $X_{\mathcal{A}}$ into another set so that the pairs with distance $\sqrt{2p}$ in $X_{\mathcal{A}}$ become pairs realizing the diameter of the new set. Such a transformation is possible, but it raises the dimension: The resulting point set, which we denote by $Q_{\mathcal{A}}$, lies in dimension n^2.

This ends the preliminary discussion. We now proceed with a statement of the result and the actual proof.

Theorem. *For every prime p there exists a point set in $\mathbb{R}^{n^2}$, $n = 4p$, that has no diameter-reducing partition into fewer than 1.1^n parts. Consequently, the answer to Borsuk's question is no.*

Proof. First we need to recall the notion of **tensor product**[1] of vectors $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^n$: It is denoted by $\mathbf{x} \otimes \mathbf{y}$, and it is the vector in $\mathbb{R}^{mn}$ whose components are all the products $x_i y_j$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$. (Sometimes it is useful to think of $\mathbf{x} \otimes \mathbf{y}$ as the $m \times n$ matrix $\mathbf{x}\mathbf{y}^T$.)

---

[1] In linear algebra, the tensor product is defined more generally, for arbitrary two vector spaces. The definition given here can be regarded as the "standard" tensor product.

We will need the following identity involving the scalar product and the tensor product: For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\langle \mathbf{x} \otimes \mathbf{x}, \mathbf{y} \otimes \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle^2, \tag{10}$$

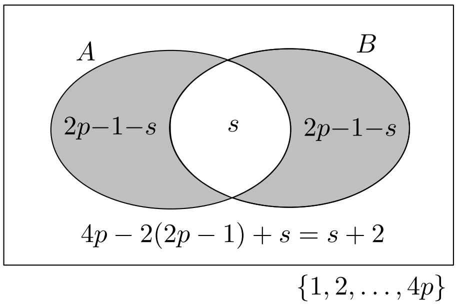
as is very easy to check.

Now we begin with the construction of the point set in the theorem. We recall that $\mathcal{A}$ consists of all $(2p-1)$-element subsets of $\{1, 2, \ldots, 4p\}$. For $A \in \mathcal{A}$, let $\mathbf{u}_A \in \{-1, 1\}^n$ be the *signed* characteristic vector of A whose ith component is is $+1$ if $i \in A$ and -1 otherwise. We set $\mathbf{q}_A := \mathbf{u}_A \otimes \mathbf{u}_A \in \mathbb{R}^{n^2}$, and the point set in the theorem is $Q_{\mathcal{A}} := \{\mathbf{q}_A : A \in \mathcal{A}\}$.

First we verify that for $A, B \in \mathcal{A}$ with $|A \cap B| = s$,

$$\langle \mathbf{u}_A, \mathbf{u}_B \rangle = 4(s - p + 1). \tag{11}$$

This can be checked using the following diagram, for instance:



Components in $(A \setminus B) \cup (B \setminus A)$ (gray) contribute -1 to the scalar product, and the remaining ones (white) contribute $+1$. Consequently, $\langle \mathbf{u}_A, \mathbf{u}_B \rangle = 0$ if and only if $|A \cap B| = p - 1$.

For the Euclidean distances in $Q_{\mathcal{A}}$ we have, using (10),

$$\begin{aligned} \|\mathbf{q}_A - \mathbf{q}_B\|^2 &= \langle \mathbf{q}_A, \mathbf{q}_A \rangle + \langle \mathbf{q}_B, \mathbf{q}_B \rangle - 2\langle \mathbf{q}_A, \mathbf{q}_B \rangle \\ &= \langle \mathbf{u}_A, \mathbf{u}_A \rangle^2 + \langle \mathbf{u}_B, \mathbf{u}_B \rangle^2 - 2\langle \mathbf{u}_A, \mathbf{u}_B \rangle^2. \end{aligned}$$

Now $\langle \mathbf{u}_A, \mathbf{u}_A \rangle^2 + \langle \mathbf{u}_B, \mathbf{u}_B \rangle^2$ is a number independent of A and B, and $\langle \mathbf{u}_A, \mathbf{u}_B \rangle^2$ is a nonnegative number that is 0 if and only if $|A \cap B| = p - 1$. Thus, the maximum possible distance of $\mathbf{q}_A$ and $\mathbf{q}_B$, equal to $\operatorname{diam}(Q_{\mathcal{A}})$, is attained exactly for $|A \cap B| = p - 1$. So by (9) $Q_{\mathcal{A}}$

has no diameter-reducing partition into fewer than 1.1^n parts as the theorem claims.

Finally, if we choose p sufficiently large so that $1.1^n > n^2 + 1$ and put $d := n^2$, we have a point set in $\mathbb{R}^d$ that has no diameter-reducing partition into $d + 1$ parts. □

What is the smallest dimension d for which Borsuk's question has a negative answer? Using only the statement of the theorem above, we get an upper bound of almost 10^4. Some improvement can be achieved by doing the calculations more precisely. At the time of writing, the best known upper bound is $d = 298$, whose proof involves additional ideas. It may still be quite far from the smallest possible value.

Sources. Borsuk's problem was stated in

> K. Borsuk, *Drei Sätze über die n-dimensionale euklidische Sphäre*, Fundamenta Mathematicae **20** (1933), 177–190.

The counterexample is from

> J. Kahn and G. Kalai, *A counterexample to Borsuk's conjecture*, Bull. Amer. Math. Soc. **29** (1993), 60–62.

The 298-dimensional counterexample is from

> A. Hinrichs and C. Richter, *New sets with large Borsuk numbers*, Disc. Math. **270** (2003), 137–147.

Miniature 19

The End of the Small Coins

An internet shop was processing m orders, each of them asking for various products. Suddenly, all coins with values below 1 Euro were taken out of circulation, and all prices had to be rounded, up or down, to whole Euros.

How can the shop round the prices so that the total price of each order is not affected by much? This rounding problem and similar questions is studied in *discrepancy theory*. Here we present a nice theorem with a linear-algebraic proof.

Theorem. *If at most t pieces of each product have been ordered in total, and if no order asks for more than one piece of each product, then it is possible to round the prices so that the total price of each order changes by no more than t Euros.*

It is interesting that the bound on the rounding error depends neither on the total number of orders nor on the number of different products.

A mathematical formulation of the problem. Let us call the products $1, 2, \ldots, n$ and let c_j be the price of the jth product. We can assume that each $c_j \in (0,1)$ (because only the rounding plays a role in the problem). No order contains more than one product of each kind, and so we can represent the ith order as a set $S_i \subseteq \{1, 2, \ldots, n\}$, $i = 1, 2, \ldots, m$. The theorem now asserts that if no j is in more than t sets, then there are numbers $z_1, z_2, \ldots, z_n \in \{0, 1\}$ such that

$$\left|\sum_{j\in S_i} c_j - \sum_{j\in S_i} z_j\right| \le t, \quad \text{for every } i = 1, 2, \ldots, m.$$

Proof. For every index $j \in \{1, 2, \ldots, n\}$ we introduce a real variable $x_j \in [0, 1]$, with initial value c_j. This variable will change during the proof and at the end, each x_j will have the value 0 or 1, which we will then use for z_j.

In each step, some of the variables x_j are already fixed, while the others are "floating". At the beginning, all the x_j are floating. The fixed x_j have values 0 or 1, and they won't change any more. The values of the floating variables are from the interval $(0, 1)$. In each step, at least one floating variable becomes fixed.

Let us call a set S_i **dangerous** if it contains more than t indices j for which x_j is still floating. The other sets are **safe**. We will keep the following condition satisfied:

$$\sum_{j\in S_i} x_j = \sum_{j\in S_i} c_j \text{ for all dangerous } S_i. \tag{12}$$

Let F be a set of indices of all floating variables and let us consider (12) as a system of linear equations with the floating variables as unknowns (while the values of the fixed variables are regarded as constants). This system surely has a solution—the current values of the floating variables. Since we assume that all floating variables lie in the interval $(0, 1)$, this solution is an interior point of the $|F|$-dimensional cube $[0, 1]^{|F|}$. We need to prove that there is a solution at the boundary of this cube as well, i.e., such that at least one of the variables attains value 0 or 1.

The crucial observation is that there are always fewer dangerous sets than floating variables, since every dangerous set needs more than t floating variables, while each floating variable contributes to at most t dangerous sets. Thus, the considered system of equations has fewer equations than unknowns, and so the solution space has dimension at least 1. Hence there is a straight line (a one-dimensional affine subspace) passing through the considered solution such that all points of this line are solutions, too. The line intersects the boundary of the cube at some point $\mathbf{y}$. We use the coordinates of $\mathbf{y}$ as the values of the floating variables in the next step, and all the floating variables x_j for which the corresponding value of $\mathbf{y}$ is 0 or 1 become fixed.

We repeat the described procedure until all variables become fixed. We claim that if we take the final value of x_j for z_j, $j = 1, 2, \ldots, n$, then $\left|\sum_{j \in S_i} c_j - \sum_{j \in S_i} z_j\right| \leq t$ for every $i = 1, 2, \ldots, m$ as we wanted.

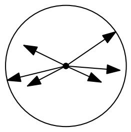
To see this, let us consider a set S_i. At the moment when it ceased to be dangerous, we still had $\sum_{j \in S_i} c_j - \sum_{j \in S_i} x_j = 0$ according to (12), and S_i contained the indices of at most t floating variables. The value of each of these floating variables has not changed by more than 1 in the rest of the process (it could have been 0.001 and later be fixed to 1). This finishes the proof. $\square$

Source. J. Beck and T. Fiala, *"Integer making" theorems*, Discr. Appl. Math **3** (1981), 1–8.

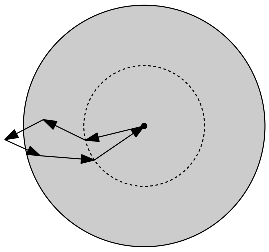
Miniature 20

Walking in the Yard

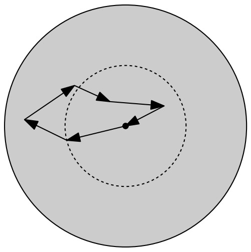
A mathematically inclined prison guard forces a prisoner to take a walk under the following strict instructions. The prisoner receives a finite set M of vectors, each of length at most 10 m. He must start the walk in the center of a circular prison yard of radius 20 m, then move by some vector $\mathbf{v}_1 \in M$, then by some other vector $\mathbf{v}_2 \in M$, etc., using each vector in M exactly once. The vectors in M sum up to $\mathbf{0}$, so the prisoner will again finish in the center. However, he must not cross the boundary of the yard any time during the walk (if he does, the guard will start shooting without warning).



the set M



a wrong order



a correct order

The following theorem shows that a safe walk is possible for every finite M, and it also works for yards that are d-dimensional balls.

Theorem. *Let M be an arbitrary set of n vectors in $\mathbb{R}^d$ such that $\|\mathbf{v}\| \le 1$ for every $\mathbf{v} \in M$, where the norm $\|\mathbf{v}\|$ of $\mathbf{v}$ is the usual*

Euclidean length, and $\sum_{\mathbf{v}\in M}\mathbf{v}=\mathbf{0}$. *Then it is possible to arrange all vectors of* M *into a sequence* $(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n)$ *in such a way that* $\|\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_k\|\le d$ *for every* $k=1,2,\ldots,n$.

In the example in the picture, the vectors can even be arranged so that the path lies within a circle of radius 1, but for an arbitrary set of vectors, radius 1 may be impossible (find an example). For the plane, the smallest possible radius of the yard is known: $\sqrt{5}/2\approx 1.118$. For a general dimension d, the best known lower bound is of order $\sqrt{d}$. It is not known whether the theorem can be improved to d-dimensional spherical yards of radius $O(\sqrt{d}\,)$.

The proof below actually yields a more general statement, for an arbitrary (not necessarily circular) yard: If $B\subset\mathbb{R}^d$ is a bounded convex set containing the origin, and M is a set of n vectors with $\mathbf{v}\in B$ for every $\mathbf{v}\in M$ and $\sum_{\mathbf{v}\in M}\mathbf{v}=0$, then there is an ordering $(\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n)$ of the vectors from M such that $\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_k\in dB$ for all $k=1,2,\ldots,n$, where $dB=\{d\mathbf{x}:\mathbf{x}\in B\}$. In this more general setting, the constant d *cannot* be improved. To see this for $d=2$, we take B as an equilateral triangle centered at the origin.

We start the proof with a simple general lemma (which was also implicitly used in Miniature 19).

Lemma. *Let* $A\mathbf{x}=\mathbf{b}$ *be a system of* m *linear equations in* $n\ge m$ *unknowns, and let us suppose that it has a solution* $\mathbf{x}_0\in[0,1]^n$. *Then there is a solution* $\tilde{\mathbf{x}}\in[0,1]^n$ *in which at least* $n-m$ *components are* 0*'s or* 1*'s.*

Proof. We proceed by induction on $n-m$. For $n=m$ there is nothing to prove, so let $n>m$. Then the solution space has dimension at least 1, and so it contains a line passing through $\mathbf{x}_0$. This line intersects the boundary of the cube $[0,1]^n$; let $\mathbf{y}$ be an intersection point. Thus, $y_i\in\{0,1\}$ for some index i.

Let us set up a new linear system with $n-1$ unknowns that is obtained from $A\mathbf{x}=\mathbf{b}$ by fixing the value of x_i to y_i. This new system satisfies the assumption of the lemma (a solution lying in $[0,1]^{n-1}$ is obtained from $\mathbf{y}$ by deleting y_i), and so, by the inductive assumption, it has a solution with at least $n-m-1$ components equal to 0 or 1.

Together with y_i this gives a solution of the original system with $n-m$ or more 0's and 1's. □

Proof of the theorem. The rough idea is this: The set M is "very good" because its vectors sum to $\mathbf{0}$, and thus the sum has norm 0. We introduce a weaker notion of a "good" set of vectors. The definition is chosen so that if K is a good set, then the sum of all of its vectors has norm at most d. Moreover, and this is the heart of the proof, we will show that every good set K of $k > d$ vectors has a good subset of $k-1$ vectors. This will allow us to find the desired ordering of the vectors of M by induction.

Here is the definition: A set $K = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ of $k \geq d$ vectors in $\mathbb{R}^d$, each of length at most 1, is called **good** if there exist coefficients $\alpha_1, \ldots, \alpha_k$ satisfying

$$\alpha_i \in [0,1], \quad i = 1, 2, \ldots, k$$

$$\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 + \cdots + \alpha_k\mathbf{w}_k = \mathbf{0} \tag{13}$$

$$\alpha_1 + \alpha_2 + \cdots + \alpha_k = k - d. \tag{14}$$

We note that if the right-hand side of (14) were k instead of $k-d$, then all the α_i would have to be 1 and thus (13) would simply mean $\sum_{i=1}^{k} \mathbf{w}_i = \mathbf{0}$. But since $\sum_{i=1}^{n} \alpha_i$ is $k-d$, most of the α_i must be close to 1, but there is some freedom left.

First let us check that if $K = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ is good, then $\|\mathbf{w}_1 + \mathbf{w}_2 + \cdots + \mathbf{w}_k\| \leq d$. Indeed, we have

$$\begin{aligned} \left\|\sum_{i=1}^{k} \mathbf{w}_i\right\| &= \left\|\sum_{i=1}^{k} \mathbf{w}_i - \sum_{i=1}^{k} \alpha_i\mathbf{w}_i\right\| \\ &\leq \sum_{i=1}^{k} \|(1-\alpha_i)\mathbf{w}_i\| = \sum_{i=1}^{k}(1-\alpha_i)\|\mathbf{w}_i\| \\ &\leq \sum_{i=1}^{k}(1-\alpha_i) = d. \end{aligned}$$

Next, we have the crucial claim.

Claim. *If $K = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k\}$ is a good set of $k > d$ vectors, then there is some i such that $K \setminus \{\mathbf{w}_i\}$ is a good set of $k-1$ vectors.*

Proof of the claim. We consider the following system of linear equations for unknowns $x_1, \dots, x_k$:

$$x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \cdots + x_k\mathbf{w}_k = \mathbf{0} \tag{15}$$

$$x_1 + x_2 + \cdots + x_k = k - d - 1. \tag{16}$$

Here (15) is an equality of two d-dimensional vectors and thus it actually represents d equations. The last equation (16) is like the condition (14), except that the right-hand side is $k-d-1$; this is a preparation for showing that a suitable subset of $k-1$ vectors in K is good.

The above system has $d+1$ equations for k unknowns. If $\alpha_1, \dots, \alpha_k$ are the coefficients witnessing that K is good, then by setting $x_i := \frac{k-d-1}{k-d}\alpha_i$ we obtain a solution of (15),(16) lying in $[0,1]^k$.

Thus by the lemma there is also a solution $\tilde{\mathbf{x}} \in [0,1]^k$ with at least $k-d-1$ components equal to 0 or 1. We want to see that at least one component of $\tilde{\mathbf{x}}$ has to be 0. Indeed, if all the $k-d-1$ components guaranteed by the lemma happen to be 1, then all of the remaining $d+1$ components must be 0, since all components add up to $k-d-1$ by (16).

Now it is easy to check that for any index i with $\tilde{x}_i = 0$ the set $K \setminus \{\mathbf{w}_i\}$ is good. Indeed, the remaining components of $\tilde{\mathbf{x}}$ can be used in the role of the α_i in the definition of a good set. This proves the claim.

The proof of the theorem is finished easily by induction. We start with the set $M_n := M$, which is obviously good. Using the claim, we find a vector in M_n whose removal produces a good set. We call this vector $\mathbf{v}_n$, and we let $M_{n-1} := M_n \setminus \{\mathbf{v}_n\}$. Similarly, having constructed the good set M_k, we find a vector $\mathbf{v}_k \in M_k$ such that $M_{k-1} := M_k \setminus \{\mathbf{v}_k\}$ is good, and so on, all the way down to M_d.

We are left with the set M_d of d vectors, and we number these arbitrarily $\mathbf{v}_1, \dots, \mathbf{v}_d$. For $k \le d$ we obviously have $\|\mathbf{v}_1 + \cdots + \mathbf{v}_k\| \le k \le d$, and for $k > d$ the norm of the sum of all vectors in M_k is at most d since M_k is a good set. The theorem is proved. □

Sources. The theorem is sometimes called the Steinitz lemma since Steinitz gave a first complete proof of a weaker version in 1913, following an incomplete proof of Lévy from 1905. The above proof is from

> V. S. Grinberg and S. V. Sevastyanov, *The value of the Steinitz constant* (in Russian), Funk. Anal. Prilozh. **14** (1980), 56–57.

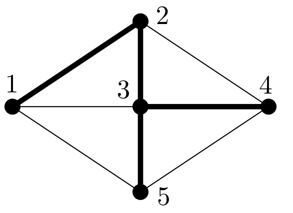
For background and several results of a similar nature see

> I. Bárány, *On the power of linear dependencies*, in Gy. O. H. Katona, M. Grötschel editors, *Building bridges*, Springer, Berlin 2008, 31–46.

Miniature 21

Counting Spanning Trees

A **spanning tree** of a graph G is a connected subgraph of G that has the same vertex set as G and contains no cycles. The next picture shows a 5-vertex graph with one of the possible spanning trees marked thick.



What is the number $\kappa(G)$ of spanning trees of a given graph G? Here is the answer:

Theorem (Matrix-tree theorem). *Let G be a graph on the vertex set $\{1, 2, \ldots, n\}$, and let L be the* **Laplace matrix** *of G, i.e., the $n \times n$ matrix whose entry ℓ_{ij} is given by*

$$\ell_{ij} := \begin{cases} \deg(i) & \textit{if } i = j, \\ -1 & \textit{if } \{i, j\} \in E(G), \\ 0 & \textit{otherwise}, \end{cases}$$

where $\deg(i)$ *is the number of neighbors (degree) of the vertex* i *in* G. *Let* L^- *be the* $(n-1)\times(n-1)$ *matrix obtained by deleting the last row and last column of* L. *Then*

$$\kappa(G) = \det(L^-).$$

For example, for the G in the picture we have

$$L = \begin{pmatrix} 3 & -1 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 & 0 \\ -1 & -1 & 4 & -1 & -1 \\ 0 & -1 & -1 & 3 & -1 \\ -1 & 0 & -1 & -1 & 3 \end{pmatrix}, \quad L^- = \begin{pmatrix} 3 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ 0 & -1 & -1 & 3 \end{pmatrix},$$

and $\det(L^-) = 45$ (can you check the number of spanning trees directly?).

I still remember my amazement when I saw the matrix-tree theorem for the first time. I believe it remains one of the most impressive uses of determinants. It is rather well known, but the forthcoming proof, hopefully, is not among those presented most often, and moreover, it resembles the proof of the *Gessel–Viennot lemma*, which is a powerful general tool in enumeration.

Proof. We begin with the usual expansion of $\det(L^-)$ according to the definition of a determinant as a sum over all permutations of $\{1, 2, \ldots, n-1\}$:

$$\det(L^-) = \sum_{\pi} \operatorname{sgn}(\pi) \prod_{i=1}^{n-1} \ell_{i,\pi(i)}. \tag{17}$$

Here $\operatorname{sgn}(\pi)$ is the sign of the permutation π, which can be defined as $(-1)^t$, where t is an integer such that one can obtain π from the identity permutation by t transpositions.

We now write each diagonal entry ℓ_{ii} of L^- in (17) as a sum of 1's, e.g., instead of 3 we write $(1+1+1)$. Then we multiply out the parentheses, so that each of the products in (17) is further expanded as a sum of products, where the factors in the products are only 1's and -1's. Let us call the resulting sum the **superexpansion** of $\det(L^-)$.

Graphically, each nonzero term in the superexpansion is obtained by selecting one 1 or -1 in each row and in each column of L^-. One of such selections is marked by circling the selected items:

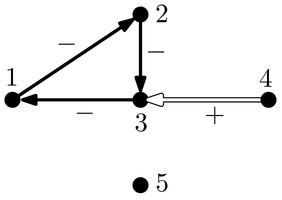
$$\begin{pmatrix} 1+1+1 & \enclose{circle}{-1} & -1 & 0 \\ -1 & 1+1+1 & \enclose{circle}{-1} & -1 \\ \enclose{circle}{-1} & -1 & 1+1+1+1 & -1 \\ 0 & -1 & -1 & 1+\enclose{circle}{1}+1 \end{pmatrix}.$$

The sign of such a term is $(-1)^m \operatorname{sgn}(\pi)$, where m is the number of -1's factors and π is the corresponding permutation. In the example, $m = 3$ and $\pi = (2, 3, 1, 4)$, with sign $+1$, so the term contributes a -1 to the superexpansion.

Next, we associate a combinatorial object with each term in the superexpansion. The object is a *directed graph* (or *digraph* for short) on the vertex set $\{1, 2, \ldots, n\}$, and moreover, each directed edge is either *positive* or *negative*. The rules for creating this signed digraph are as follows:

- If there is a circled -1 in row i and column j, make a *negative* directed edge from i to j.
- If the kth "1" in the diagonal entry ℓ_{ii} is circled, make a *positive* directed edge from i to the kth smallest neighbor of i in G (the vertices of G are numbered, so we can talk about the kth smallest neighbor).

For the term shown by the circles above, we thus obtain the following signed digraph (negative edges are shown black and positive edges white):



Let $\mathcal{D}$ denote the set of all signed digraphs D obtained in this way from the terms of the superexpansion. It is easy to see that each

$D \in \mathcal{D}$ comes from exactly one term of the superexpansion. We can thus talk about $\mathrm{sgn}(D)$, meaning the sign of the corresponding term, and write π_D for the associated permutation.

We divide $\mathcal{D}$ into three parts as follows:

- $\mathcal{T}$, the $D \in \mathcal{D}$ with no directed cycle.
- $\mathcal{D}^+$, the $D \in \mathcal{D}$ with $\mathrm{sgn}(D) = +1$ and at least one directed cycle.
- $\mathcal{D}^-$, the $D \in \mathcal{D}$ with $\mathrm{sgn}(D) = -1$ and at least one directed cycle.

Here is a plan for the rest of the proof. We will show that the "acyclic objects" in $\mathcal{T}$ all have positive signs and they are in one-to-one correspondence with the spanning trees of G—thus they count what we want. Then, by constructing a suitable bijection, we will prove that $|\mathcal{D}^+| = |\mathcal{D}^-|$—so the "cyclic objects" cancel out. We then have $\det(L^-) = \sum_{D\in\mathcal{D}} \mathrm{sgn}(D) = |\mathcal{T}| + |\mathcal{D}^+| - |\mathcal{D}^-| = |\mathcal{T}|$ and the theorem follows.

To realize this plan, we first collect several easy properties of the signed digraphs in D.

(i) *If $i \to j$ is a directed edge, then $\{i, j\}$ is an edge of G.* (Clear.)

(ii) *Every vertex, with the exception of n, has exactly one outgoing edge, while n has no outgoing edge.* (Obvious.)

(iii) *All ingoing edges of n are positive.* (Clear.)

(iv) *No vertex has more than one negative ingoing edge.* This is because two negative ingoing edges $j \to i$ and $k \to i$ would mean two circled entries ℓ_{ji} and ℓ_{ki} in the ith column.

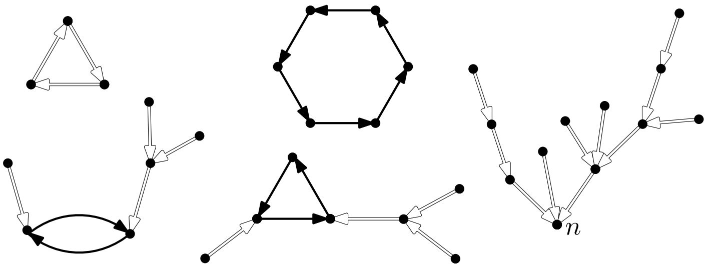
(v) *If a vertex i has a negative ingoing edge, then the outgoing edge is also negative.* Indeed, a negative ingoing edge $j \to i$ means that the off-diagonal entry ℓ_{ji} is circled, and hence none of the 1's in the diagonal entry ℓ_{ii} may be circled (which would be the only way of getting a positive outgoing edge from i).

Claim A. *These properties characterize $\mathcal{D}$. That is, if D is a signed digraph satisfying (i)–(v), then $D \in \mathcal{D}$.*

Proof. Given D, we determine the circled entry in each row i, $1 \le i \le n-1$, of L^-. We look at the single outgoing edge $i \to j$. If it is positive, we circle the appropriate 1 in ℓ_{ii}, and if it is negative, we circle ℓ_{ij}. We can't have two circled entries in a single column, since they would correspond to the situations excluded in (iv) or (v). $\square$

Next, we use (i)–(v) to describe the structure of D.

Claim B. *Each $D \in \mathcal{D}$ has the following structure (illustrated in the next picture):*



(a) *The vertex set is partitioned into one or more subsets V_1, V_2,…,V_k corresponding to the components of D, with no edges connecting different V_i. If V_1 is the subset containing the vertex n, then the subgraph on V_0 is a tree with all edges directed towards n. The subgraph on every other V_i contains a single directed cycle of length at least 2, and a tree (possibly empty) attached to each vertex of the cycle, with edges directed towards the cycle.*

(b) *The edges not belonging to the directed cycles are all positive, and in each directed cycle either all edges are positive or all edges are negative.*

(c) *Conversely, each possible D with this structure and satisfying (i) above belongs to $\mathcal{D}$.*

Sketch of proof. Part (a), describing the structure of the digraph, is a straightforward consequence of (ii) (a single outgoing edge for every vertex except for n), and we leave it as an exercise. (If we added a directed loop to n, then every vertex has exactly one outgoing edge, and we get a so-called *functional digraph*, for which the structure as in (a) is well known.)

Concerning (b), if we start at a negative edge and walk on, condition (v) implies that we are going to encounter only negative edges. We thus can't reach n, since its incoming edges are positive, and so at some point we start walking around a negative cycle. Finally, a negative edge can't enter such a negative cycle from outside by (iv).

As for (c), if D has the structure as described in (a) and (b), the conditions (ii)–(iv) are obviously satisfied and Claim A applies. This proves Claim B. $\square$

The first item in our plan of the proof is now very easy to complete.

Corollary. *All $D \in \mathcal{T}$ have a positive sign and they are in one-to-one correspondence with the spanning trees of G.*

Proof. If $D \in \mathcal{D}$ has no directed cycles, then D is a tree with positive edges directed towards the vertex n. Moreover, π_D is the identity permutation since all the circled elements in the term corresponding to D lie on the diagonal of L^-. Thus $\operatorname{sgn}(D) = +1$, and if we forget the orientations of the edges, we arrive at a spanning tree of D. Conversely, given a spanning tree of G, we can orienting its edges towards n, and we obtain a $D \in \mathcal{T}$. $\square$

It remains to deal with the "cyclic objects". For $D \in \mathcal{D}^+ \cup \mathcal{D}^-$, let the *smallest cycle* be the directed cycle that contains the vertex with the smallest number (among all vertices in cycles). Let $\overline{D}$ be obtained from D by changing the signs of all edges in the smallest cycle.

Obviously $\overline{\overline{D}} = D$, and for $D \in \mathcal{D}$ we have $\overline{D} \in \mathcal{D}$ as well, as can be seen using Claim B. The following claim then shows that the mapping sending D to $\overline{D}$ is a bijection between $\mathcal{D}^+$ and $\mathcal{D}^-$, which is all that we need to finish the proof of the theorem.

Claim C. $\operatorname{sgn}(\overline{D}) = -\operatorname{sgn}(D)$.

Proof. We have $\operatorname{sgn}(D) = \operatorname{sgn}(\pi_D)(-1)^m$, where m is the number of negative edges of D and π_D is the associated permutation.

Let $i_1, i_2, \ldots, i_s$ be the vertices of the smallest cycle of D, numbered so that the directed edges of the cycle are $i_1 \to i_2$, $i_2 \to i_3, \ldots$, $i_{s-1} \to i_s$, $i_s \to i_1$.

In one of D and $\overline{D}$, the smallest cycle is positive; say in D (if it is positive in $\overline{D}$, the argument is similar). Positive edges correspond to entries on the diagonal of L^-, and thus the i_j are fixed points of the permutation π_D, i.e., $\pi_D(i_j) = i_j$, $j = 1, 2, \ldots, s$. In $\overline{D}$, the smallest cycle is negative, and so for $\pi_{\overline{D}}$ we have $\pi_{\overline{D}}(i_1) = i_2, \ldots$, $\pi_{\overline{D}}(i_{s-1}) = i_s$, $\pi_{\overline{D}}(i_s) = i_1$. Otherwise, π_D and $\pi_{\overline{D}}$ coincide.

So $\pi_{\overline{D}}$ has an extra cycle of length s, and thus it can be obtained from π_D by $s-1$ transpositions. Hence $\operatorname{sgn}(\pi_{\overline{D}}) = (-1)^{s-1} \operatorname{sgn}(\pi_D)$, and

$$\operatorname{sgn}(\overline{D}) = \operatorname{sgn}(\pi_{\overline{D}})(-1)^{m+s} = (-1)^{s-1} \operatorname{sgn}(\pi_D)(-1)^{m+s} = -\operatorname{sgn}(D).$$

Claim C, and thus also the theorem, are proved. □

Sources. The theorem is usually attributed to

> G. Kirchhoff, *Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird*, Ann. Phys. Chem. **72** (1847), 497–508,

while

> J. J. Sylvester, *On the change of systems of independent variables*, Quart. J. Pure Appl. Math. **1** (1857), 42–56

is regarded as the first complete proof.

The above proof mostly follows

> A. T. Benjamin and N. T. Cameron, *Counting on determinants*, Amer. Math. Monthly **112** (2005), 481–492.

Benjamin and Cameron attribute the proof to

> S. Chaiken, *A Combinatorial proof of the all-minors matrix tree theorem*, SIAM J. Alg. Disc. Methods **3** (1982), 319–329,

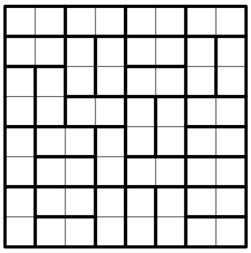
but it may not be easy to find it there, since the paper deals with a more general setting.

Miniature 22

In How Many Ways Can a Man Tile a Board?

The answer, my friend, is a determinant,[1] at least in many cases of interest.

There are 12988816 tilings of the 8×8 chessboard by 2×1 rectangles (dominoes). Here is one of them:



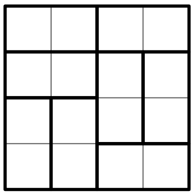
How can they all be counted?

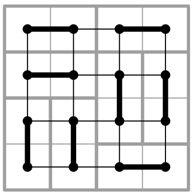
As the next picture shows, domino tilings of a chessboard are in one-to-one correspondence with **perfect matchings**[2] in the underlying **square grid** graph:

---

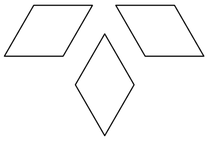
[1] With apologies to Mr. Dylan.

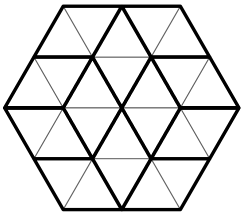
[2] A perfect matching in a graph G is a subset $M \subseteq E(G)$ of the edge set such that each vertex of G is contained in exactly one edge of M.

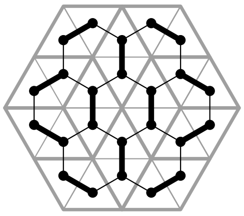




Another popular kind of tilings are **lozenge tilings** (or *rhombic tilings*). Here the board is made of equilateral triangles, and the tiles are the three rhombi obtained by gluing two adjacent triangles:







As the right picture illustrates, these tilings correspond to perfect matchings in **honeycomb graphs**.

We will explain how one can express the number of perfect matchings in these graphs, and many others, by a determinant. First we need to introduce some notions.

The bipartite adjacency matrix and Kasteleyn signings. We recall that a graph G is **bipartite** if its vertices can be divided into two classes $\{u_1, u_2, \ldots, u_n\}$ and $\{v_1, v_2, \ldots, v_m\}$ so that the edges go only between the two classes, never within the same class.

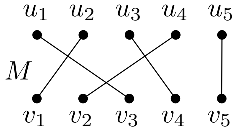
We may assume that $m = n$, i.e., the classes have the same size, for otherwise, G has no perfect matching.

We define the **bipartite adjacency matrix** of such G as the $n \times n$ matrix B given by

$$b_{ij} := \begin{cases} 1 & \text{if } \{u_i, v_j\} \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

Let S_n denote the set of all permutations of the set $\{1, 2, \ldots, n\}$. Every perfect matching M in G corresponds to a unique permutation

$\pi \in S_n$, where $\pi(i)$ is defined as the index j such that the edge $\{u_i, v_j\}$ lies in M. Here is an example:



$\pi(1) = 3, \pi(2) = 1, \pi(3) = 4, \pi(4) = 2, \pi(5) = 5$

In the other direction, when does G have a perfect matching corresponding to a given permutation $\pi \in S_n$? Exactly if $b_{1,\pi(1)} = b_{2,\pi(2)} = \cdots = b_{n,\pi(n)} = 1$. Therefore, the number of perfect matchings in G equals

$$\sum_{\pi \in S_n} b_{1,\pi(1)} b_{2,\pi_2} \cdots b_{n,\pi(n)}.$$

This expression is called the **permanent** of the matrix B and denoted by $\mathrm{per}(B)$. The permanent makes sense for arbitrary square matrices, but here we stick to bipartite adjacency matrices, i.e., matrices made of 0's and 1's.

The above formula for the permanent looks very similar to the definition of the determinant; the determinant has "only" the extra factor $\mathrm{sgn}(\pi)$ in front of each term. But the difference is actually a crucial one: The permanent lacks the various pleasant properties of the determinant, and while the determinant can be computed reasonably fast even for large matrices, the permanent is computationally hard, even for matrices consisting only of 0's and 1's.[3]

Here is the key idea of this section. Couldn't we cancel out the effect of the factor $\mathrm{sgn}(\pi)$ by changing the signs of some carefully selected subset of the b_{ij}, and thereby turn the *permanent* of B into the *determinant* of some other matrix? As we will see, for many graphs this can be done. Let us introduce a definition capturing this idea more formally.

We let a **signing** of G be an arbitrary assignment of signs to the edges of G, i.e., a mapping $\sigma\colon E(G) \to \{-1, +1\}$, and we define a matrix B^σ, which is a “signed version” of B, by

---

[3] In technical terms, computing the permanent of a 0-1 matrix, which is equivalent to computing the number of perfect matchings in a bipartite graph, is #P-complete.

$$b^\sigma_{ij} := \begin{cases} \sigma(u_i, v_j) & \text{if } \{u_i, v_j\} \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

We call σ a **Kasteleyn signing** for G if

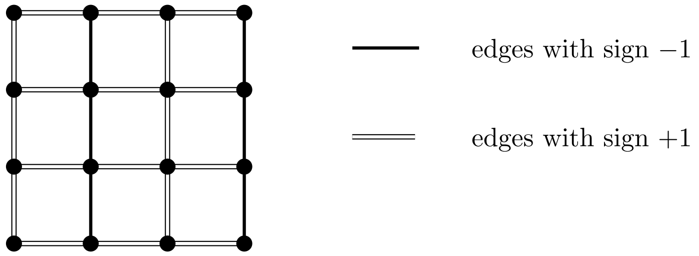
$$|\det(B^\sigma)| = \mathrm{per}(B).$$

Not all bipartite graphs have a Kasteleyn signing; for example, the complete bipartite graph $K_{3,3}$ doesn’t have one, as a diligent and energetic reader can check. But it turns out that all *planar*[4] bipartite graphs do.

In order to focus on the essence and avoid some technicalities, we will deal only with **2-connected** graphs, which means that every edge is contained in at least one cycle (which holds for the square grids and for the honeycomb graphs). As is not difficult to see, and well known, in a planar drawing of a 2-connected graph G, the boundary of every face forms a cycle in G.

Theorem. *Every 2-connected planar bipartite graph G has a Kasteleyn signing, which can be found efficiently.*[5] *Consequently, the number of perfect matchings in such a graph can be computed in polynomial time.*

For the grid graphs derived from the tiling examples above, Kasteleyn signings happen to be very simple. Here is one for the square grid graph,



---

[4] We recall that a graph is **planar** if it can be drawn in the plane without edge crossings.

[5] The proof will obviously give a polynomial-time algorithm, but with some more work one can obtain even a linear-time algorithm.

and for the hexagonal grid we can even give all edges the sign $+1$. Both of these facts will immediately follow from Lemma B below.

The restriction to 2-connected graphs in the theorem can easily be removed with a little more work. The restriction to *bipartite* graphs is also not essential. It makes the presentation slightly simpler, but an analogous theory can be developed for the non-bipartite case along similar lines—the interested readers will find this in the literature.

On the other hand, the assumption of *planarity* is more substantial: The method certainly breaks down for a general nonplanar graph, and as was mentioned above, counting the number of perfect matchings in a general graph is computationally hard. The class of graphs where this approach works, the so-called *Pfaffian graphs*, is somewhat wider than all planar graphs, but not easy to describe, and most applications deal with planar graphs anyway.

Properly signed cycles. As a first step towards the proof, we give a sufficient condition for a signing to be Kasteleyn. It may look mysterious at first sight, but in the proof we will see where it comes from.

Let C be a cycle in a bipartite graph G. Then C has an even length, which we write as 2ℓ. Let σ be a signing of G, and let n_C be the number of *negative edges* (i.e., edges with sign -1) in C. Then we call C **properly signed** with respect to σ if $n_C \equiv \ell - 1 \pmod 2$. In other words, a properly signed cycle of length $4, 8, 12, \ldots$ contains an odd number of negative edges, while a properly signed cycles of length $6, 10, 14, \ldots$ contains an even number of negative edges.

Further let us say that a cycle C is **evenly placed** if the graph obtained from G by deleting all vertices of C (and the adjacent edges) has a perfect matching.

Lemma A. *Suppose that σ is a signing of a bipartite graph G (no planarity assumed here) such that every evenly placed cycle in G is properly signed. Then σ is a Kasteleyn signing for G.*

Proof. This is straightforward. Let the signing σ as in the lemma be fixed, and let M be a perfect matching in G, corresponding to a permutation π. We define the **sign** of M as the sign of the corresponding term in $\det(B^\sigma)$; explicitly,

$$\operatorname{sgn}(M) := \operatorname{sgn}(\pi) b^\sigma_{1,\pi(1)} b^\sigma_{2,\pi(2)} \cdots b^\sigma_{n,\pi(n)} = \operatorname{sgn}(\pi) \prod_{e \in M} \sigma(e).$$

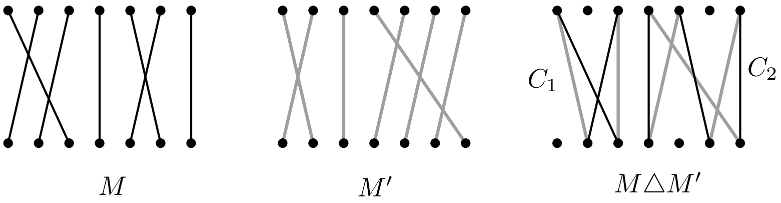
It is easy to see that σ is a Kasteleyn signing if (and only if) all perfect matchings in G have the same sign.

Let M and M' be two perfect matchings in G, with the corresponding permutations π and π'. Then

$$\begin{aligned} \operatorname{sgn}(M)\operatorname{sgn}(M') &= \operatorname{sgn}(\pi)\operatorname{sgn}(\pi')\Big(\prod_{e\in M}\sigma(e)\Big)\Big(\prod_{e\in M'}\sigma(e)\Big) \\ &= \operatorname{sgn}(\pi)\operatorname{sgn}(\pi') \prod_{e \in M \triangle M'} \sigma(e), \end{aligned}$$

where $\triangle$ denotes the symmetric difference.

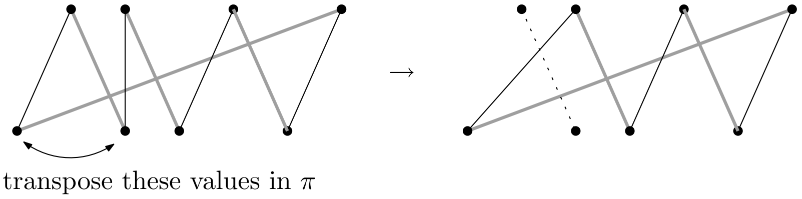
The symmetric difference $M \triangle M'$ is a disjoint union of evenly placed cycles, as the picture illustrates:



Let these cycles be $C_1, C_2, \ldots, C_k$, and let the length of C_i be $2\ell_i$. Since C_i is evenly placed, it must be properly signed by the assumption in the lemma, and so we have $\prod_{e\in C_i} \sigma(e) = (-1)^{\ell_i - 1}$. Thus $\prod_{e \in M\triangle M'} \sigma(e) = (-1)^t$ with $t := \ell_1 - 1 + \ell_2 - 1 + \cdots + \ell_k - 1$.

It remains to check that π can be converted to π' by t transpositions (then, by the properties of the sign of a permutation, we have $\operatorname{sgn}(\pi) = (-1)^t \operatorname{sgn}(\pi')$, and thus $\operatorname{sgn}(M) = \operatorname{sgn}(M')$ as needed).

This can be done one cycle C_i at a time. As the next picture illustrates for a cycle of length $2\ell_i = 8$, by modifying π with a suitable transposition we can "cancel" two edges of the cycle and pass to a cycle of length $2\ell_i - 2$ (black edges belong to M, gray edges to M', and the dotted edge in the right drawing now belongs to both M and M').



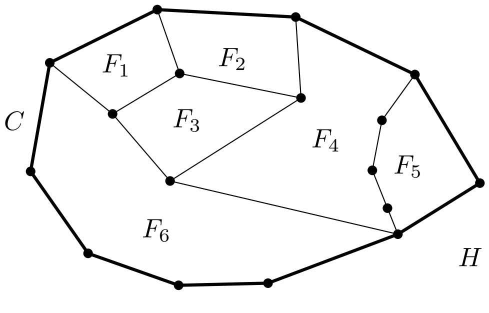
Continuing in this way for $\ell_i - 1$ steps, we cancel C_i, and we can proceed with the next cycle. Lemma A is proved. □

The rest of the proof of the theorem is a simple graph theory. First we show that for graphs as in the theorem, it is sufficient to check the condition in Lemma A only for special cycles, namely, face boundaries. Clearly it is enough to deal with *connected* graphs.

Lemma B. *Let G be a planar bipartite graph that is both connected and 2-connected, and let us fix a planar drawing of G. If σ is a signing of G such that the boundary cycle of every inner face in the drawing is properly signed, then σ is a Kasteleyn signing.*

Proof of Lemma B. Let C be an evenly placed cycle in G; we need to prove that it is properly signed.

Let the length of C be 2ℓ. Let $F_1, \ldots, F_k$ be the inner faces enclosed in C in the drawing, and let C_i be the boundary cycle of F_i, of length $2\ell_i$. Let H be the subgraph of G obtained by deleting all vertices and edges drawn outside C; in other words, H is the union of the C_i.



We want to see how the parity of ℓ is related to the parities of the ℓ_i. To this end, we need to do some counting. The number of vertices of H is $r + 2\ell$, where r is the number of vertices lying in the interior of C. Every edge of H belongs to exactly two cycles among $C, C_1, \ldots, C_k$, and so the number of edges of H equals $\ell+\ell_1+\cdots+\ell_k$. Finally, the drawing of H has $k + 1$ faces: $F_1, \ldots, F_k$ and the outer one.

Now we apply **Euler's formula**, which tells us that for every drawing of a connected planar graph, the number of vertices plus the number of faces equals the number of edges plus 2. Thus

$$r + 2\ell + k + 1 = \ell + \ell_1 + \cdots + \ell_k + 2. \tag{18}$$

Next, we use the assumption that C is evenly placed. Since the complement of C in G has a perfect matching, the number r of vertices inside C must be even. Therefore, from (18) we get

$$\ell - 1 \equiv \ell_1 + \cdots + \ell_k - k \pmod 2. \tag{19}$$

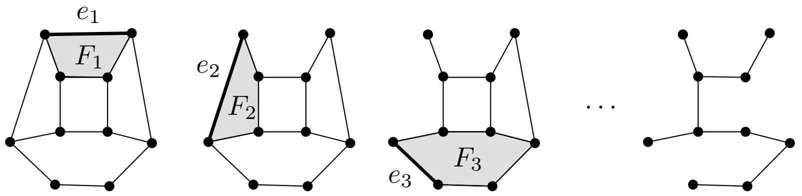
Let n_C be the number of negative edges in C, and similarly for n_{C_i}. The sum $n_C + n_{C_1} + \cdots + n_{C_k}$ is even because it counts every negative edge twice, and so

$$n_C \equiv n_{C_1} + \cdots + n_{C_k} \pmod 2. \tag{20}$$

Finally, we have $n_{C_i} \equiv \ell_i - 1 \pmod 2$ since the C_i are properly signed. Combining this with (19) and (20) gives $n_C \equiv \ell - 1 \pmod 2$. Hence C is properly signed. Lemma B now follows from Lemma A. □

Proof of the theorem. Given a connected, 2-connected, planar, bipartite G, we fix some planar drawing, and we want to construct a signing as in Lemma B, with the boundary of every inner face properly signed.

First we start deleting edges from G, as the next picture illustrates:



$G_1 = G$ G_2 G_3 G_6

We set $G_1 := G$, and G_{i+1} is obtained from G_i by deleting an edge e_i that separates an inner face F_i from the outer (unbounded) face (in the current drawing). The procedure finishes with some G_k that has no such edge. Then the drawing of G_k has only the outer face.

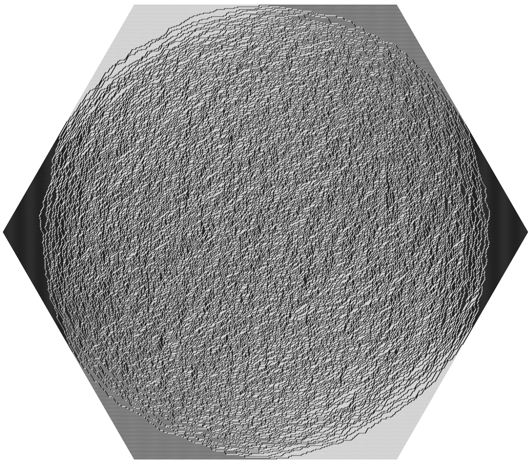
Now we choose the signs of the edges of G_k arbitrarily, and we extend this to a signing of G by going backwards, choosing the signs for $e_{k-1}, e_{k-2}, \ldots, e_1$ in this order. When we consider e_i, it is contained in the boundary of the single inner face F_i in the drawing of G_i, so we can set $\sigma(e_i)$ so that the boundary of F_i is properly signed. The theorem is proved. $\square$

From the determinant formula one can obtain, with some effort, the following amazing formula for the number of domino tilings of an $m \times n$ chessboard:

$$\left[\prod_{k=1}^{m}\prod_{\ell=1}^{n}\left(2\cos\frac{\pi k}{m+1} + 2\mathrm{i}\cos\frac{\pi \ell}{n+1}\right)\right]^{1/2},$$

where i is the imaginary unit. But the determinants can be used not only for counting, but also for generating a *random* perfect matching (chosen uniformly among all possible perfect matchings), and for analyzing its typical properties. Such results are relevant for questions in theoretical physics.

Here is a quick illustration of an interesting phenomenon for random tilings. The next picture shows a random lozenge tiling of a large hexagon:



The three types of tiles are painted black, white, and gray. One can see that, while the tiling looks "chaotic" in the central circle, the regions outside this circle are "frozen", i.e., tiled by rhombi of a single type. (This is a typical property of a *random* tiling—definitely not *all* tilings look like this.) This is called the "arctic circle" phenomenon.

Depending on the board's shape, various complicated curves may play the role of the arctic circle. In some cases, there are no frozen regions at all, e.g., for domino tilings of rectangular chessboards—these look chaotic everywhere. The determinant formula provides a crucial starting point for analyzing such phenomena.

Sources. Counting perfect matchings is considered in several areas; mathematicians often talk about *tilings*, computer scientists about *perfect matchings*, and physicists about the *dimer model* (which is a highly simplified but still interesting model in solid-state physics). The idea of counting perfect matching in a square grid via determinants was invented in the dimer context, in

> P. W. Kasteleyn, *The statistics of dimers on a lattice I. The number of dimer arrangements on a quadratic lattice*, Physica **27** (1961), 1209–1225

and independently in

> H. N. V. Temperley and M. E. Fisher, *Dimer problem in statistical mechanics—An exact result*, Philos. Magazine **6** (1961), 1061–1063.

The material covered in this section is just the beginning of amazing theories going in several directions. As starting points one can use, e.g.,

> R. Kenyon, *The planar dimer model with boundary: A survey*, Directions in mathematical quasicrystals, CRM Monograph Ser. 13, Amer. Math. Soc., Providence, R.I., 2000, pp. 307–328

(discussing tilings, dimers, the arctic circle, random surfaces, and such) and

> R. Thomas, *A survey of Pfaffian orientations of graphs*, in International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 963–984

(with graph-theoretic and algorithmic aspects of Pfaffian graphs).

Miniature 23

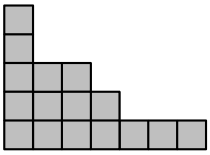
More Bricks—More Walls?

One of the classical topics in enumeration are **integer partitions**. For example, there are five partitions of the number 4:

$$\begin{aligned} 4 &= 1+1+1+1+1 \\ 4 &= 2+1+1 \\ 4 &= 2+2 \\ 4 &= 3+1 \\ 4 &= 4. \end{aligned}$$

The order of the addends in a partition doesn't matter, and it is customary to write them in a nonincreasing order as we did above.

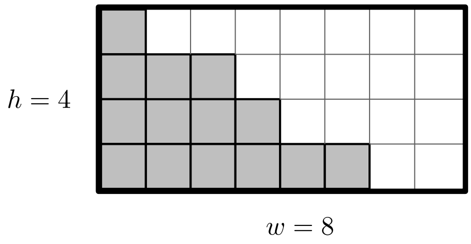
A partition of n is often represented graphically by its **Ferrers diagram**, which one can think of as a nondecreasing wall built of n bricks. For example, the following Ferrers diagram



corresponds to $16 = 5+3+3+2+1+1+1$.

How can we determine or estimate $p(k)$, the number of partitions of the integer k? This is a surprisingly difficult enumeration problem, ultimately solved by a formula of Hardy and Ramanujan. The asymptotics of $p(k)$ is $p(k) \sim \frac{1}{4k\sqrt{3}} e^{\pi\sqrt{2k/3}}$, where $f(k) \sim g(k)$ means $\lim_{k\to\infty} \frac{f(k)}{g(k)} = 1$.

Here we consider another matter, the number $p_{w,h}(k)$ of partitions of k with at most w addends, none of them exceeding h. In other words, $p_{w,h}(k)$ is the number of ways to build a nonincreasing wall out of k bricks inside a box of width w and height h:



Here is the main result of this section:

Theorem. *For every* $w \geq 1$ *and* $h \geq 1$ *we have*

$$p_{w,h}(0) \leq p_{w,h}(1) \leq \cdots \leq p_{w,h}\left(\lfloor \tfrac{wh}{2} \rfloor\right)$$

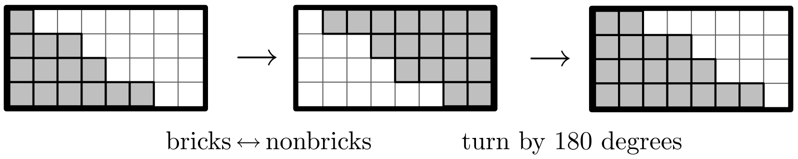
and

$$p_{w,h}\left(\lceil \tfrac{wh}{2} \rceil\right) \geq p_{w,h}\left(\lceil \tfrac{wh}{2} \rceil + 1\right) \geq \cdots \geq p_{w,h}(wh-1) \geq p_{w,h}(wh).$$

That is, $p_{w,h}(k)$ *as a function of* k *is nondecreasing for* $k \leq \frac{wh}{2}$ *and nonincreasing for* $k \geq \frac{wh}{2}$.

So the first half of the theorem tells us that with more bricks we can build more (or rather, at least as many) walls. This goes on until half of the box is filled with bricks; after that, we already have too little space and the number of possible walls starts decreasing.

Actually, once we know that $p_{w,h}(k)$ is nondecreasing for $k \leq \frac{wh}{2}$, then it must be nonincreasing for $k \geq \frac{wh}{2}$, because $p_{w,h}(k) = p_{w,h}(wh-k)$, as can be seen using the following bijection transforming walls with k bricks into walls with $wh - k$ bricks:



The theorem is one of the results that look intuitively obvious but are surprisingly hard to prove. The great Cayley used this as a fact requiring no proof in his 1856 memoir, and only about twenty years later did Sylvester discover the first proof.

One would naturally expect such a combinatorial problem to have a combinatorial solution, perhaps simply an injective map assigning to every wall of k bricks a wall of $k+1$ bricks (for $k+1 \leq \frac{wh}{2}$). But to my knowledge, nobody has managed to discover a proof of this kind, and estimating $p_{w,h}(k)$ or expressing it by a formula doesn't seem to lead to the goal either.

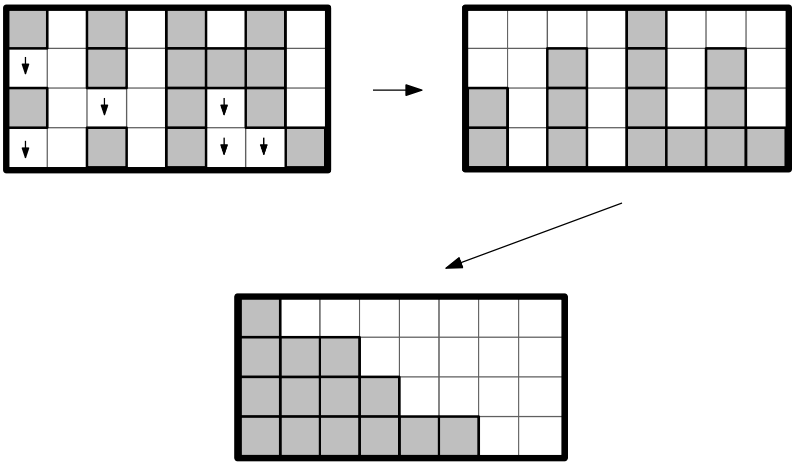
Earlier proofs of the theorem used relatively heavy mathematical tools, essentially representations of Lie algebras. The proof shown here is a result of several simplifications of the original ideas, and it uses "only" matrix-rank arguments.

Functions, or sequences, that are first nondecreasing and then, from some point on, nonincreasing, are called **unimodal** (and so are functions that begin as nonincreasing and continue as nondecreasing). There are many important results and conjectures in various areas of mathematics asserting that certain quantities form an unimodal sequence, and the proof below contains tools of general applicability.

Preliminary considerations. Let us write $n := wh$ for the area of the box, and let us fix a numbering of the n squares in the box by the numbers $1, 2, \ldots, n$.

To prove the theorem, we will show that $p_{w,h}(k) \leq p_{w,h}(\ell)$ for $0 \leq k < \ell \leq \frac{n}{2}$.

The first step is to view a wall in the box as an *equivalence class.* Namely, we start with an arbitrary set of k bricks filling some k squares in the box, and then we tidy them up into a nonincreasing wall:



First we push down the bricks in each column, and then we rearrange the columns into a nonincreasing order.

Let us call two k-element subsets $K, K' \subseteq \{1, 2, \ldots, n\}$, understood as sets of k squares in the box, **wall-equivalent** if they lead to the same nonincreasing wall. This indeed defines an equivalence on the set $\mathcal{K}$ of all k-element subsets of $\{1, 2, \ldots, n\}$. Let the equivalence classes be $\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_r$, where $r := p_{w,h}(k)$.

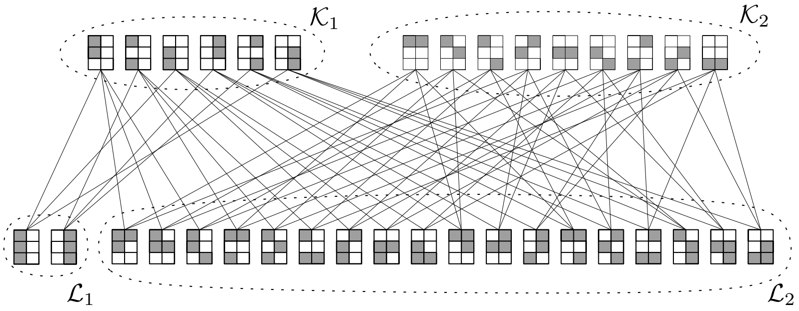
Let us phrase the definition of the wall-equivalence differently, in a way that will be more convenient latter. Let π be a permutation of the n squares in the box; let us say that π **doesn't break columns** if it corresponds to first permuting the squares in each column arbitrarily, and then permuting the columns. It is easily seen that two subsets $K, K' \in \mathcal{K}$ are wall-equivalent exactly if $K' = \pi(K)$ for some permutation that doesn't break columns.[1]

Next, let $\mathcal{L}$ be the set of all ℓ-element subsets of $\{1, 2, \ldots, n\}$, and let it be divided similarly into $s := p_{w,h}(\ell)$ classes $\mathcal{L}_1, \ldots, \mathcal{L}_s$ according to wall-equivalence. The goal is to prove that $r \leq s$.

---

[1] In a more mature mathematical language, the permutations that don't break columns form a permutation group acting on $\mathcal{K}$, and the classes of the wall-equivalence are the orbits of this action. Some things in the sequel could (should?) also be phrased in the language of actions of permutation groups, but I decided to avoid this terminology, with the hope of deterring slightly fewer students.

Let us consider the bipartite graph G with vertex set $\mathcal{K} \cup \mathcal{L}$ and with edges corresponding to inclusion; i.e., a k-element set $K \in \mathcal{K}$ is connected to an ℓ-element set $L \in \mathcal{L}$ by an edge if $K \subseteq L$. A small-scale illustration with $w = 2$, $h = 3$, $k = 2$, and $\ell = 3$ follows:



Claim. *For every i and j, all $L \in \mathcal{L}_j$ have the same number d_{ij} of neighbors in $\mathcal{K}_i$.*

Proof. Let $L, L' \in \mathcal{L}_j$, and let us fix some permutation π that doesn't break columns and such that $L' = \pi(L)$. For $K \in \mathcal{K}_i$, we have $\pi(K) \in \mathcal{K}_i$ as well (by the alternative description of the wall-equivalence), and it is easily seen that $K \mapsto \pi(K)$ defines a bijection between the neighbors of L lying in $\mathcal{K}_i$ and the neighbors of L' lying in $\mathcal{K}_i$. $\square$

Let us now pass to a more general setting for a while: Let U, V be disjoint finite sets, let $(U_1, \ldots, U_r, V_1, \ldots, V_s)$ be a partition of $U \cup V$ with $U = U_1 \cup \cdots \cup U_r$ and $V = V_1 \cup \cdots \cup V_s$, where the U_i and V_j are all nonempty, and let G be a bipartite graph on the vertex set $U \cup V$ (with all edges going between U and V). We call the partition $(U_1, \ldots, U_r, V_1, \ldots, V_s)$ **$\boldsymbol{V}$-degree homogeneous** w.r.t. G if the condition as in the claim holds, i.e., all vertices in V_j have the same number d_{ij} of neighbors in U_i, for all i and j. In such case, we call the matrix $D = (d_{ij})_{i=1}^{r}{}_{j=1}^{s}$ the **$\boldsymbol{V}$-degree matrix** of the partition (with respect to G).

In the setting introduced above, we have a bipartite graph with a V-degree homogeneous partition, and we would like to conclude that r, the number of the U-pieces, can't be smaller than s, the number of

V-pieces. The next lemma gives a sufficient condition, which we will then be able to verify for our particular G. The condition essentially says that V is at least as large as U for a "linear-algebraic reason".

To formulate the lemma, we set up a $|U| \times |V|$ matrix B (the **bipartite adjacency matrix** of G), with rows indexed by the vertices in U and columns indexed by the vertices in V, whose entries b_{uv} are given by

$$b_{uv} := \begin{cases} 1 & \text{if } \{u,v\} \in E(G) \\ 0 & \text{otherwise.} \end{cases}$$

Lemma. *Let G be a bipartite graph as above, let $(U_1, U_2, \ldots, U_r, V_1, V_2, \ldots, V_s)$ be a V-degree homogeneous partition of its vertices, and let us suppose that the rows of the matrix B are linearly independent. Then $r \leq s$.*

Proof. This powerful statement is quite easy to prove. We will show that the $r \times s$ V-degree matrix D has linearly independent rows, which means that it can't have fewer columns than rows, and thus $r \leq s$ indeed.

Let $B[U_i, V_j]$ denote the submatrix of B consisting of the entries b_{uv} with $u \in U_i$ and $v \in V_j$; schematically

$$B = \begin{array}{c|c|c|c|c|}
\multicolumn{1}{c}{} & \multicolumn{1}{c}{V_1} & \multicolumn{1}{c}{V_2} & \multicolumn{1}{c}{V_3} & \multicolumn{1}{c}{V_4} \\ \cline{2-5}
U_1 & B[U_1, V_1] & B[U_1, V_2] & & \\ \cline{2-5}
U_2 & & & & \\ \cline{2-5}
U_3 & & & & B[U_3, V_4] \\ \cline{2-5}
\end{array}$$

The V-degree homogeneity condition translates to the matrix language as follows: The sum of each of the columns of $B[U_i, V_j]$ equals d_{ij}.

For a vector $\mathbf{x} \in \mathbb{R}^r$, let $\tilde{\mathbf{x}} \in \mathbb{R}^{|U|}$ be the vector indexed by the vertices in U obtained by replicating $|U_i|$-times the component x_i; that is, $\tilde{x}_u = x_i$ for all $u \in U_i$, $i = 1, 2, \ldots, r$.

For this $\tilde{\mathbf{x}}$, we consider the product $\tilde{\mathbf{x}}^T B$. Its vth component equals $\sum_{u\in U}\tilde{x}_u b_{uv} = \sum_{i=1}^r x_i \sum_{u\in U_i} b_{uv} = \sum_{i=1}^r x_i d_{ij} = (\mathbf{x}^T D)_j$. Hence $\mathbf{x}^T D = \mathbf{0}$ implies $\tilde{\mathbf{x}}^T B = \mathbf{0}$.

Let us assume for contradiction that the rows of D are linearly dependent; that is, there is a nonzero $\mathbf{x} \in \mathbb{R}^r$ with $\mathbf{x}^T D = \mathbf{0}$. Then $\tilde{\mathbf{x}} \neq \mathbf{0}$ but, as we've just seen, $\tilde{\mathbf{x}}^T B = \mathbf{0}$. This contradicts the linear independence of the rows of B and proves the lemma. □

Proof of the theorem. We return to the particular bipartite graph G introduced above, with vertex set $\mathcal{K} \cup \mathcal{L}$ and with the $\mathcal{L}$-degree homogeneous partition $(\mathcal{K}_1, \ldots, \mathcal{K}_r, \mathcal{L}_1, \ldots, \mathcal{L}_s)$ according to the wall-equivalence. For applying the lemma, it remains to show that the rows of the corresponding matrix B are linearly independent.

This result, known as **Gottlieb's theorem**,[2] has proved useful in several other applications as well. Explicitly, it tells us that *for* $0 \le k < \ell \le \frac{n}{2}$, *the zero-one matrix B with rows indexed by $\mathcal{K}$ (all k-subsets of $\{1, 2, \ldots, n\}$), columns indexed by $\mathcal{L}$ (all ℓ-subsets), and the nonzero entries corresponding to containment, has linearly independent rows.*

Several proofs are known; here we present one resembling the proof of the lemma above.

Proof of Gottlieb's theorem. For contradiction, we assume that $\mathbf{y}^T B = \mathbf{0}$ for some nonzero vector $\mathbf{y}$. The components of $\mathbf{y}$ are indexed by k-element sets; let us fix some $K_0 \in \mathcal{K}$ with $y_{K_0} \neq 0$.

Next, we partition both $\mathcal{K}$ and $\mathcal{L}$ into $k+1$ classes according to the size of the intersection with K_0 (this partition has nothing to do with the partition of $\mathcal{K}$ and $\mathcal{L}$ considered earlier—we just re-use the same letters):

$$\begin{aligned} \mathcal{K}_i &:= \{K \in \mathcal{K} : |K \cap K_0| = i\}, \quad i = 0, 1, \ldots, k \\ \mathcal{L}_j &:= \{L \in \mathcal{L} : |L \cap K_0| = j\}, \quad j = 0, 1, \ldots, k. \end{aligned}$$

Every $\mathcal{K}_i$ and every $\mathcal{L}_j$ is nonempty—here we use the assumption $k < \ell \le \frac{n}{2}$ (if, for example, we had $k + \ell > n$, we would get $\mathcal{L}_0 = \emptyset$, since there wouldn't be enough room for an ℓ-element L disjoint from K_0).

---

[2] This is not the only theorem associated with Gottlieb's name, though.

Here, for a change, we will need that this partition is $\mathcal{K}$-degree homogeneous (with respect to the same bipartite graph as above, with edges representing inclusion). That is, every $K \in \mathcal{K}_i$ has the same number d_{ij} of neighbors in $\mathcal{L}_j$. More explicitly, d_{ij} is the number of ways of extending a k-element set K with $|K \cap K_0| = i$ to an ℓ-element $L \supset K$ with $|L \cap K_0| = j$; this number is clearly independent of the specific choice of K. (We could compute d_{ij} explicitly, but we don't need it.)

By this description, we have $d_{ij} = 0$ for $i > j$, and thus the $\mathcal{K}$-degree matrix D is upper triangular. Moreover, $d_{ii} \neq 0$ for all $i = 0, 1, \ldots, k$, and so D is non-singular.

Using the vector $\mathbf{y}$, we are going to exhibit a nonzero $\mathbf{x} = (x_0, x_1, \ldots, x_k)$ with $\mathbf{x}^T D = \mathbf{0}$, which will be a contradiction. A suitable $\mathbf{x}$ is obtained by summing the components of $\mathbf{y}$ over the classes $\mathcal{K}_i$:

$$x_i := \sum_{K \in \mathcal{K}_i} y_K.$$

We have $\mathbf{x} \neq \mathbf{0}$, since the class $\mathcal{K}_k$ contains only K_0, and so $x_k = y_{K_0} \neq 0$.

For every j we calculate

$$\begin{aligned} 0 &= \sum_{L \in \mathcal{L}_j} (\mathbf{y}^T B)_L = \sum_{L \in \mathcal{L}_j} \sum_{K \in \mathcal{K}} y_K b_{KL} = \sum_{K \in \mathcal{K}} y_K \sum_{L \in \mathcal{L}_j} b_{KL} \\ &= \sum_{i=0}^{k} \sum_{K \in \mathcal{K}_i} y_K d_{ij} = \sum_{i=0}^{k} x_i d_{ij} = (\mathbf{x}^T D)_j. \end{aligned}$$

Hence $\mathbf{x}^T D = \mathbf{0}$, and this is the promised contradiction to the non-singularity of D. Gottlieb's theorem, as well as our main theorem, are proved. □

Another example. For readers familiar with the notion of graph isomorphism, the following might be a rewarding exercise in applying the method shown above: Prove that if $g_n(k)$ stands for the number of nonisomorphic graphs with n vertices and k edges, then the sequence $g_n(0), g_n(1), \ldots, g_n(\binom{n}{2})$ is unimodal.

Sources. As was mentioned above, the theorem was implicitly assumed without proof in

A. Cayley, *A second memoir on quantics*, Phil. Trans. Roy. Soc. **146** (1856), 101–126.

The word "quantic" in the title means, in today's terminology, a homogeneous multivariate polynomial, and Cayley was interested in quantics that are invariant under the action of linear transformations. The first proof of the theorem was obtained in

J. J. Sylvester, *Proof of the hitherto undemonstrated fundamental theorem of invariants*, Philos. Mag. **5** (1878), 178–188.

A substantially more elementary proof than the previous ones, phrased in terms of group representations, was obtained in

R. P. Stanley, *Some aspects of groups acting on finite posets*, J. Combinatorial Theory Ser. A **32** (1982), 132–161.

Our presentation is based on that of Babai and Frankl in their textbook cited in the introduction.

Gottlieb's theorem was first proved in

D. H. Gottlieb, *A certain class of incidence matrices*, Proc. Amer. Math. Soc **17** (1066), 1233–1237.

The proof presented above rephrases an argument from

C. D. Godsil, *Tools from linear algebra*, Chapter 31 of R. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, North-Holland, Amsterdam, 1995, pp. 1705–1748.

For an introduction to integer partitions see

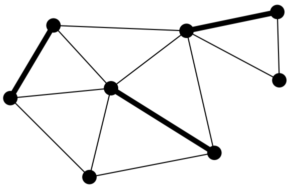
G. Andrews and K. Eriksson, *Integer partitions*, Cambridge University Press, Cambridge 2004,

(this is a very accessible source), or Wilf's lecture notes at `http://www.math.upenn.edu/~wilf/PIMS/PIMSLectures.pdf`.

Miniature 24

Perfect Matchings and Determinants

A **matching** in a graph G is a set of edges $F \subseteq E(G)$ such that no vertex of G is incident to more than one edge of F.



A **perfect matching** is a matching covering all vertices. The reader may want to find a perfect matching in the graph in the picture.

In Miniature 22, we counted perfect matchings in certain graphs via determinants. Here we will employ determinants in a simple algorithm for testing whether a given graph has a perfect matching. The basic approach is similar to the approach to testing matrix multiplication from Miniature 11. We consider only the bipartite case, which is simpler.

Consider a bipartite graph G. Its vertices are divided into two classes $\{u_1, u_2, \ldots, u_n\}$ and $\{v_1, v_2, \ldots, v_n\}$ and the edges go only between the two classes, never within one class. Both of the classes have the same size, for otherwise, the graph has no perfect matching. Let m stand for the number of edges of G.

Let S_n be the set of all permutations of the set $\{1, 2, \ldots, n\}$. Every perfect matching of G uniquely corresponds to a permutation $\pi \in S_n$. We can describe it in the form $\{\{u_1, v_{\pi(1)}\}, \{u_2, v_{\pi(2)}\}, \ldots, \{u_n, v_{\pi(n)}\}\}$.

We express the existence of a perfect matching by a determinant, but not of an ordinary matrix of numbers, but rather of a matrix whose entries are *variables.* We introduce a variable x_{ij} for every edge $\{u_i, v_j\} \in E(G)$ (so we have m variables altogether), and we define an $n \times n$ matrix A by

$$a_{ij} := \begin{cases} x_{ij} & \text{if } \{u_i, v_j\} \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

The determinant of A is a polynomial in the m variables x_{ij}. By the definition of a determinant, we get

$$\begin{aligned} \det(A) &= \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)} \\ &= \sum_{\substack{\pi \text{ describes a perfect} \\ \text{matching of } G}} \operatorname{sgn}(\pi) \cdot x_{1,\pi(1)} x_{2,\pi(2)} \cdots x_{n,\pi(n)}. \end{aligned}$$

Lemma. *The polynomial* $\det(A)$ *is identically zero if and only if* G *has no a perfect matching.*

Proof. The formula above the formula makes it clear that if G has no perfect matching, then $\det(A)$ is the zero polynomial.

To show the converse, we fix a permutation π that defines a perfect matching, and we substitute for the variables in $\det(A)$ as follows: $x_{i,\pi(i)} := 1$ for every $i = 1, 2, \ldots, n$, and all the remaining x_{ij} are 0. We have $\operatorname{sgn}(\pi) \cdot x_{1,\pi(1)} x_{2,\pi(2)} \cdots x_{n,\pi(n)} = \pm 1$ for this permutation π.

For every other permutation $\sigma \neq \pi$ there is an i with $\sigma(i) \neq \pi(i)$, thus $x_{i,\sigma(i)} = 0$, and therefore, all other terms in the expansion of $\det(A)$ are 0. For this choice of the x_{ij} we thus have $\det(A) = \pm 1$. □

Now we would like to test whether the polynomial $\det(A)$ is the zero polynomial. We can't afford to compute it explicitly as a polynomial, since it has the same number of terms as the number of perfect matchings of G and that can be exponentially many. But if we substitute any specific numbers for the variables x_{ij}, we can easily calculate the determinant, e.g., by the Gaussian elimination. So we can imagine that $\det(A)$ is available to us through a black box, from which we can obtain the value of the polynomial at any specified point.

For an arbitrary function given by a black box, we can never be sure that it is identically 0 unless we check its values at all points. But a polynomial has a wonderful property: Either it equals 0 everywhere, or almost nowhere. The following theorem expresses this quantitatively.

Theorem (The Schwartz–Zippel theorem[1]). *Let $\mathbb{K}$ be an arbitrary field, and let S be a finite subset of $\mathbb{K}$. Then for every non-zero polynomial $p(x_1, \ldots, x_m)$ of degree d in m variables and with coefficients from $\mathbb{K}$, the number of m-tuples $(r_1, r_2, \ldots, r_m) \in S^m$ with $p(r_1, r_2, \ldots, r_m) = 0$ is at most $d|S|^{m-1}$. In other words, if r_1, $r_2, \ldots,$ $r_m \in S$ are chosen independently and uniformly at random, then the probability of $p(r_1, r_2, \ldots, r_m) = 0$ is at most $\frac{d}{|S|}$.*

Before we prove this theorem, we get back to bipartite matchings. Let us assume that G has a perfect matching and thus $\det(A)$ is a nonzero polynomial. Then the Schwartz–Zippel theorem shows that if we calculate $\det(A)$ for values of the variables x_{ij} chosen independently at random from $S := \{1, 2, \ldots, 2n\}$, then the probability of getting 0 is at most $\frac{1}{2}$.

But in order to decide whether the determinant is 0 for a given substitution, we have to compute it exactly. In such a computation, we may encounter huge numbers, with about n digits, and then arithmetic operations would become quite expensive.

It is better to work with a finite field. The simplest way is to choose a prime number p, $2n \le p < 4n$ (by a theorem from number theory called Bertrand's postulate such a number always exists and it can be found sufficiently quickly) and operate in the finite field $\mathbb{F}_p$ of integers modulo p. Then the arithmetic operations are fast (if we prepare a table of inverse elements in advance).

---

[1] This Schwartz is really spelled with "t", unlike the one from the Cauchy–Schwarz inequality.

Using the Gaussian elimination for computing the determinant, we get a probabilistic algorithm for testing the existence of a bipartite matching in a given graph running in $O(n^3)$ time. It fails with a probability at most $\frac{1}{2}$. As usual, the probability of the failure can be reduced to 2^{-k} by repeating the algorithm k times.

The determinant can also be computed by the algorithms for fast matrix multiplication (mentioned in Miniature 10), and in this way we obtain the asymptotically fastest known algorithm for testing the existence of a perfect bipartite matching, with running time $O(n^{2.376})$.

But we should honestly admit that a deterministic algorithm is known that always finds a maximum matching in $O(n^{2.5})$ time. This algorithm is much faster in practice. Moreover, the algorithm discussed above can decide whether a perfect matching exists, but it doesn't find one (however, there are more complicated variants that can also find the matching). On the other hand, this algorithm can be implemented very efficiently on a parallel computeri, and no other known approach yields comparably fast parallel algorithms.

Proof of the Schwartz–Zippel theorem. We proceed by induction on m. The univariate case is clear, since there are at most d roots of $p(x_1)$ by a well-known theorem of algebra. (That theorem is proved by induction on d: If $p(\alpha) = 0$, then we can divide $p(x)$ by $x - \alpha$ and reduce the degree.)

Let $m > 1$. Let us suppose that x_1 occurs in at least one term of $p(x_1, \ldots, x_n)$ with a nonzero coefficient (if not, we rename the variables). Let us write $p(x_1, \ldots, x_m)$ as a polynomial in x_1 with coefficients being polynomials in $x_2, \ldots, x_n$:

$$p(x_1, x_2, \ldots, x_m) = \sum_{i=0}^{k} x_1^i p_i(x_2, \ldots, x_m),$$

where k is the maximum exponent of x_1 in $p(x_1, \ldots, x_n)$.

We divide the m-tuples $(r_1, \ldots, r_m)$ with $p(r_1, r_2 \ldots, r_m) = 0$ into two classes. The first class, called R_1, are those with $p_k(r_2, \ldots, r_m) =$

0. Since the polynomial $p_k(x_2, \ldots, x_m)$ is not identically zero and has degree at most $d - k$, the number of choices for $(r_2, \ldots, r_m)$ is at most $(d-k)|S|^{m-2}$ by the induction hypothesis, and so $|R_1| \leq (d-k)|S|^{m-1}$.

The second class R_2 are the remaining m-tuples, that is, those with $p(r_1, r_2, \ldots, r_m) = 0$ but $p_k(r_2, \ldots, r_m) \neq 0$. Here we count as follows: r_2 through r_m can be chosen in at most $|S|^{m-1}$ ways, and if $r_2, \ldots, r_m$ are fixed with $p_k(r_2, \ldots, r_m) \neq 0$, then r_1 must be a root of the univariate polynomial $q(x_1) = p(x_1, r_2, \ldots, r_m)$. This polynomial has degree (exactly) k, and hence it has at most k roots. Thus the second class has at most $k|S|^{m-1}$ m-tuples, which gives $d|S|^{m-1}$ altogether, finishing the induction step and the proof of the Schwartz–Zippel theorem. □

Sources. The idea of the algorithm for testing perfect matchings via determinants is from

> J. Edmonds, *Systems of distinct representatives and linear algebra*, J. Res. Nat. Bur. Standards Sect. B **71B** (1967), 241–245.

There are numerous papers on algebraic matching algorithms; a recent one is

> N. J. A. Harvey, *Algebraic Algorithms for Matching and Matroid Problems*, Proc. 47th IEEE Symposium on Foundations of Computer Science (FOCS), 2006, 531–542.

The Schwartz–Zippel theorem (or lemma) appeared in

> J. Schwartz, *Fast probabilistic algorithms for verification of polynomial identities*, J. ACM **27** (1980), 701–717

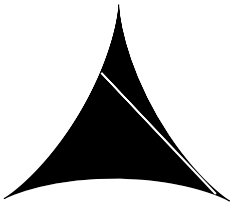
and in

> R. Zippel, *Probabilistic algorithms for sparse polynomials*, Proc. International Symposium on Symbolic and Algebraic Computation, vol. 72 of LNCS, Springer, Berlin, 1979, 216–226.

Miniature 25

Turning a Ladder Over a Finite Field

We want to turn around a ladder of length 10 m inside a garden (without lifting it). What is the smallest area of a garden in which this is possible? For example, here is a garden that, area-wise, looks quite economical (the ladder is drawn as a white segment):

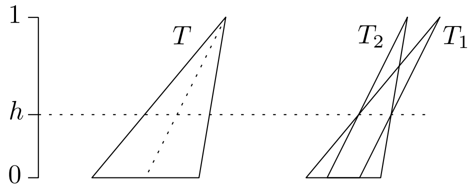


This question is commonly called the **Kakeya needle problem**; Kakeya phrased it with rotating a needle but, while I've never seen any reason for trying to rotate a needle, I did have some quite memorable experiences with turning a long and heavy ladder, so I will stick to this alternative formulation.

One of the fairly counter-intuitive results in mathematics, discovered by Besicovitch in the 1920s, is that there are gardens of *arbitrarily small* area that still allow the ladder to be rotated. Let me sketch the beautiful construction, although it is not directly related to the topic of this book.

A necessary condition for turning a unit-length ladder inside a set X is that X contains a unit-length segment of every direction. An X satisfying this latter, weaker condition is called a **Kakeya set**; unlike the ladder problem, this definition has an obvious generalization to higher dimensions. We begin by constructing a planar Kakeya set of arbitrarily small area (actually, one can get a zero-measure Kakeya set with a little more effort).

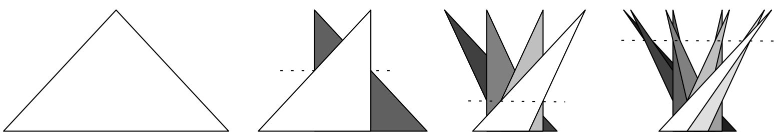
Let us consider a triangle T of height 1 with base on the x-axis, and let $h \in [0, 1)$. The **thinning** of T at height h means replacing T with the two triangles T_1 and T_2 obtaining by slicing T through the top vertex and the middle of its base and translating left T_2 so that it exactly overlaps with T_1 at height h:



More generally, thinning a collection of triangles at height h means thinning each of them separately, so from k triangles we obtain $2k$ triangles.

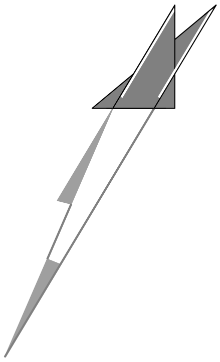
We will construct a small-area set in the plane that contains segments of all directions with slope at least 1 in absolute value (more vertical than horizontal); to get a Kakeya set, we need to add another copy rotated by 90 degrees.

We choose a sequence $(h_1, h_2, h_3, \ldots)$ that is dense in the interval $[0, 1)$ and contains every member infinitely often, e.g., the sequence $(\frac{1}{2}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \frac{1}{8}, \frac{2}{8}, \ldots)$. We start with the triangle with top angle 90 degrees, perform thinning at height h_1, then at height h_2, etc.



Let B_i be the union of all the 2^i triangles after the ith thinning. We claim that the area of B_i gets arbitrarily small as i grows. The idea of the proof is that after k thinnings at height h, the total length of the intersection of the current collection of triangles with the horizontal line of height h is at most 2^{-k} times the original length. Then we need a "continuity" argument, showing that the length is very small not only at height exactly h, but also in a sufficiently large neighborhood. We leave the details to an ambitious reader.

How can we use B_i to turn the ladder? We need to enlarge it so that the ladder can move from one triangle to the next. For that, we add "translation corridors" of the following kind to B_i:



The dark gray triangles are from B_i, and the lighter gray corridor can be used to transport the ladder between the two marked positions. If we're willing to walk with the ladder far enough, then the translation tunnels add an arbitrarily small area.

Kakeya's conjecture. A similar construction produces zero-measure Kakeya sets in all higher dimensions too. However, a statement known as Kakeya's conjecture asserts that they can't be *too* small. Namely, a Kakeya set K in $\mathbb{R}^n$ should have Hausdorff dimension n (for readers not familiar with Hausdorff dimension: roughly speaking, this means that it is not possible to cover K with sets of small diameter much more economically than the n-dimensional cube, say).

While the Kakeya needle problem has a somewhat recreational flavor, Kakeya's conjecture is regarded as a fundamental mathematical question, mainly in harmonic analysis, and it is related to several other serious problems. Although many partial results have been achieved, by the effort of many great mathematicians, the conjecture still seems far from solution (it has been proved only for $n = 2$).

Kakeya for finite fields. Recently, however, an analogue of Kakeya's conjecture, with the field $\mathbb{R}$ replaced by a finite field $\mathbb{F}$, has been settled by a short algebraic argument (after previous, weaker results involving *much* more complicated mathematics). A set K in the vector space $\mathbb{F}^n$ is a **Kakeya set** if it contains a "line" in every possible "direction"; that is, for every nonzero $\mathbf{u} \in \mathbb{F}^n$ there is $\mathbf{a} \in \mathbb{F}^n$ such that $\mathbf{a} + t\mathbf{u}$ belongs to K for all $t \in \mathbb{F}$.

Theorem (Kakeya's conjecture for finite fields)**.** *Let $\mathbb{F}$ be a q-element field. Then any Kakeya set K in $\mathbb{F}^n$ has at least $\binom{q+n-1}{n}$ elements.*

For n fixed and q large, $\binom{q+n-1}{n}$ behaves roughly like $q^n/n!$, so a Kakeya set occupies at least about $\frac{1}{n!}$ of the whole space. Hence, unlike in the real case, a Kakeya set over a finite field occupies a substantial part of the "n-dimensional volume" of the whole space.

The binomial coefficient enters through the following easy lemma.

Lemma. *Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N$ be points in $\mathbb{F}^n$, where $N < \binom{d+n}{n}$. Then there exists a nonzero polynomial $p(x_1, x_2, \ldots, x_n)$ of degree at most d such that $p(\mathbf{a}_i) = 0$ for all i.*

Proof. A general polynomial of degree at most d in variables $x_1, x_2, \ldots, x_n$ can be written as $p(\mathbf{x}) = \sum_{\alpha_1+\cdots+\alpha_n \le d} c_{\alpha_1,\ldots,\alpha_n} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, where the sum is over all n-tuples of nonnegative integers $(\alpha_1, \ldots, \alpha_n)$ summing to at most d, and the $c_{\alpha_1,\ldots,\alpha_n} \in \mathbb{F}$ are coefficients.

We claim that the number of the n-tuples $(\alpha_1, \ldots, \alpha_n)$ as above is $\binom{d+n}{n}$. Indeed, we can think of choosing $(\alpha_1, \ldots, \alpha_n)$ as distributing d identical balls into $n+1$ numbered boxes (the last box is for the $d - \alpha_1 - \cdots - \alpha_n$ "unused" balls). A simple way of seeing that the number of distribution is as claimed is to place the d balls in a row, and then insert n separators among them defining the groups:

○ | ○ ○ ○ | | ○ | ○ ○ |

So among $n + d$ positions for balls and separators we choose the n positions that will be occupied by separators, and the count follows.

A requirement of the form $p(\mathbf{a}) = 0$ translates to a *homogeneous* linear equation with the $c_{\alpha_1,\ldots,\alpha_n}$ as unknowns. Since $N < \binom{n+d}{d}$, we have fewer equations than unknowns, and such a homogeneous system always has a nonzero solution. So there is a polynomial with at least one nonzero coefficient. □

Proof of the theorem. We proceed by contradiction, assuming $|K| < \binom{q+n-1}{n}$. Then by the lemma, there is a nonzero polynomial p of degree $d \le q - 1$ vanishing at all points of K.

Let us consider some nonzero $\mathbf{u} \in \mathbb{F}^n$. Since K is a Kakeya set, there is $\mathbf{a} \in \mathbb{F}^n$ with $\mathbf{a} + t\mathbf{u} \in K$ for all $t \in \mathbb{F}$. Let us define $f(t) := p(\mathbf{a} + t\mathbf{u})$; this is a polynomial in the single variable t of degree at most d. It vanishes for all the q possible values of t, and since a univariate polynomial of degree d over a field has at most d roots, it follows that $f(t)$ is the zero polynomial. In particular, the coefficient of t^d in $f(t)$ is 0.

Now let us see what is the meaning of this coefficient in terms of the original polynomial p: It equals $\overline{p}(\mathbf{u})$, where $\overline{p}$ is the *homogeneous part* of p, i.e., the polynomial obtained from p by omitting all monomials of degree strictly smaller than d. Clearly, $\overline{p}$ is also a nonzero polynomial, for otherwise, the degree of p would be smaller than d.

Hence $\overline{p}(\mathbf{u}) = 0$, and since $\mathbf{u}$ was arbitrary, $\overline{p}$ is 0 on all of $\mathbb{F}^n$. But this contradicts the Schwartz–Zippel theorem from Miniature 24, which implies that a nonzero polynomial of degree d can vanish on at most $dq^{n-1} \le (q-1)q^n < |\mathbb{F}^n|$ points of $\mathbb{F}^n$. The resulting contradiction proves the theorem. □

Sources. Zero-measure Kakeya sets were constructed in

A. Besicovitch, *Sur deux questions d'integrabilite des fonctions*, J. Soc. Phys. Math. **2** (1919), 105–123.

After hearing about Kakeya's needle problem, Besicovitch solved it by modifying his method, in

A. Besicovitch, *On Kakeya's problem and a similar one*, Math. Zeitschrift **27** (1928), 312–320.

There are several simplifications of Besicovitch's original construction (e.g., by Perron and by Schoenberg). The above proof of the Kakeya conjecture for finite fields is from

> Z. Dvir, *On the size of Kakeya sets in finite fields*, J. Amer. Math. Soc. **22** (2009), 1093–1097.

(the above result includes a simple improvement of Dvir's original lower bound, noticed independently by Alon and by Tao).

Miniature 26

Counting Compositions

We consider the following algorithmic problem: P is a given set of permutations of the set $\{1, 2, \ldots, n\}$, and we would like to compute the cardinality of the set $P \circ P := \{\sigma \circ \tau : \sigma, \tau \in P\}$ of all compositions of pairs of permutations from P.

We recall that a **permutation** of $\{1, 2, \ldots, n\}$ is a bijective mapping $\sigma\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. For instance, with $n = 4$, we may have $\sigma(1) = 3$, $\sigma(2) = 2$, $\sigma(3) = 4$, and $\sigma(4) = 1$. It is customary to write a permutation by listing its values in a row; i.e., for our example, we write $\sigma = (3, 2, 4, 1)$. In this way, as an array indexed by $\{1, 2, \ldots, n\}$, a permutation can also be stored in a computer.

Permutations are composed as mappings: In order to obtain the composition $\rho := \sigma \circ \tau$ of two permutations σ and τ, we first apply τ and then σ, i.e., $\rho(i) = \sigma(\tau(i))$. For example, for σ as above and $\tau = (2, 3, 4, 1)$, we have $\sigma \circ \tau = (2, 4, 1, 3)$, while $\tau \circ \sigma = (4, 3, 1, 2) \neq \sigma \circ \tau$. Using the array representation of permutations, the composition can be computed in $O(n)$ time.

As an aside, we recall that the set of all permutations of $\{1, \ldots, n\}$ equipped with the operation of composition forms a group, called the **symmetric group** and denoted by S_n. This is an important object in group theory, both in itself and also because every finite group can be represented as a subgroup of some S_n. The problem of computing

$|P \circ P|$ efficiently is a natural basic question in computational group theory.

How large can $P \circ P$ be? One extreme case is when P forms a subgroup of S_n, and in particular, $\sigma \circ \tau \in P$ for all $\sigma, \tau \in P$—then $|P \circ P| = |P|$. The other extreme is that the compositions are all distinct, i.e., $\sigma_1 \circ \tau_1 \neq \sigma_2 \circ \tau_2$ whenever $\sigma_1, \sigma_2, \tau_1, \tau_2 \in P$ and $(\sigma_1, \tau_1) \neq (\sigma_2, \tau_2)$—then $|P \circ P| = |P|^2$.

A straightforward way of computing $|P \circ P|$ is to compute the composition $\sigma \circ \tau$ for every $\sigma, \tau \in P$, obtaining a list of $|P|^2$ permutations, in $O(|P|^2 n)$ time. In this list, some permutations may occur several times. A standard algorithmic approach to counting the number of *distinct* permutations on such a list is to sort the list lexicographically, and then remove multiplicities by a single pass through the sorted list. With some ingenuity, the sorting can also be done in $O(|P|^2 n)$ time; we will not elaborate on the details since our goal is to discuss another algorithm.

It is not easy to come up with an asymptotically faster algorithm (to appreciate this, of course, the reader may want to try for a while). Yet, by combining tools we have already met in some of the previous miniatures, we can do better, at least if we are willing to tolerate some (negligibly small) probability of error.

To develop the faster algorithm, we first relate the composition of permutations to a scalar product of certain vectors. Let $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ be variables. For a permutation σ, we define the vector $\mathbf{x}(\sigma) := (x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})$; e.g., for $\sigma = (3, 2, 4, 1)$ we have $\mathbf{x}(\sigma) = (x_3, x_2, x_4, x_1)$. Similarly we set $\mathbf{y}(\sigma) := (y_{\sigma(1)}, \ldots, y_{\sigma(n)})$.

Next, we recall that τ^{-1} denotes the **inverse** of the permutation τ, i.e., the unique permutation such that $\tau^{-1}(\tau(i)) = i$ for all i. For $\tau = (2, 3, 4, 1)$ as above, $\tau^{-1} = (4, 1, 2, 3)$.

Now we look at the scalar product

$$\mathbf{x}(\sigma)^T \mathbf{y}(\tau^{-1}) = x_{\sigma(1)} y_{\tau^{-1}(1)} + \cdots + x_{\sigma(n)} y_{\tau^{-1}(n)};$$

this is a polynomial (of degree 2) in the variables $x_1, \ldots, x_n, y_1, \ldots, y_n$. All nonzero coefficients of this polynomial are 1's; for definiteness, let us interpret them as integers. For our concrete σ and τ we have $\mathbf{x}(\sigma)^T\mathbf{y}(\tau^{-1}) = x_3y_4 + x_2y_1 + x_4y_2 + x_1y_3$.

The polynomial $\mathbf{x}(\sigma)^T\mathbf{y}(\tau^{-1})$ as above contains exactly one term with y_1, exactly one term with y_2, etc. (since τ^{-1} is a permutation). What is the term with y_1? We can write it as $x_{\sigma(k)}y_{\tau^{-1}(k)}$, where k is the index with $\tau^{-1}(k) = 1$; that is, $k = \tau(1)$. Therefore, the term with y_1 is $x_{\sigma(\tau(1))}y_1$, and similarly, the term with y_i is $x_{\tau(\sigma(i))}y_i$. So, setting $\rho := \sigma \circ \tau$, we can rewrite

$$\mathbf{x}(\sigma)^T\mathbf{y}(\tau^{-1}) = \sum_{i=1}^{n} x_{\rho(i)}y_i.$$

This shows that the polynomial $\mathbf{x}(\sigma)^T\mathbf{y}(\tau^{-1})$ encodes the composition $\sigma \circ \tau$, in the following sense:

Observation. *Let* $\sigma_1, \sigma_2, \tau_1, \tau_2$ *be permutations of* $\{1, 2, \ldots, n\}$. *Then* $\mathbf{x}(\sigma_1)^T\mathbf{y}(\tau_1^{-1})$ *and* $\mathbf{x}(\sigma_2)^T\mathbf{y}(\tau_2^{-1})$ *are equal (as polynomials) if and only if* $\sigma_1 \circ \tau_1 = \sigma_2 \circ \tau_2$. $\square$

Let $P = \{\sigma_1, \sigma_2, \ldots, \sigma_m\}$ be a set of permutations as in our original problem. Let X be the $n \times m$ matrix whose jth column is the vector $\mathbf{x}(\sigma_j)$, $j = 1, 2, \ldots, m$, and let Y be the $n \times m$ matrix with $\mathbf{y}(\sigma_j^{-1})$ as the jth column. Then the matrix product X^TY has the polynomial $\mathbf{x}(\sigma_i)^T\mathbf{y}(\sigma_j^{-1})$ at position (i, j). In view of the observation above, *the cardinality of the set* $P \circ P$ *equals the number of distinct entries of* X^TY.

It may not be clear why this strange-looking reformulation should be algorithmically any easier than the original problem of computing $|P \circ P|$. However, the Schwartz–Zippel theorem from Miniature 24 and fast matrix multiplication come to our aid.

Let $s := 4m^4$ (later we will see why it is chosen this way), and let $S := \{1, 2, \ldots, s\}$. Our algorithm for computing $|P \circ P|$ is going to work as follows:

(1) Choose integers $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ at random; each a_i and each b_i are chosen from S uniformly at random, and all of these choices are independent.

(2) Set up a matrix A, obtained from X by substituting the integer a_i for the variable x_i, $i = 1, 2, \ldots, n$. Similarly, B is

obtained from Y by replacing each y_i by b_i, $i = 1, 2, \ldots, n$. Compute the product $C := A^T B$.

(3) Compute the number of distinct entries of C (by sorting), and output it as the answer.

Lemma. *The output of this algorithm is never larger than* $|P \circ P|$, *and with probability at least* $\frac{1}{2}$ *it equals* $|P \circ P|$.

Proof. If two entries of $X^T Y$ are equal polynomials, then they also yield equal entries in $A^T B$, and thus the number of distinct entries of $A^T B$ is never larger than $|P \circ P|$.

Next, suppose that the entries at positions (i_1, j_1) and (i_2, j_2) of $X^T Y$ are distinct polynomials. Then their difference is a nonzero polynomial p of degree 2. The Schwartz–Zippel theorem tells us that by substituting independent random elements of S for the variables into p we obtain 0 with probability at most $2/|S| = 1/(2m^4)$.

Hence every two given distinct entries of $X^T Y$ become equal in $A^T B$ with probability at most $1/(2m^4)$. Now $X^T Y$ is an $m \times m$ matrix and thus it definitely cannot have more than m^4 pairs of distinct entries. The probability that *any* pair of distinct entries of $X^T Y$ becomes equal in $A^T B$ is no more than $m^4/(2m^4) = \frac{1}{2}$. So with probability at least $\frac{1}{2}$, the number of distinct entries in $A^T B$ and in $X^T Y$ are the same, and this proves the lemma. □

The lemma shows that the algorithm works correctly with probability at least $\frac{1}{2}$. If we run the algorithm k times and take the largest of the answers, the probability that we don't get $|P \circ P|$ is at most 2^{-k}.

How fast can the algorithm be implemented? For simplicity, let us consider only the case $m = n$, i.e., n permutations of n numbers. We recall that the straightforward algorithm we outlined first needs time of order n^3.

In the randomized algorithm we have just described, the most time-consuming step is the computation of the matrix product $A^T B$. For $m = n$, A and B are square matrices whose entries are integers no larger than $s = 4n^4$, and as we mentioned in Miniature 10, such matrices can in theory be multiplied in time $O(n^{2.376})$. This is a considerable asymptotic gain compared to $O(n^3)$.

Source. R. Yuster, *Efficient algorithms on sets of permutations, dominance, and real-weighted APSP*, Proc. 20th Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2009, pp. 950–957.

Miniature 27

Is It Associative?

In mathematics, one often deals with sets equipped with one or several binary operations. Familiar examples include groups, fields, and rings.

Let us now consider a completely arbitrary **binary operation** "$\odot$" on a set X. Formally, "$\odot$" is an arbitrary mapping $X \times X \to X$. Less formally, every two elements $x, y \in X$ are assigned some element $z \in X$, and this z is denoted by $x \odot y$. Algebraists sometimes study binary operations at this level of generality; a set X together with a completely arbitrary binary operation is called a **groupoid**.

One of the most basic properties of binary operations is **associativity**; the operation "$\odot$" is associative if $(x \odot y) \odot z = x \odot (y \odot z)$ holds for all $x, y, z \in X$. Practically all binary operations in everyday mathematics are associative; multiplication of the Cayley octonions is a honorable exception proving the rule. Once a groupoid is proved to be associative, its social status is immediately upgraded, and it has the right to be addressed **semigroup**.

Here we investigate an algorithmic problem, which might be useful to an algebraist studying finite groupoids and semigroups: Is a given binary operation "$\odot$" on a finite set X associative?

We assume that X has n elements and that "$\odot$" is given by a table with rows and columns indexed by X, where the entry in row x

and column y stores the element $x \odot y$. For $X = \{\heartsuit, \diamondsuit, \spadesuit, \clubsuit\}$, such a table may look as follows:

$\odot$	$\heartsuit$	$\diamondsuit$	$\spadesuit$	$\clubsuit$
$\heartsuit$	$\heartsuit$	$\heartsuit$	$\heartsuit$	$\heartsuit$
$\diamondsuit$	$\heartsuit$	$\diamondsuit$	$\spadesuit$	$\clubsuit$
$\spadesuit$	$\heartsuit$	$\spadesuit$	$\heartsuit$	$\spadesuit$
$\clubsuit$	$\heartsuit$	$\clubsuit$	$\heartsuit$	$\diamondsuit$

Let us call a triple $(x, y, z) \in X^3$ **associative** if $(x \odot y) \odot z = x \odot (y \odot z)$ holds, and **nonassociative** otherwise. An obvious method of checking associativity of "$\odot$" is to test each triple $(x, y, z) \in X^3$. For each triple (x, y, z), we need two lookups in the table to find $(x \odot y) \odot z$ and two more lookups to compute $x \odot (y \odot z)$. Hence the running time of this straightforward algorithm is of order n^3.

We will present an ingenious algorithm with much better running time.

Theorem. *There is a probabilistic algorithm that accepts a binary operation "$\odot$" on an n-element set given by a table, runs for time at most $O(n^2)$, and outputs one of the answers YES or NO. If "$\odot$" is associative, then the answer is always YES. If "$\odot$" is not associative, then the answer can be either YES or NO, but YES is output with probability at most $\frac{1}{2}$.*

The probability of an incorrect answer can be made arbitrarily small by repeating the algorithm sufficiently many times, similar to the algorithm in Miniature 11.

An obvious randomized algorithm for associativity testing would be to repeatedly pick a random triple $(x, y, z) \in X^3$ and to test its associativity. But the catch is that the nonassociativity need not manifest itself on many triples. For example, the operation specified in the above table has only two nonassociative triples, namely $(\clubsuit, \clubsuit, \spadesuit)$ and $(\clubsuit, \spadesuit, \clubsuit)$, while there are $4^3 = 64$ triples altogether. Actually, it is not hard to construct an example of an operation on an n-element set with a single nonassociative triple for every $n \geq 3$. In such case, even if we test n^2 random triples, the chance of detecting nonassociativity is only $\frac{1}{n}$, very far from the constant $\frac{1}{2}$ in the theorem.

The heart of the better algorithm from the theorem is the following mathematical construction. First we fix a suitable field $\mathbb{K}$. We want it to have at least 6 elements, and it is also convenient to assume that the addition and multiplication in $\mathbb{K}$ can be done in constant time. Thus, we can take the 7-element field for $\mathbb{K}$ (with some care, though, we could also implement the algorithm with $\mathbb{K} = \mathbb{R}$ or with many other fields).

We consider the vector space $\mathbb{K}^X$, whose vectors are n-tuples of numbers from $\mathbb{K}$ indexed by the elements of X.

We let $\mathbf{e}\colon X \to \mathbb{K}^X$ be the following mapping: For every $x \in X$, $\mathbf{e}(x)$ is the vector in $\mathbb{K}^X$ that has 1 at the position corresponding to x and 0's elsewhere. Thus $\mathbf{e}$ defines a bijective correspondence of X with the standard basis of $\mathbb{K}^X$.

We now come to the key part of the construction: We define a binary operation "$\boxdot$" on $\mathbb{K}^X$. Informally, it is a linear extension of "$\odot$" to $\mathbb{K}^X$. Two arbitrary vectors $\mathbf{u}, \mathbf{v} \in \mathbb{K}^X$ can be written in the standard basis as

$$\mathbf{u} = \sum_{x \in X} \alpha_x \mathbf{e}(x), \quad \mathbf{v} = \sum_{y \in X} \beta_y \mathbf{e}(y),$$

where the coefficients α_x and β_y are elements of $\mathbb{K}$, uniquely determined by $\mathbf{u}$ and $\mathbf{v}$. To determine $\mathbf{u} \boxdot \mathbf{v}$, we first "multiply out" the parentheses:

$$\mathbf{u} \boxdot \mathbf{v} = \Big(\sum_{x \in X} \alpha_x \mathbf{e}(x)\Big) \boxdot \Big(\sum_{y \in X} \beta_y \mathbf{e}(y)\Big) = \sum_{x,y \in X} \alpha_x \beta_y (\mathbf{e}(x) \boxdot \mathbf{e}(y)).$$

Then we replace each $\mathbf{e}(x) \boxdot \mathbf{e}(y)$ with $\mathbf{e}(x \odot y)$, obtaining

$$\mathbf{u} \boxdot \mathbf{v} = \sum_{x,y \in X} \alpha_x \beta_y \mathbf{e}(x \odot y). \tag{21}$$

The right-hand side is a linear combination of basis vectors, thus a well-defined vector of $\mathbb{K}^X$, and we take it as the definition of $\mathbf{u} \boxdot \mathbf{v}$. Of course, one could define "$\boxdot$" by stating only (21), but the above calculation shows how one arrives at this definition starting from the idea that "$\boxdot$" should be a linear extension of "$\odot$".

It is easy to check that if "$\odot$" is associative, then "$\boxdot$" is associative as well (we leave this to the reader). On the other hand, if (a, b, c)

is a nonassociative triple for "$\odot$", then $(\mathbf{e}(a), \mathbf{e}(b), \mathbf{e}(c))$ is clearly a nonassociative triple for "$\boxdot$".

However, the key feature of this construction is that there are many more nonassociative triples for "$\boxdot$": Even if "$\odot$" has a single nonassociative triple, "$\boxdot$" has very many, and we are quite likely to hit one by a random test, as we will see.

Now we are ready to describe the algorithm for associativity testing. Let us fix a 6-element set $S \subset \mathbb{K}$.

(1) For every $x \in X$, choose elements $\alpha_x, \beta_x, \gamma_x \in S$ uniformly at random, all of these choices independent.

(2) Let us set $\mathbf{u} := \sum_{x \in X} \alpha_x \mathbf{e}(x)$, $\mathbf{v} := \sum_{y \in X} \beta_y \mathbf{e}(y)$, and $\mathbf{w} := \sum_{z \in X} \gamma_z \mathbf{e}(z)$.

(3) Compute the vectors $(\mathbf{u} \boxdot \mathbf{v}) \boxdot \mathbf{w}$ and $\mathbf{u} \boxdot (\mathbf{v} \boxdot \mathbf{w})$. If they are equal, answer YES, and otherwise, answer NO.

Given two arbitrary vectors $\mathbf{u}, \mathbf{v} \in \mathbb{K}^X$, the vector $\mathbf{u} \boxdot \mathbf{v}$ can be computed, following the definition (21), using $O(n^2)$ lookups in the table of the operation "$\odot$" and $O(n^2)$ operations in the field $\mathbb{K}$. If we assume that each operation in $\mathbb{K}$ takes constant time, it is clear that the algorithm can be executed in time $O(n^2)$.

Since "$\boxdot$" is associative for an associative "$\odot$", it is also clear that algorithm always answers YES for an associative operation. For establishing the theorem, it is now sufficient to prove the following claim.

Claim. *If "$\odot$" is not associative and* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *are chosen randomly as in the algorithm, then* $(\mathbf{u} \boxdot \mathbf{v}) \boxdot \mathbf{w} \neq \mathbf{u} \boxdot (\mathbf{v} \boxdot \mathbf{w})$ *with probability at least* $\frac{1}{2}$.

Proof. Let us fix a nonassociative triple $(a, b, c) \in X^3$. Let us consider the random choice of the $\alpha_x, \beta_y, \gamma_z \in S$ in the algorithm, and let us imagine that α_a, β_b, and γ_c are chosen last, after all of the other $\alpha_x, \beta_y, \gamma_z$ have already been fixed. We will actually show that if we fix all $\alpha_x, \beta_y, \gamma_z$, $x \neq a$, $y \neq b$, $z \neq c$ to completely arbitrary values and then choose α_a, β_b, and γ_c at random, the probability of $(\mathbf{u} \boxdot \mathbf{v}) \boxdot \mathbf{w} \neq \mathbf{u} \boxdot (\mathbf{v} \boxdot \mathbf{w})$ is at least $\frac{1}{2}$.

To this end, we show that with probability at least $\frac{1}{2}$, these vectors differ in the component indexed by the element $r := (a \odot b) \odot c$, i.e., $((\mathbf{u} \boxdot \mathbf{v}) \boxdot \mathbf{w})_r \neq (\mathbf{u} \boxdot (\mathbf{v} \boxdot \mathbf{w}))_r$. To emphasize that we treat α_a, β_b, and γ_c as (random) variables, while all the other $\alpha_x, \beta_y, \gamma_z$ are considered constant, we write $f(\alpha_a, \beta_b, \gamma_c) := ((\mathbf{u} \boxdot \mathbf{v}) \boxdot \mathbf{w})_r$, $g(\alpha_a, \beta_b, \gamma_c) := (\mathbf{u} \boxdot (\mathbf{v} \boxdot \mathbf{w}))_r$.

Using the definition of "$\boxdot$", we obtain

$$f(\alpha_a, \beta_b, \gamma_c) = \sum_{x,y,z \in X, (x \odot y) \odot z = r} \alpha_x \beta_y \gamma_z.$$

Thus, $f(\alpha_a, \beta_b, \gamma_c)$ is a polynomial in $\alpha_a, \beta_b, \gamma_c$ of degree at most 3. Since $(a \odot b) \odot c = r$, the monomial $\alpha_a \beta_b \gamma_c$ appears with coefficient 1 (and thus the degree equals 3).

Similarly, we have

$$g(\alpha_a, \beta_b, \gamma_c) = \sum_{x,y,z \in X, x \odot (y \odot z) = r} \alpha_x \beta_y \gamma_z.$$

But now $a \odot (b \odot c) \neq r$ since (a, b, c) is a nonassociative triple, and thus the coefficient of $\alpha_a \beta_b \gamma_c$ in $g(\alpha_a, \beta_b, \gamma_c)$ is 0.

Now we can use the services of our reliable ally, the Schwartz–Zippel theorem from Miniature 24: The difference $f(\alpha_a, \beta_b, \gamma_c) - g(\alpha_a, \beta_b, \gamma_c)$ is a nonzero polynomial of degree 3, and so the probability that substituting independent random elements of S for the variables $\alpha_a, \beta_b, \gamma_c$ yields value 0 is at most $3/|S| = \frac{1}{2}$. Hence, for random $\alpha_a, \beta_b, \gamma_c$ we have $f(\alpha_a, \beta_b, \gamma_c) \neq g(\alpha_a, \beta_b, \gamma_c)$ with probability at least $\frac{1}{2}$. This finishes the proof the claim and also of the theorem. $\square$

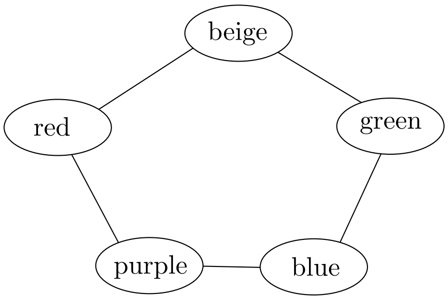
Source. S. Rajagopalan and L. Schulman, *Verification of Identities*, SIAM J. Computing **29,4** (2000), 1155–1163.

Miniature 28

The Secret Agent and the Umbrella

A secret government agent in a desert training camp of a terrorist group has very limited possibilities of sending messages. He has five scarves: red, beige, green, blue, and purple, and he wears one of them with his uniform every day. The analysts at the headquarters then determine the color of his scarf from a satellite photography.

But since the scarves are not really clean, it turned out that certain pairs of colors can't be distinguished reliably. The possibilities of confusion are shown in the next picture:



For example, one cannot reliably tell purple from blue nor from red, but there is no danger to confuse purple with beige or green.

In order to transmit reliably, the agent can, for example, use only the blue and red scarves, and thereby send one of two possible messages every day—one bit in the computer science language. He can communicate one of 2^k possible messages in k days.

Among every three scarves there are some two that can be confused, and so it may seem that there is no chance to send more than one bit per day. But there is a better way! In two successive days, the agent can send one of five messages, e.g., as follows:

	the first day	the second day
message 1	red	red
message 2	beige	green
message 3	green	purple
message 4	blue	beige
message 5	purple	blue

Indeed, there is no chance of mistaking any of these two-day combinations for another, as the reader can easily check. So the agent can transmit one of $5^{k/2} = \sqrt{5}^k$ possible messages in k days (for k even), and the efficiency per day has increased from 2 to $\sqrt{5}$.

Can the efficiency be increased further using three-day or ten-day combinations, say? This is a difficult mathematical problem. The answer is no, and the following masterpiece is the only known proof.

First we formulate the problem in mathematical terms (and generalize it). We consider some **alphabet** S; in our case S consists of the five possible colors of the scarf. Some pairs of symbols can be confused (in other words, are *interchangeable*), and this is expressed by a graph $G = (S, E)$, where the interchangeable pairs of symbols of S are connected by edges. For the situation with five scarves, the graph is drawn in the above picture and it is a cycle of length 5, i.e., C_5.

Let us consider two messages of length k, a message $a_1 a_2 \cdots a_k$ and a message $b_1 b_2 \cdots b_k$. In the terminology of coding theory, these are the words of length k over the alphabet S; see Miniature 5. These messages are interchangeable if and only if a_i is interchangeable with b_i (meaning that $a_i = b_i$ or $\{a_i, b_i\} \in E$) for every $i = 1, 2, \ldots, k$.

Let $\alpha_k(G)$ be the maximum size of a set of messages of length k with no interchangeable pair. In particular, $\alpha_1(G)$ is the maximum size of an **independent set** in G, i.e., a subset of vertices in which no pair is connected by an edge. This quantity is usually denoted by $\alpha(G)$. For our example, we have $\alpha_1(C_5) = \alpha(C_5) = 2$. Our table proves that $\alpha_2(C_5) \geq 5$, and actually equality holds—the inequality $\alpha_2(C_5) \leq 5$ is a very special case of the result we are about to prove.

The **Shannon capacity** of a graph G is defined as follows:

$$\Theta(G) := \sup\left\{\alpha_k(G)^{1/k} : k = 1, 2, \ldots\right\}.$$

It represents the maximum possible efficiency of message transmission per symbol. For a sufficiently large k, the agent can send one from approximately $\Theta(C_5)^k$ possible messages in k days, and not more. We prove the following:

Theorem. $\Theta(C_5) = \sqrt{5}$.

First we observe that $\alpha_k(G)$ can be expressed as the maximum size of an independent set of a suitable graph. The vertex set of this graph is S^k, meaning that the vertices are all possible messages (words) of the length k, and two vertices $a_1a_2\cdots a_k$ and $b_1b_2\cdots b_k$ are connected by an edge if they are interchangeable. We denote this graph by G^k, and we call it the **strong product** of k copies of G.

The strong product $H \cdot H'$ of two arbitrary graphs H and H' is defined as follows:

$$\begin{aligned} V(H\cdot H') &= V(H)\times V(H'),\\ E(H\cdot H') &= \{\{(u,u'),(v,v')\} : (u=v \text{ or } \{u,v\}\in E(H))\\ &\qquad\qquad \text{and at the same time}\\ &\qquad\qquad (u'=v' \text{ or } \{u',v'\}\in E(H'))\}. \end{aligned}$$

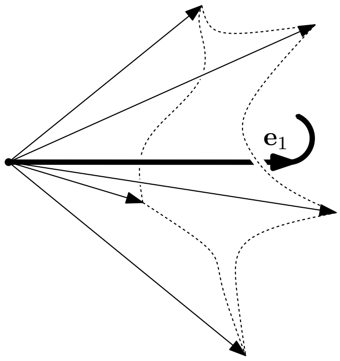
For bounding $\Theta(C_5)$, we thus need to bound above the maximum size of an independent set in each of the graphs C_5^k.

We are going to establish two general results relating independent sets in graphs to certain systems of vectors. Let $H = (V, E)$ be an arbitrary graph. An **orthogonal representation** of H is a mapping $\rho\colon V \to \mathbb{R}^n$, for some n, that assigns a *unit* vector $\rho(v)$ to every vertex $v \in V(H)$ (i.e., $\|\rho(v)\| = 1$), such that the following holds:

> If two distinct vertices u, v are *not connected* by an edge, then the corresponding vectors are *orthogonal.* In symbols, $\{u, v\} \notin E$ implies $\langle \rho(u), \rho(v) \rangle = 0$.

(We use $\langle ., . \rangle$ for the standard scalar product in $\mathbb{R}^n$.)

To prove our main theorem we will need an interesting orthogonal representation ρ_{LU} of the graph C_5 in $\mathbb{R}^3$, the "Lovász umbrella". Let us imagine a collapsed umbrella with five ribs and the tube made of the vector $\mathbf{e}_1 = (1, 0, 0)$. Now we open it slowly until the moment when all pairs of non-neighboring ribs become orthogonal:



At this moment, the ribs define unit vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_5$. When we assign the ith vertex of the graph C_5 to the vector $\mathbf{v}_i$, we get an orthogonal representation ρ_{LU}. It is easy to calculate the opening angle of the umbrella: We obtain $\langle \mathbf{v}_i, \mathbf{e}_1 \rangle = 5^{-1/4}$, which we will soon need.

Any orthogonal representation of a graph G provides an upper bound on $\alpha(G)$:

Lemma A. *If H is a graph and ρ is an orthogonal representations of H, then $\alpha(H) \leq \vartheta(H, \rho)$, where*

$$\vartheta(H, \rho) := \max_{v \in V(H)} \frac{1}{\langle \rho(v), \mathbf{e}_1 \rangle^2}.$$

Proof. Producing an orthogonal representation ρ with $\vartheta(H, \rho)$ minimum geometrically means packing all the unit vectors $\rho(\mathbf{v})$ into a spherical cap centered at $\mathbf{e}_1$ and having the minimum possible radius.

The vectors resist such a packing since pairs corresponding to non-edges must be orthogonal. In particular, the vectors corresponding to an independent set in H form an orthonormal system, and for such a system the minimum cap radius can be calculated exactly.

For a formal proof we need to know that for an arbitrary orthonormal system of vectors $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ in some $\mathbb{R}^n$ and an arbitrary vector $\mathbf{u}$, we have

$$\sum_{i=1}^{m} \langle \mathbf{v}_i, \mathbf{u} \rangle^2 \leq \|\mathbf{u}\|^2.$$

Indeed, the given system $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m)$ can be extended to an orthonormal basis $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ of $\mathbb{R}^n$, by adding $n-m$ other suitable vectors $(\mathbf{v}_{m+1}, \mathbf{v}_{m+2}, \ldots, \mathbf{v}_n)$. The ith coordinate of $\mathbf{u}$ with respect to this basis is $\langle \mathbf{v}_i, \mathbf{u} \rangle$, and we have $\|\mathbf{u}\|^2 = \sum_{i=1}^{n} \langle \mathbf{v}_i, \mathbf{u} \rangle^2$ by the Pythagorean theorem. The required inequality is obtained by omitting the last $n - m$ terms on the right-hand side.

Now if $I \subseteq V(H)$ is an independent set in H, then, as noted above, $(\rho(v) : v \in I)$ forms an orthonormal system, and so

$$\sum_{v \in I} \langle \rho(v), \mathbf{e}_1 \rangle^2 \leq \|\mathbf{e}_1\|^2 = 1.$$

Hence there exists $v \in I$ with $\langle \rho(v), \mathbf{e}_1 \rangle^2 \leq \frac{1}{|I|}$, and thus $\vartheta(H, \rho) \geq |I|$. □

The lemma together with the Lovász umbrella gives

$$\alpha(C_5) \leq \vartheta(C_5, \rho_{\mathrm{LU}}) = \sqrt{5},$$

but this is not an earth-shattering result yet, since everyone knows that $\alpha(C_5) = 2$. We need to complement this with the following lemma, showing that orthogonal representations behave well with respect to the strong product.

Lemma B. *Let H_1, H_2 be graphs, and let ρ_i be an orthogonal representation of H_i, $i = 1, 2$. Then there is an orthogonal representation ρ of the strong product $H_1 \cdot H_2$ such that $\vartheta(H_1 \cdot H_2, \rho) = \vartheta(H_1, \rho_1) \cdot \vartheta(H_2, \rho_2)$.*

Applying Lemma B inductively to the strong product of k copies of C_5, we obtain

$$\alpha(C_5^k) \leq \vartheta(C_5, \rho_{\mathrm{LU}})^k = \sqrt{5}^k,$$

which proves that $\Theta(C_5) \leq \sqrt{5}$ and thus yields the theorem.

Proof of Lemma B. We recall the operation of **tensor product**, already used in Miniature 18. The tensor product of two vectors $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$ is a vector in $\mathbb{R}^{mn}$, denoted by $\mathbf{x} \otimes \mathbf{y}$, with coordinates corresponding to all products $x_i y_j$ for $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$. For example, for $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2)$, we have

$$\mathbf{x} \otimes \mathbf{y} = (x_1y_1, x_2y_1, x_3y_1, x_1y_2, x_2y_2, x_3y_2) \in \mathbb{R}^6.$$

We need the following fact, whose routine proof is left to the reader:

$$\langle \mathbf{x} \otimes \mathbf{y}, \mathbf{x}' \otimes \mathbf{y}' \rangle = \langle \mathbf{x}, \mathbf{x}' \rangle \cdot \langle \mathbf{y}, \mathbf{y}' \rangle \tag{22}$$

for arbitrary $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^m$, $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^n$.

Now we can define an orthogonal representation ρ of the strong product $H_1 \cdot H_2$. The vertices $H_1 \cdot H_2$ are pairs (v_1, v_2), $v_1 \in H_1$, $v_2 \in H_2$. We put

$$\rho((v_1, v_2)) := \rho_1(v_1) \otimes \rho_2(v_2).$$

Using (22) we can easily verify that ρ is an orthogonal representation of $H_1 \cdot H_2$, and the equality $\vartheta(H_1 \cdot H_2, \rho) = \vartheta(H_1, \rho_1) \cdot \vartheta(H_2, \rho_2)$ follows as well. This completes the proof of Lemma B. □

Remarks. The quantity

$$\vartheta(G) = \inf\{\vartheta(G, \rho) : \rho \text{ an orthogonal representation of } G\}$$

is called the **Lovász theta function** of the graph G. As we have seen, it gives an upper bound for $\alpha(G)$, the independence number of the graph. It is not hard to prove that it also provides a lower bound on the **chromatic number** of the complement of the graph G, or in other words, the minimum number of complete subgraphs needed to cover G. Computing the independence number or the chromatic number of a given graph is algorithmically hard (NP-complete), but surprisingly, $\vartheta(G)$ can be computed in polynomial time (more precisely, approximated with arbitrary required precision). Because of this and several other remarkable properties, the theta function is very important.

The Shannon capacity of a graph is a much harder nut. *No algorithm at all*, polynomial or not, is known for computing or approximating it. And we need not go far for an unsolved case—$\Theta(C_7)$ is not known! If the agent had seven scarves, nobody can tell him the best way of transmitting.

Source. L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Th. **IT-25** (1979), 1–7.

Miniature 29

Shannon Capacity of the Union: A Tale of Two Fields

Here we continue with the topic of Miniature 28: the Shannon capacity of a graph. However, for convenience, we will repeat the relevant definitions. Reading the beginning of Miniature 28 may still be useful for understanding the motivation of the Shannon capacity.

So we first recall that if G is a graph, then $\alpha(G)$ denotes the maximum possible size of an independent set in G, that is, of a set $I \subseteq V(G)$ such that no two vertices of I are connected by an edge in G. The **strong product** $H \cdot H'$ of graphs H and H' has vertex set $V(H) \times V(H')$, and two vertices in this vertex set, which we can write as (u, u') and (v, v'), are connected by an edge if $u = v$ or $\{u, v\} \in E(H)$, and at the same time, $u' = v'$ or $\{u', v'\} \in E(H')$.

The **Shannon capacity** of a graph G is denoted by $\Theta(G)$ and defined by

$$\Theta(G) := \sup\left\{\alpha(G^k)^{1/k} : k = 1, 2, \ldots\right\},$$

where G^k stands for the strong product of k copies of G.

The Shannon capacity is quite important in coding theory and graph theorists have been studying it with great interest for a long time, but it remains one of the most mysterious graph parameters. The aim of this section is to prove a surprising result concerning the behavior of the Shannon capacity with respect to the operation of disjoint union of graphs.

Informally, the **disjoint union** of graphs G and H, denoted by $G + H$, is the graph obtained by putting G and H "side by side." A formal definition of the disjoint union is easy if the vertex sets $V(G)$ and $V(H)$ are disjoint; then we can simply set $V(G + H) := V(G) \cup V(H)$ and $E(G + H) := E(G) \cup E(H)$. However, in general G and H may happen to have some vertices in common, or their vertex sets may even coincide. Then we first have to construct an isomorphic copy H' of H such that $V(G) \cap V(H') = \emptyset$, and then we put

$$V(G + H) := V(G) \cup V(H'), \quad E(G + H) := E(G) \cup E(H').$$

(In this way the graph $G + H$ is defined only up to isomorphism, but this is just fine for our purposes.)

It is a nice and not entirely trivial exercise, which we don't even urge the reader to undertake, to prove that

$$\Theta(G + H) \geq \Theta(G) + \Theta(H) \tag{23}$$

for every two graphs G and H.

In the coding-theoretic interpretation from Miniature 28, $\Theta(G)$ is the number of distinguishable messages "per symbol" that can be sent using (arbitrarily long) messages composed of symbols from the alphabet $V(G)$, and similarly for $\Theta(H)$. Then the inequality (23) tells us that if no symbol from the alphabet $V(G)$ can ever be confused with a symbol from $V(H)$, and if we are allowed to send messages composed of symbols from $V(G)$ and from $V(H)$, then the number of distinguishable messages per symbol is at least $\Theta(G) + \Theta(H)$. The reader will probably agree that this sounds quite plausible, if not "intuitively obvious."

However, it may seem equally plausible, or intuitively obvious, that (23) should always hold with equality (and this was conjectured by Shannon). But it has turned out to be false, and this is the result we have announced above as "surprising."

Theorem. *There exist graphs G and H such that* $\Theta(G+H) > \Theta(G)+\Theta(H)$.

For the proof, we will introduce two tools: the first will be used for bounding $\Theta(G+H)$ from below, and the second for bounding $\Theta(G)$ and $\Theta(H)$ from above. The first tool is the following simple lemma.

Lemma. *Let G be a graph on m vertices, and let $\overline{G}$ denote the complement of G, i.e., the graph on the vertex set $V(G)$ in which two distinct vertices u, v are adjacent exactly if they are not adjacent in G. Then*

$$\Theta(G+\overline{G}) \geq \sqrt{2m}.$$

Proof of the lemma. By the definition of the Shannon capacity, it suffices to find an independent set of size $2m$ in the strong product $(G+\overline{G})^2$.

Let $v_1, v_2, \ldots, v_m$ be the vertices of G and let $v'_1, v'_2, \ldots, v'_m$ be the vertices of the isomorphic copy of $\overline{G}$ used in forming the disjoint union $G+\overline{G}$. We set $I := \{(v_1, v'_1), (v_2, v'_2), \ldots, (v_m, v'_m)\} \cup \{(v'_1, v_1), (v'_2, v_2), \ldots, (v'_m, v_m)\}$. Then I is independent in $(G+\overline{G})^2$: Indeed, (v_i, v'_i) is not adjacent to (v'_j, v_j) since v_i and v'_j are nonadjacent in $G+\overline{G}$, and (v_i, v'_i) and (v_j, v'_j), $i \neq j$, are not adjacent since either v_i and v_j are not adjacent in G or v'_i and v'_j are not adjacent in (the isomorphic copy of) $\overline{G}$. The lemma is proved. □

Functional representations. The second tool, which will be used for bounding $\Theta(.)$ from above, is algebraic. Let $\mathbb{K}$ be a field (such as $\mathbb{R}$, $\mathbb{Q}$, or $\mathbb{F}_2$), and let $G = (V, E)$ be a graph. A **functional representation** $\mathcal{F}$ of G over $\mathbb{K}$ consists of the following:

- A ground set X (an arbitrary set, not necessarily related to G or $\mathbb{K}$ in any way),
- an element $c_v \in X$ for every vertex $v \in V$, and
- a function $f_v\colon X \to \mathbb{K}$ for every vertex $v \in V$.

These objects have to satisfy

(i) $f_v(c_v) \neq 0$ for every $v \in V$, and

(ii) if u, v are distinct and *non-adjacent* vertices of G, then $f_u(c_v) = 0$. (If u and v are adjacent, then $f_u(c_v)$ can be anything.)

We write $\mathcal{F} = (X, (c_v, f_v)_{v\in V})$.

The **dimension** $\dim \mathcal{F}$ of a functional representation $\mathcal{F}$ is the dimension of the subspace generated by all the functions f_v, $v \in V$, in the vector space $\mathbb{K}^X$ of all functions $X \to \mathbb{K}$. (As usual, the functions are added pointwise, $(f+g)(x) = f(x)+g(x)$, and similarly for multiplication by elements of $\mathbb{K}$.)

Proposition. *If G has a functional representation of dimension d over some field $\mathbb{K}$, then $\Theta(G) \le d$.*

The proof of this proposition follows immediately from the definition of the Shannon capacity and the next two claims.

Claim A. *If G has a functional representation $\mathcal{F} = (X, (c_v, f_v)_{v\in V})$ over some field $\mathbb{K}$, then $\alpha(G) \le \dim \mathcal{F}$.*

Claim B. *Suppose that a graph $G = (V, E)$ has a functional representation $\mathcal{F}$ over some field $\mathbb{K}$ and $G' = (V', E')$ has a functional representation $\mathcal{F}'$ over the same $\mathbb{K}$. Then the strong product $G \cdot G'$ has a functional representation over $\mathbb{K}$ of dimension at most $(\dim \mathcal{F})(\dim \mathcal{F}')$.*

Proof of Claim A. It suffices to show that whenever $I \subseteq V(G)$ is an independent set, the functions f_v, $v \in I$, are linearly independent.

This is done in a fairly standard way. We suppose that

$$\sum_{v\in I} t_v f_v = 0 \tag{24}$$

for some scalars t_v, $v \in I$ (the 0 on the right-hand side is the function assigning 0 to every $x \in X$). We fix $u \in V$ and evaluate the left-hand side of (24) at c_u. Since no two distinct $u, v \in I$ are connected by an edge, we have $f_v(c_u) = 0$ for $v \ne u$, and we obtain $\sum_{v\in I} \alpha_v f_v(c_u) = \alpha_u f_u(c_u)$. Since $f_u(c_u) \ne 0$, we have $t_u = 0$, and since u was arbitrary, the f_v are linearly independent as claimed. □

Proof of Claim B. We are going to define a functional representation $\mathcal{G}$ of $G \cdot G'$ in a quite natural way (it can be regarded as a "tensor product" of $\mathcal{F}$ and $\mathcal{F}'$). Let $\mathcal{F} = (X, (c_v, f_v)_{v \in V})$ and $\mathcal{F}' = (X', (c'_{v'}, f'_{v'})_{v' \in V'})$. The ground set of $\mathcal{G}$ is $X \times X'$. A vertex of $G \cdot G'$ has the form $(v, v') \in V \times V'$, and we complete the definition of $\mathcal{G}$ by setting

$$c_{(v,v')} = (c_v, c'_{v'}) \in X \times X' \quad f_{(v,v')} := f_v \otimes f'_{v'},$$

where $f_v \otimes f'_{v'}$ stands for the function $X \times X' \to \mathbb{K}$ defined by $f_v \otimes f'_{v'}(x, x') := f_v(x) f'_{v'}(x')$. It is straightforward to check that this $\mathcal{G}$ indeed satisfies the axioms (i) and (ii) of a functional representation (and we leave it to the reader).

It remains to verify $\dim \mathcal{G} \leq (\dim \mathcal{F})(\dim \mathcal{F}')$, which is equally straightforward: If all the f_v are linear combinations of some basis functions $b_1, \ldots, b_d$, and the $f'_{v'}$ are linear combinations of $b'_1, \ldots, b'_{d'}$, then it is almost obvious that each $f_v \otimes f'_{v'}$ is a linear combination of functions of the form $b_i \otimes b'_j$, $i = 1, 2, \ldots, d$, $j = 1, 2, \ldots, d'$. (It can be checked that $\dim \mathcal{G}$ actually equals $(\dim \mathcal{F})(\dim \mathcal{F}')$). □

Proof of the theorem. It remains to exhibit suitable graphs G and H and apply the tools above. Several constructions are known, and some of them show that $\Theta(G+H)$ can actually be *much* larger than $\Theta(G) + \Theta(H)$. Here, for simplicity, we present only a single very concrete construction, for which $\Theta(G+H)$ is only "somewhat larger" than $\Theta(G) + \Theta(H)$.

We let s be an integer parameter; later on we will calculate that for proving the theorem it suffices to set $s = 16$. The vertices of G are all 3-element subsets of the set $\{1, 2, \ldots, s\}$, and two such vertices A and B are connected by an edge of G if $|A \cap B| = 1$. (Graphs of this kind, where the vertices are sets and the edges are defined based on the cardinality of the intersection, serve as very interesting examples for many graph-theoretic questions.)

The graph H is the complement $\overline{G}$ of G.

First of all, G has $\binom{s}{3}$ vertices, and so $\Theta(G + \overline{G}) \geq \sqrt{2\binom{s}{3}}$ by the lemma.

Now we define suitable functional representations. For G, we use the field $\mathbb{F}_2$, and we let the ground set X be $\mathbb{F}_2^s$, so that its elements are s-component vectors of 0's and 1's. For a vertex (3-element set) $A \in$

$V(H)$, we let $\mathbf{c}_A$ be the characteristic vector of A; that is, $(\mathbf{c}_A)_i = 1$ for $i \in A$ and $(\mathbf{c}_A)_i = 0$ for $i \notin A$. Finally, the function $f_A\colon \mathbb{F}_2^s \to \mathbb{F}_2$ is given by $f_A(\mathbf{x}) = \sum_{i\in A} x_i$ (addition in $\mathbb{F}_2$, i.e., modulo 2).

To see that this is indeed a functional representation of G, we observe that $f_A(\mathbf{c}_B)$ equals $|A\cap B|$ modulo 2. In particular, $f_A(\mathbf{c}_A) = 1 \neq 0$. Now if $A \neq B$ are not adjacent in G, then $|A \cap B|$ can be 2 or 0, and so $f_A(\mathbf{c}_B) = 0$ in this case.

The dimension of this functional representation is at most s, since each f_A is a linear combination of the coordinate functions $\mathbf{x} \mapsto x_i$. Therefore, $\Theta(G) \leq s$.

For $\overline{G}$ we use a very similar functional representation, but over a different field, say $\mathbb{R}$ (or any other field of characteristic distinct from 2). Namely, we let $X' := \mathbb{R}^s$, $\mathbf{c}'_A$ is again the characteristic vector of A (interpreted as a real vector this time), and $f'_A(\mathbf{x}) := \bigl(\sum_{i\in A} x_i\bigr) - 1$. Now $f'_A(\mathbf{c}'_B) = |A \cap B| - 1$, and so $f'_A(\mathbf{c}'_A) = 2 \neq 0$, while for $A \neq B$ non-adjacent in $\overline{G}$ we have $|A \cap B| = 1$ and $f'_A(\mathbf{c}'_B) = 0$, as needed. The dimension is at most $s + 1$ this time (in addition to the coordinate functions $\mathbf{x} \mapsto x_i$ we also need the constant function -1 in the basis), and so $\Theta(\overline{G}) \leq s + 1$.

The proof of the theorem is finished by choosing s sufficiently large so that $\sqrt{2\binom{s}{3}} > 2s+1$. A numerical calculation shows that the smallest suitable s is 16. Then the graphs G and $\overline{G}$ have 560 vertices each. □

Remark. It is interesting to compare the functional representations treated here with the orthogonal representations discussed in Miniature 28. These notions, and the proofs that they both yield upper bounds for $\Theta(G)$, are basically similar. However, functional representations can yield only bounds that are integers, and thus they cannot establish, e.g., that $\Theta(C_5) \leq \sqrt{5}$. On the other hand, orthogonal representations don't appear suitable for the proof in the present section, since the use of two different fields in it is essential, as we will illustrate next.

Indeed, reasoning similar to that in the proof of the lemma shows that $\alpha(G \cdot \overline{G}) \geq m$ for every m-vertex graph G. Thus, if $\mathcal{F}$ is a functional representation of G and $\mathcal{F}'$ is a functional representation of $\overline{G}$

over the same field, we have $(\dim \mathcal{F})(\dim \mathcal{F}') \geq m$ by Claims A and B. Consequently, $\dim \mathcal{F} + \dim \mathcal{F}' \geq 2\sqrt{m}$ (by the inequality between the arithmetic and geometric means), and thus functional representations over the same field can never give an upper bound for $\Theta(G) + \Theta(\overline{G})$ smaller than $\sqrt{2m}$, which is what the lemma yields for $\Theta(G + \overline{G})$.

Source. N. Alon, *The Shannon capacity of a union*, Combinatorica **18** (1998), 301–310.

Our presentation achieves a weaker result and is slightly simpler.

Miniature 30

Equilateral Sets

An **equilateral set** in $\mathbb{R}^d$ is a set of points $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ such that all pairs $\mathbf{p}_i, \mathbf{p}_j$ of distinct points have the same distance.

Intentionally we haven't said what distance we mean: This will play a key role in this section. If one considers the most usual *Euclidean* distance, then it is not too hard to prove that an equilateral set in $\mathbb{R}^d$ can have $d+1$ points but no more.

As an aside, let us sketch the classical proof that there can't be more than $d+1$ points; it proceeds in a way very similar to Miniature 6: Let the points be $\mathbf{p}_1$ through $\mathbf{p}_{n+1}$, translate them so that $\mathbf{p}_{n+1} = \mathbf{0}$, re-scale so that the interpoint distances are 1, and set up the matrix (the Gram matrix) G with $g_{ij} = \langle \mathbf{p}_i, \mathbf{p}_j \rangle$ (scalar product). Using the equilaterality condition one calculates that $G = \frac{1}{2}(I_n + J_n)$, where I_n is the identity matrix and J_n is the all 1's matrix, and thus $\operatorname{rank}(G) = n$. On the other hand, we have $G = P^T P$, where P is the $d \times n$ matrix with the vector $\mathbf{p}_i$ as the ith column, and thus $\operatorname{rank}(G) \leq d$, which gives $n \leq d$.

(And, by the way, how do we prove rigorously that a $(d+1)$-point equilateral set is possible? We can take, e.g., the vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d$ of the standard basis plus the point $(-t, -t, \ldots, -t)$ for a suitable $t > 0$—even if we are too lazy to calculate the right t, it is easy to prove its existence.)

Approximately equilateral sets. We will now relax the condition that all interpoint distances must be exactly the same: We will require them to be only approximately the same. With a sufficiently strong notion of "approximately the same" we will show that the size of such an approximately equilateral set in $\mathbb{R}^d$ is bounded by a constant multiple of d. The proof relies on a neat linear algebra trick. We will then use the result in the proof of the main theorem of this section.

Rank Lemma. *Let A be a real symmetric $n\times n$ matrix, not equal to the zero matrix. Then*

$$\operatorname{rank}(A) \geq \frac{\left(\sum_{i=1}^n a_{ii}\right)^2}{\sum_{i,j=1}^n a_{ij}^2}.$$

Proof. Linear algebra teaches us that A as in the lemma has n real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. If $\operatorname{rank}(A) = r$, then exactly r of the λ_i are nonzero; we may suppose that $\lambda_i \neq 0$ for $1 \leq i \leq r$, while $\lambda_i = 0$ for $i > r$.

Let us write down the Cauchy–Schwarz inequality $(\sum_{i=1}^r x_i y_i)^2 \leq (\sum_{i=1}^r x_i^2)(\sum_{i=1}^r y_i^2)$ for $x_i = \lambda_i$, $y_i = 1$; we obtain $(\sum_{i=1}^r \lambda_i)^2 \leq r \sum_{i=1}^r \lambda_i^2$. Dividing by $\sum_{i=1}^r \lambda_i^2$ yields the following inequality for the rank in terms of eigenvalues:

$$\operatorname{rank}(A) \geq \frac{\left(\sum_{i=1}^n \lambda_i\right)^2}{\sum_{i=1}^n \lambda_i^2}. \tag{25}$$

(We have extended the summation all the way to n, since λ_{r+1} through λ_n are 0's.)

The last inequality can be converted to the inequality in the Rank Lemma in three easy steps: First, the sum of all eigenvalues of A equals the **trace** of A, i.e. $\sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii}$ (a standard linear algebra fact). This takes care of the numerator in (25). Second, the eigenvalues of A^2 are $\lambda_1^2, \ldots, \lambda_n^2$, as one can recall or immediately check, and thus $\sum_{i=1}^n \lambda_i^2 = \operatorname{trace}(A^2)$. Third, $\operatorname{trace}(A^2) = \sum_{i,j=1}^n a_{ij}^2$, as one easily calculates. This brings the denominator into the desired form. $\square$

Corollary (A small perturbation of I_n has a large rank)**.** *Let A be a symmetric $n \times n$ matrix with $a_{ii} = 1$, $i = 1, 2, \ldots, n$, and $|a_{ij}| \leq 1/\sqrt{n}$ for all $i \neq j$. Then* $\operatorname{rank}(A) \geq \frac{n}{2}$. □

Proposition (on approximately equilateral sets)**.** *Let $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n \in \mathbb{R}^d$ be points such that for every $i \neq j$ we have*

$$1 - \frac{1}{\sqrt{n}} \leq \|\mathbf{p}_i - \mathbf{p}_j\|^2 \leq 1 + \frac{1}{\sqrt{n}}.$$

Then $n \leq 2(d+2)$. (Note that, for technical convenience, we bound the squared *Euclidean distances.)*

Proof. Let A be the $n \times n$ matrix with $a_{ij} = 1 - \|\mathbf{p}_i - \mathbf{p}_j\|^2$. The assumptions of the proposition immediately give that A meets the assumptions of the above corollary, and thus $\operatorname{rank}(A) \geq \frac{n}{2}$.

It remains to bound $\operatorname{rank}(A)$ from above in terms of d. Here we proceed as in Miniature 15. For $i = 1, 2, \ldots, n$ let $f_i \colon \mathbb{R}^d \to \mathbb{R}$ be the function defined by $f_i(\mathbf{x}) = 1 - \|\mathbf{x} - \mathbf{p}_i\|^2$; so the ith row of A is $\bigl(f_i(\mathbf{p}_1), f_i(\mathbf{p}_2), \ldots, f_i(\mathbf{p}_n)\bigr)$.

We rewrite $f_i(\mathbf{x}) = 1 - \|\mathbf{x}\|^2 - \|\mathbf{p}_i\|^2 + 2(p_{i1}x_1 + p_{i2}x_2 + \cdots + p_{id}x_d)$, where p_{ik} is the kth coordinate of $\mathbf{p}_i$. Then it becomes clear that each f_i is a linear combination of the following $d+2$ functions: the constant function 1, the function $\mathbf{x} \mapsto \|\mathbf{x}\|^2$, and the "coordinate functions" $\mathbf{x} \mapsto x_k$, $k = 1, 2, \ldots, d$. Hence the vector space generated by the f_i has dimension at most $d+2$, and so has the column space of A. Thus $\operatorname{rank}(A) \leq d + 2$, and the proposition is proved. □

Other kinds of distance. Equilateral sets become much more puzzling if one considers other notions of distance in $\mathbb{R}^d$.

First, as a cautionary tale, let us consider the ℓ_∞ ("el infinity") distance, where the distance of two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ is defined as $\|\mathbf{x} - \mathbf{y}\|_\infty = \max\{|x_i - y_i| : i = 1, 2, \ldots, d\}$. Then the "cube" $\{0, 1\}^d$ is an equilateral set with as many as 2^d points! (Which turns out to be the largest possible example in $\mathbb{R}^d$ with the ℓ_∞ distance, but this is not the story we want to narrate here.)

The distance we really want to focus on here is the **ℓ_1 distance**, given by

$$\|\mathbf{x}-\mathbf{y}\|_1 = |x_1-y_1| + |x_2-y_2| + \cdots + |x_d-y_d|.$$

Then the following is an example of an equilateral set with $2d$ points: $\{\mathbf{e}_1, -\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_2, \ldots, \mathbf{e}_d, -\mathbf{e}_d\}$. A widely believed conjecture states that this is as many as one can ever get, but until about 2001, no upper bound better than 2^d-1 (exponential!) was known.

We will present an ingenious proof of a polynomial upper bound, $O(d^4)$. The proof of the current best bound, $O(d\log d)$, uses a number of additional ideas and it is considerably more technical.

Theorem. *For every $d \ge 1$, no equilateral set in $\mathbb{R}^d$ with the ℓ_1 distance has more than $100d^4$ points.*

The main reason why for the ℓ_1 distance one can't imitate the proof for the Euclidean case sketched above or something similar seems to be this: The functions $\varphi_a\colon \mathbb{R}\to\mathbb{R}$, $\varphi_a(x) = |x-a|$, $a\in\mathbb{R}$, are all linearly independent—unlike the functions $\psi_a(x) = (x-a)^2$ that generate a vector space of dimension only 3.

The forthcoming proof has an interesting twist: In order to establish a bound on *exactly* equilateral sets for the "unpleasant" ℓ_1 distance, we use *approximately* equilateral sets but for the "pleasant" Euclidean distance. Here is a tool for such a passage.

Lemma (on approximate embedding). *For every two natural numbers d, q there exists a mapping $f_{d,q}\colon [0,1]^d \to \mathbb{R}^{dq}$ such that for every $\mathbf{x}, \mathbf{y} \in [0,1]^d$*

$$\|\mathbf{x}-\mathbf{y}\|_1 - \tfrac{2d}{q} \le \tfrac{1}{q}\|f_{d,q}(\mathbf{x}) - f_{d,q}(\mathbf{y})\|^2 \le \|\mathbf{x}-\mathbf{y}\|_1 + \tfrac{2d}{q}.$$

Let us stress that we take *squared* Euclidean distances in the target space. If we wanted instead that the ℓ_1 distance $\|\mathbf{x}-\mathbf{y}\|_1$ be reasonably close to the Euclidean distance of the images for all $\mathbf{x}, \mathbf{y}$, the task becomes impossible.

Our proof of the lemma is somewhat simple-minded. By more sophisticated methods one can reduce the dimension of the target space considerably, and this is also how the d^4 bound in the theorem can be improved.

Proof of the lemma. First we consider the case $d = 1$. For $x \in [0, 1]$, $f_{1,q}(x)$ is the q-component zero/one vector starting with a segment of $\lfloor qx \rfloor$ ones, followed by $q - \lfloor qx \rfloor$ zeros. Then $\|f_{1,q}(x) - f_{1,q}(y)\|^2$ is the number of position where one of $f_{1,q}(x)$, $f_{1,q}(y)$ has 1 and the other 0, and thus it equals $\Bigl|\lfloor qx \rfloor - \lfloor qy \rfloor\Bigr|$. This differs from $q|x - y|$ by at most 2, and we are done with the $d = 1$ case.

For larger d, $f_{d,1}(\mathbf{x})$ is defined as the dq-component vector obtained by concatenating $f_{1,q}(x_1)$, $f_{1,q}(x_2)$,..., $f_{1,q}(x_d)$. The error bound is obvious using the 1-dimensional case. $\square$

Proof of the theorem. For contradiction, let us assume that there exists an equilateral set in $\mathbb{R}^d$ with the ℓ_1 distance that has at least $100d^4$ points. After possibly discarding some points we may assume that it has exactly $n := 100d^4$ points.

We re-scale the set so that the interpoint distances become $\frac{1}{2}$, and we translate it so that one of the points is $(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2})$. Then the set is fully contained in $[0, 1]^d$.

We use the lemma on approximate embedding with $q := 40d^3$. Applying the mapping $f_{d,q}$ to our set, we obtain an n-point set in $\mathbb{R}^{qd}$, for which the squared Euclidean distance of every two points is between $\frac{q}{2} - 2d$ and $\frac{q}{2} + 2d$. After re-scaling by $\sqrt{2/q}$, we get an approximately equilateral set with squared Euclidean interpoint distances between $1 - \frac{4d}{q}$ and $1 + \frac{4d}{q}$. We have $\frac{4d}{q} = \frac{1}{10d^2} = \frac{1}{\sqrt{n}}$, and thus the proposition on approximately equilateral sets applies and shows that $n \leq 2(dq + 2)$. But this is a contradiction, since $n = 100d^4$, while $2(dq + 2) = 2(40d^4 + 2) < 100d^4$. The theorem is proved. $\square$

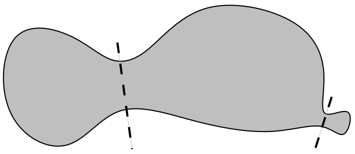
Source. N. Alon and P. Pudlák, *Equilateral sets in* l_p^n, Geometric and Functional Analysis **13** (2003), 467–482.

Our presentation via an approximate embedding is slightly different.

Miniature 31

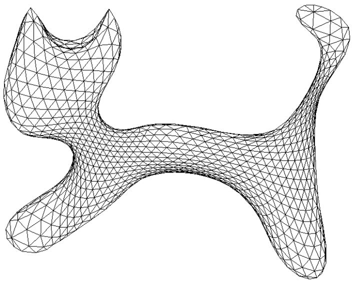
Cutting Cheaply Using Eigenvectors

In many practical applications we are given a large graph G and we want to cut off a piece of the vertex set by removing as few edges as possible. For a large piece we can afford to remove more edges than for a small one, as the next picture schematically indicates:



We can imagine that removing an edge costs one unit and we want to cut off some vertices, at most half of all vertices, at the smallest possible price *per vertex*.

This problem is closely related to the *divide and conquer* paradigm in algorithms design. For example, in areas like computer graphics, computer-aided design, or medical image processing we may have a two-dimensional surface represented by a *triangular mesh*:



For various computations we often need to divide a large mesh into smaller parts that are interconnected as little as possible.

Or more abstractly, the vertices of the graph G may correspond to some objects, edges may express dependences or interactions, and again we would like to partition the problem into smaller subproblems with few mutual interactions.

Sparsest cut. Let us state the problem more formally. Let G be a given graph with vertex set V, $|V| = n$, and edge set E. Let us call a partition of V into two subsets A and $V \setminus A$, with both A and $V \setminus A$ nonempty, a **cut**, and let $E(A, V \setminus A)$ stand for the set of all edges in G connecting a vertex of A to a vertex of $V \setminus A$.

The "price per vertex" for cutting of A, alluded to above, can be defined as $\Phi(A, V \setminus A) := |E(A, V \setminus A)|/|A|$, assuming $|A| \leq \frac{n}{2}$. We will work with a different but closely related quantity: We define the **density** of the cut $(A, V \setminus A)$ as

$$\phi(A, V \setminus A) := n \cdot \frac{|E(A, V \setminus A)|}{|A| \cdot |V \setminus A|}$$

(this is n times the ratio of the number of edges connecting A and $V \setminus A$ in G and in the complete graph on V). Since $|A| \cdot |V \setminus A|$ is between $\frac{1}{2}n|A|$ and $n|A|$ (again with $|A| \leq \frac{n}{2}$), we always have $\Phi(A, V \setminus A) \leq \phi(A, V \setminus A) \leq 2\Phi(A, V \setminus A)$. So it doesn't make much of a difference if we look for a cut minimizing Φ or one minimizing ϕ, and we will stick to the latter.

Thus, let ϕ_G denote the smallest possible density of a cut in G. We would like to compute a **sparsest cut**, i.e., a cut of density ϕ_G.

This problem is known to be computationally difficult (NP-complete), and various approximation algorithms have been proposed for it. One of such algorithms, or rather a class of algorithms, called *spectral partitioning*, is based on eigenvectors of a certain matrix associated with the graph. It is widely and successfully used in practice, and thanks to modern methods for computing eigenvalues, it is also quite fast even for large graphs.

Before we proceed with formulating the algorithm, a remark is in order. In some applications, a sparsest cut is not really what we are interested in—we want a sparse cut that is also *approximately balanced*, i.e., it cuts off at least $\frac{1}{3}$ of all vertices (say). To this end, we can use a sparsest cut algorithm iteratively: We cut off pieces, possibly small ones, repeatedly until we have accumulated at least $\frac{1}{3}$ of all vertices. It can be shown that with a good sparsest cut algorithm this strategy leads to a good approximately balanced cut. We will not elaborate on the details, since this would distract us from the main topic.

Now we can begin with preparations for the algorithm.

The Laplace matrix. For notational convenience let us assume that the vertices of G are numbered $1, 2, \ldots, n$. We define the **Laplace matrix** L of G (also used in Miniature 21) as the $n\times n$ matrix with entries ℓ_{ij} given by

$$\ell_{ij} := \begin{cases} \deg(i) & \text{if } i = j, \\ -1 & \text{if } \{i,j\} \in E(G), \\ 0 & \text{otherwise,} \end{cases}$$

where $\deg(i)$ is the number of neighbors (degree) of i in G.

We will need the following identity: For every $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x}^T L\mathbf{x} = \sum_{\{i,j\}\in E} (x_i - x_j)^2. \tag{26}$$

Indeed, we have

$$\mathbf{x}^T L\mathbf{x} = \sum_{i,j=1}^{n} \ell_{ij}x_ix_j = \sum_{i=1}^{n} \deg(i)x_i^2 - 2 \sum_{\{i,j\}\in E} x_ix_j,$$

the right-hand side simplifies to $\sum_{\{i,j\}\in E}(x_i-x_j)^2$, and so (26) holds.

The right-hand side of (26) is always nonnegative, and thus L is positive semidefinite. So it has n nonnegative real eigenvalues, which we write nondecreasing order as $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_n$.

Since the row sums of L are all 0, we have $L\mathbf{1} = \mathbf{0}$ (where $\mathbf{1}$ is the vector of all 1's), and thus $\mu_1 = 0$ is an eigenvalue with eigenvector $\mathbf{1}$. The key role in the forthcoming algorithm, as well as in many other graph problems, is played by the second eigenvalue μ_2 (sometimes called the *Fiedler value* of G).

Spectral partitioning. The algorithm for finding a sparse cut works as follows.

(1) Given a graph G, compute an eigenvector $\mathbf{u}$ belonging to the second smallest eigenvalue μ_2 of the Laplace matrix.

(2) Sort the components of $\mathbf{u}$ descendingly. Let π be a permutation such that $u_{\pi(1)} \geq u_{\pi(2)} \geq \cdots \geq u_{\pi(n)}$.

(3) Set $A_k := \{\pi(1), \pi(2), \ldots, \pi(k)\}$. Among the cuts $(A_k, V \setminus A_k)$, $k = 1, 2, \ldots, n-1$, output one with the smallest density.

Theorem. *The following hold for every graph G:*
(i) $\phi_G \geq \mu_2$.
(ii) *The algorithm always finds a cut of density at most $4\sqrt{d_{\max}\mu_2}$, where $d_{\max}$ is the maximum vertex degree in G. In particular, $\phi_G \leq 4\sqrt{d_{\max}\mu_2}$.*

Remarks. This theorem is a fundamental result, whose significance goes far beyond the spectral partitioning algorithm. For instance, it is a crucial ingredient in constructions of expander graphs.[1]

The constant 4 in (ii) can be improved by doing the proof more carefully. There can be a large gap between the upper bound for ϕ_G in (i) and the lower bound in (ii), but both of the bounds are essentially tight in general. That is, for some graphs the lower bound is more or less the truth, while for others the upper bound is attained.

For planar graphs of degree bounded by a constant, such as the cat mesh depicted above, it is known that $\mu_2 = O(\frac{1}{n})$ (a proof is beyond the scope of this text), and thus the spectral partitioning algorithm always finds a cut of density $O(n^{-1/2})$. This density is the smallest possible, up to a constant factor, for many planar graphs (e.g., consider the $n \times n$ square grid). Thus, one can say that "spectral partitioning works" for planar graphs of bounded degree. Similar results are known for several other graph classes.

---

[1] Part (ii) is often called the *Cheeger–Alon–Milman inequality*, where Cheeger's inequality is an analogous "continuous" result in the geometry of Riemannian manifolds.

Proof of part (i) of the theorem. Let us say that a vector $\mathbf{x} \in \mathbb{R}^n$ is **nonconstant** if it is not a multiple of $\mathbf{1}$.

For a nonconstant $\mathbf{x} \in \mathbb{R}^n$ let us put

$$Q(\mathbf{x}) := n \cdot \frac{\sum_{\{i,j\} \in E} (x_i - x_j)^2}{\sum_{1 \le i < j \le n} (x_i - x_j)^2}.$$

First let $(A, V \setminus A)$ be a cut in G and let $\mathbf{c}_A$ be the characteristic vector of A (with the ith component 1 for $i \in A$ and 0 otherwise). Then $Q(\mathbf{c}_A)$ is exactly the density of $(A, V \setminus A)$, and so ϕ_G is the minimum of $Q(\mathbf{x})$ over all nonconstant vectors $\mathbf{x} \in \{0,1\}^n$.

Next, we will show that μ_2 is the minimum of $Q(\mathbf{x})$ over a larger set of vectors, namely,

$$\mu_2 = \min\{Q(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n \text{ nonconstant}\} \tag{27}$$

(computer scientists would say that μ_2 is a *relaxation* of ϕ_G). This, of course, implies $\phi_G \ge \mu_2$.

Since $Q(\mathbf{x}) = Q(\mathbf{x} + t\mathbf{1})$ for all $t \in \mathbb{R}$, we can change (27) to

$$\mu_2 = \min\{Q(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}, \langle \mathbf{x}, \mathbf{1} \rangle = 0\}.$$

Claim. *For $\mathbf{x}$ orthogonal to $\mathbf{1}$, the denominator of $Q(\mathbf{x})$ equals $n\|\mathbf{x}\|^2$.*

> *Proof of the claim.* The denominator of $Q(\mathbf{x})$ is the sum of $(x_i - x_j)^2$ over all edges of the complete graph on $\{1, 2, \ldots, n\}$, whose Laplace matrix is $nI_n - J_n$. Then the identity (26) for the Laplace matrix then tells us that $\sum_{1 \le i < j \le n} (x_i - x_j)^2 = \mathbf{x}^T (nI_n - J_n)\mathbf{x} = n\|\mathbf{x}\|^2$ since $J_n\mathbf{x} = \mathbf{0}$ by the assumption. $\square$

Thus, we can further rewrite (27) to

$$\mu_2 = \min\left\{\mathbf{x}^T L\mathbf{x} : \|\mathbf{x}\| = 1, \langle \mathbf{1}, \mathbf{x}\rangle = 0\right\}. \tag{28}$$

But this is a (special case of a) standard result in linear algebra, the *variational characterization of eigenvalues* (or the *Courant–Fisher theorem*). It is also easy to check: We write $\mathbf{x}$ in an orthonormal basis of eigenvectors of L and expand $\mathbf{x}^T L\mathbf{x}$; we leave this to the reader. We just remark that the proof also shows that the minimum in (28) is attained by an eigenvector of L belonging to μ_2, which will be useful in the sequel. This concludes the proof of part (i) of the theorem. □

One of the main steps in the proof of part (ii) is the next lemma.

Lemma. *Let $A_k = \{1, 2, \ldots, k\}$, and let α be a real number such that each of the cuts $(A_k, V \setminus A_k)$, $k = 1, 2, \ldots, n$, has density at least α. Let $\mathbf{z} \in \mathbb{R}^n$ be any vector with $z_1 \geq z_2 \geq \cdots \geq z_n$. Then*

$$\sum_{\{i,j\}\in E, i<j} (z_i - z_j) \geq \frac{\alpha}{n} \sum_{1\leq i<j\leq n} (z_i - z_j). \tag{29}$$

Proof. In the left-hand side of (29) we rewrite each $z_i - z_j$ as $(z_i - z_{i+1}) + (z_{i+1} - z_{i+2}) + \cdots + (z_{j-1} - z_j)$. How many times does the term $z_{k+1} - z_k$ occur in the resulting sum? The answer is the number of edges $\{i, j\} \in E$ such that $i \leq k < j$, i.e. $|E(A_k, V \setminus A_k)|$. Thus

$$\sum_{\{i,j\}\in E, i<j} (z_i - z_j) = \sum_{k=1}^{n-1} (z_k - z_{k+1}) \cdot |E(A_k, V \setminus A_k)|.$$

Exactly the same kind of calculation shows that $\sum_{1\leq i<j\leq n} (z_i - z_j) = \sum_{k=1}^{n-1} (z_k - z_{k+1})|A_k| \cdot |V \setminus A_k|$. The lemma follows by using the density assumption $|E(A_k, V \setminus A_k)| \geq \frac{\alpha}{n}|A_k| \cdot |V \setminus A_k|$ for all k. □

Proof of part (ii) of the theorem. To simplify notation, we assume from now on that the vertices of G have been renumbered so that $u_1 \geq u_2 \geq \cdots \geq u_n$, where $\mathbf{u}$ is the eigenvector in the algorithm (then $\pi(i) = i$ for all i).

Let α be the density of the cut returned by the algorithm; we want to prove $\alpha \leq 4\sqrt{d_{\max}\mu_2}$. In the proof of part (i) we showed

$\mu_2 = Q(\mathbf{u}) = \big(\sum_{\{i,j\}\in E}(u_i - u_j)^2\big)/\|\mathbf{u}\|^2$, and so it suffices to prove

$$\alpha\|\mathbf{u}\| \le 4\bigg(d_{\max}\sum_{\{i,j\}\in E}(u_i - u_j)^2\bigg)^{1/2}. \tag{30}$$

We will obtain this inequality from the lemma above with a suitable $\mathbf{z}$, $z_1 \ge z_2 \ge \cdots \ge z_n$. Choosing the right $\mathbf{z}$ is perhaps the trickiest part of the proof; it may look like a magic but the calculations below will show why it makes sense.

First we set $\mathbf{v} := \mathbf{u} - u_{\lceil n/2\rceil}\mathbf{1}$. That is, we shift all coordinates so that $v_i \ge 0$ for $i \le n/2$ and $v_i \le 0$ for $i > n/2$. For later use, we record that $\|\mathbf{v}\| \ge \|\mathbf{u}\|$ (because $\mathbf{u}$ is orthogonal to $\mathbf{1}$).

Let us now assume that $\sum_{i:1\le i\le n/2} v_i^2 \ge \sum_{i:n/2<i\le n} v_i^2$; if it is not the case, we start the whole proof with $-\mathbf{u}$ instead of $\mathbf{u}$ (which obviously doesn't influence the outcome of the algorithm).

Next, we define $\mathbf{w}$ by $w_i := \max(v_i, 0)$; thus, $\mathbf{w}$ consists of the first half of $\mathbf{v}$ and then 0's. By the assumption made in the preceding paragraph we have $\|\mathbf{w}\|^2 \ge \frac{1}{2}\|\mathbf{v}\|^2 \ge \frac{1}{2}\|\mathbf{u}\|^2$.

Now, finally, we define $\mathbf{z}$ by $z_i := w_i^2$, and we substitute it into the inequality of the lemma (and swap the sides for convenience):

$$\frac{\alpha}{n}\sum_{1\le i<j\le n}(w_i^2 - w_j^2) \le \sum_{\{i,j\}\in E}(w_i^2 - w_j^2) \tag{31}$$

We will estimate both sides and finally arrive at (30).

First we deal with the right-hand side of (31). Factoring $w_i^2 - w_j^2 = (w_i - w_j)(w_i + w_j)$ and using the Cauchy–Schwarz inequality $\sum_{i=1}^n a_ib_i \le \big(\sum_{i=1}^n a_i^2\big)^{1/2}\big(\sum_{i=1}^n b_i^2\big)^{1/2}$ with $a_i = w_i - w_j$, $b_i = w_i + w_j$ yields

$$\begin{aligned}
\sum_{\{i,j\}\in E}(w_i^2 - w_j^2) &\le \bigg(\sum_{\{i,j\}\in E}(w_i - w_j)^2\bigg)^{\frac12}\bigg(\sum_{\{i,j\}\in E}(w_i + w_j)^2\bigg)^{\frac12}\\
&\le \bigg(\sum_{\{i,j\}\in E}(v_i - v_j)^2\bigg)^{\frac12}\bigg(\sum_{\{i,j\}\in E}2(w_i^2 + w_j^2)\bigg)^{\frac12}\\
&\le \bigg(\sum_{\{i,j\}\in E}(u_i - u_j)^2\bigg)^{\frac12}\sqrt{2d_{\max}}\,\|\mathbf{w}\|.
\end{aligned}$$

It remains to deal with the left-hand side of (31), which is quite simple:

$$\begin{aligned}\sum_{1\le i<j\le n}(w_i^2-w_j^2) &\ge \sum_{1\le i\le n/2}\ \sum_{n/2<j\le n}(w_i^2-w_j^2)\\ &= \sum_{1\le i\le n/2}\ \sum_{n/2<j\le n} w_i^2\\ &\ge \frac{n}{2}\|\mathbf{w}\|^2 \ge \frac{n}{4}\|\mathbf{u}\|^2.\end{aligned}$$

Putting this together with (31) and the previous estimate for its right-hand side, we arrive at (30) and finish the proof of part (ii) of the theorem. □

Sources. The continuous analog of the theorem is due to

> J. Cheeger, *A lower bound for the smallest eigenvalue of the Laplacian*, in *Problems in analysis* (Papers dedicated to Salomon Bochner, 1969), Princeton Univ. Press, Princeton, NJ, 1970, 195–199.

The discrete version was proved in

> N. Alon and V. D. Milman, λ_1, *isoperimetric inequalities for graphs, and superconcentrators*, J. Combin. Theory Ser. B **38,1** (1985), 73–88.

and

> N. Alon, *Eigenvalues and expanders*, Combinatorica **6,2** (1986), 83–96

and independently in

> J. Dodziuk, *Difference equations, isoperimetric inequality and transience of certain random walks*, Trans. Amer. Math. Soc. **284,2** (1984), 787–794.

A somewhat different version of the proof of part (ii) of the theorem can be found, e.g., in the wonderful survey

> S. Hoory, N. Linial, and A. Wigderson, *Expander graphs and their applications*, Bull. Amer. Math. Soc. (N.S.) **43,4** (2006), 439–561.

It is shorter, but to me it looks even slightly more “magical” than the proof above. A still different and interesting approach, regarding the proof as an analysis of a certain randomized algorithm, was provided in

> L. Trevisan, *Max cut and the smallest eigenvalue*, preprint, `http://arxiv.org/abs/0806.1978`, 2008.

The result concerning the second eigenvalue of planar graphs is from

> D. A. Spielman and S.-H. Teng, *Spectral partitioning works: planar graphs and finite element meshes*, Linear Algebra Appl. **421,2–3** (2007), 284–305.

A generalization and a new proof was given in

> P. Biswal, J. R. Lee, and S. Rao, *Eigenvalue bounds, spectral partitioning, and metrical deformations via flows*, in Proc. 49th Annual IEEE Symposium on Foundations of Computer Science, 2008, 751–760.

Approximation algorithms for the sparsest cut form an active research area.

Miniature 32

Rotating the Cube

First we state two beautiful geometric theorems. Since we need them only for motivation, we will not discuss the proofs, which involve methods of algebraic topology. Let $S^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$ stand for the unit sphere in $\mathbb{R}^n$, where $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ denotes the Euclidean norm. Thus, for example, S^2 is the usual 2-dimensional unit sphere in the 3-dimensional space.

(T1) For every continuous function $f\colon S^2 \to \mathbb{R}$ there exist three mutually orthogonal unit vectors $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ with $f(\mathbf{p}_1) = f(\mathbf{p}_2) = f(\mathbf{p}_3)$.

(T2) Let $\alpha \in (0, 2]$ and let $f\colon S^{n-1} \to \mathbb{R}^{n-1}$ be an arbitrary continuous mapping. Then there are two points $\mathbf{p}, \mathbf{q} \in S^n$ whose Euclidean distance is exactly α and such that $f(\mathbf{p}) = f(\mathbf{q})$. In popular terms, at any given moment there are two places on the Earth's surface that are exactly 1234 km apart and have the same temperature and the same barometric pressure.

Theorem (T2) probably motivated Bronisław Knaster to pose the following question in 1947:

Knaster's question. *Is it true that for every continuous mapping* $f\colon S^{n-1} \to \mathbb{R}^m$*, where* $n - 1 \geq m \geq 1$*, and every set* K *of* $n - m + 1$

points on S^{n-1} there exists a rotation ρ of $\mathbb{R}^n$ around the origin such that all points of the rotated set ρK have the same value of f?

It is easily seen that a positive answer to Knaster's question for all m, n would contain both (T1) and (T2) as special cases. In particular, the second theorem deals exactly with the case $m = n-1$ of Knaster's question.

Somewhat disappointingly, though, the claim in Knaster's question does *not* hold for all n, m, as was discovered in the 1980s. Actually, it almost *never* holds: By now counterexamples are known for every n and m such that $n - 1 > m \geq 2$, and also for $m = 1$ and all n sufficiently large.[1]

Here we discuss a counterexample for the last of these cases, namely, $m = 1$ (with some suitable large n). It was found only in 2003, after almost all of the other cases had been settled.

Theorem. *There exist an integer n, a continuous function $f\colon S^{n-1} \to \mathbb{R}$, and an n-point set $K \subset S^{n-1}$ such that for every rotation ρ of $\mathbb{R}^n$ around $\mathbf{0}$, the function f attains at least two distinct values on ρK.*

The function f in the proof is very simple, namely, $f(\mathbf{x}) = \|\mathbf{x}\|_\infty := \max\{|x_1|, |x_2|, \ldots, |x_n|\}$. The sophistication is in constructing K and proving the required property.

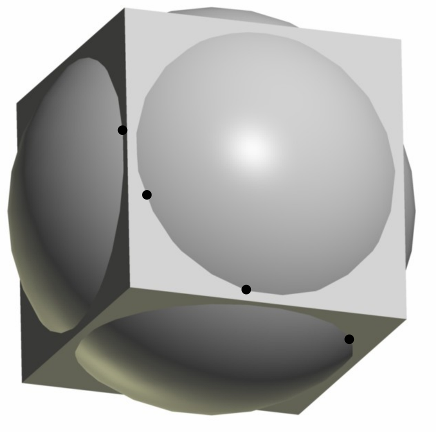
Some geometric intuition, not really necessary. The maximum value of f on S^{n-1} is obviously 1, attained at the points $\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n$. With a little more effort one finds that the minimum of f on S^{n-1} equals $n^{-1/2}$, attained at points of the form $(\pm n^{-1/2}, \pm n^{-1/2}, \ldots, \pm n^{-1/2})$.

Let us now consider the function $f(\mathbf{x}) = \|\mathbf{x}\|_\infty$ on all of $\mathbb{R}^n$. Then the set $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_\infty = 1\}$ is the surface of the unit cube $[-1, 1]^n$, and more generally, the level set $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_\infty = t\}$ is the surface of the scaled cube $[-t, t]^n$. Thus, if K is a point set on S^{n-1}, finding a rotation ρ such that f is constant on ρK can be reformulated as follows: Find a scaling factor t and a rotation of the scaled cube

---

[1] This doesn't kill the question, though: It remains to understand for which sets K the claim does hold, and this question is very interesting and very far from solved.

$[-t,t]^n$ such that all points of K lie on the surface of the rotated and scaled cube.



In the proof of the theorem, K is chosen as the disjoint union of two sets K_1 and K_2. These are constructed in such a way that if K_1 should lie on the surface of a rotated and scaled cube, then the scaling factor t has to be *large* (which means, geometrically, that the points of K_1 must be placed far from the corners of the cube), while for K_2 the scaling factor has to be *small* (the points of K_2 must be close to the corners). Hence it is impossible for both K_1 and K_2 to lie on the surface of the same scaled and rotated cube.

Preliminaries. In the theorem we deal with a point set K in the $(n-1)$-dimensional unit sphere and with rotated copies ρK. In the proof it will be more convenient to work with a set $\overline{K}$ living in the unit sphere S^{d-1} of a suitable lower dimension. Then, instead of rotations, we consider **isometries** $\varphi\colon \mathbb{R}^d \to \mathbb{R}^n$, that is, linear maps such that $\|\varphi(\mathbf{x})\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^d$. If φ_0 is one such isometry, then $K := \varphi_0(\overline{K})$ is a point set in S^{n-1}, and the sets $\varphi(\overline{K})$ for all other isometries $\varphi\colon \mathbb{R}^d \to \mathbb{R}^n$ are exactly all rotated copies of K (and their mirror reflections—but for the purposes of the proof we can ignore the mirror reflections).

We need one more definition. Let $X \subseteq \mathbb{R}^n$ be a set and let $\delta > 0$ be a real number. A set $N \subseteq X$ is called **$\boldsymbol{\delta}$-dense in $\boldsymbol{X}$** if for every $\mathbf{x} \in X$ there exists $\mathbf{y} \in N$ such that $\|\mathbf{x} - \mathbf{y}\| \le \delta$.

Lemma K1. (i) *Let $\varphi\colon \mathbb{R}^d \to \mathbb{R}^n$ be an isometry. Then there exists $\mathbf{x} \in S^{d-1}$ such that $\|\varphi(\mathbf{x})\|_\infty \ge \sqrt{d/n}$.*

(ii) *Let, moreover, $\overline{K}_1 \subset S^{d-1}$ be a $\frac{1}{2}$-dense set in S^{d-1}. Then there exists $\overline{\mathbf{p}} \in \overline{K}_1$ with $\|\varphi(\overline{\mathbf{p}})\|_\infty \ge \frac{1}{2}\sqrt{d/n}$.*

Proof. We begin with part (i). Let A be the matrix of the isometry φ with respect to the standard bases; i.e., the ith column of A is the vector $\varphi(\mathbf{e}_i) \in \mathbb{R}^n$, $i = 1, 2, \ldots, d$. Since φ preserves the Euclidean norm, the columns of A are unit vectors in $\mathbb{R}^n$, and thus

$$\sum_{i=1}^{n}\sum_{j=1}^{d} a_{ij}^2 = d. \tag{32}$$

Let $\mathbf{a}_i \in \mathbb{R}^d$ denote the ith *row* of A. For $\mathbf{x} \in \mathbb{R}^d$, the ith coordinate of $\varphi(\mathbf{x})$ is the scalar product $\langle \mathbf{a}_i, \mathbf{x}\rangle$, and thus $\|\varphi(\mathbf{x})\|_\infty = \max\{|\langle \mathbf{a}_i, \mathbf{x}\rangle| : i = 1, 2, \ldots, n\}$.

Now (32) tells us that $\sum_{i=1}^n \|\mathbf{a}_i\|^2 = d$, and thus there is an i_0 with $\|\mathbf{a}_{i_0}\| \ge \sqrt{d/n}$. Setting $\mathbf{x} := \mathbf{a}_{i_0}/\|\mathbf{a}_{i_0}\|$, we have $\|\varphi(\mathbf{x})\|_\infty \ge \langle \mathbf{a}_{i_0}, \mathbf{x}\rangle = \|\mathbf{a}_{i_0}\| \ge \sqrt{d/n}$, which finishes the proof of part (i).

We proceed with part (ii), which is the result that we will actually use later on. The proof is somewhat more clever than one might perhaps expect at first sight.

In the setting of (ii), we let $M := \sup\{\|\varphi(\mathbf{x})\|_\infty : \mathbf{x} \in S^{d-1}\}$, and let $\mathbf{x}_0 \in S^{d-1}$ be a point where M is attained.[2] By part (i) we have $M \ge \sqrt{d/n}$.

Since $\overline{K}_1$ is $\frac{1}{2}$-dense, we can choose a point $\overline{\mathbf{p}} \in \overline{K}_1$ with $\|\mathbf{x}_0 - \overline{\mathbf{p}}\| \le \frac{1}{2}$. If, by chance, $\overline{\mathbf{p}} = \mathbf{x}_0$, we are done, and so we may assume $\overline{\mathbf{p}} \ne \mathbf{x}_0$ and let $\mathbf{v} := (\mathbf{x}_0 - \overline{\mathbf{p}})/\|\mathbf{x}_0 - \overline{\mathbf{p}}\| \in S^{d-1}$ be the unit vector in direction $\mathbf{x}_0 - \overline{\mathbf{p}}$. Then $\|\varphi(\mathbf{v})\|_\infty \le M$ by the choice of M, and thus $\|\varphi(\mathbf{x}_0 - \overline{\mathbf{p}})\|_\infty \le \frac{1}{2}M$. Then, using the triangle inequality for the $\|.\|_\infty$ norm, we have

---

[2] The supremum is attained because S^{d-1} is compact. Readers not familiar enough with compactness may as well consider $\mathbf{x}_0$ such that $\|\varphi(\mathbf{x}_0)\|_\infty \ge 0.99M$, say, which clearly exists. Then the constants in the proof need a minor adjustment.

$$\|\varphi(\overline{\mathbf{p}})\|_\infty \geq \|\varphi(\mathbf{x}_0)\|_\infty - \|\varphi(\mathbf{x}_0 - \overline{\mathbf{p}})\|_\infty \geq M - \tfrac{1}{2}M = \tfrac{1}{2}M \geq \tfrac{1}{2}\sqrt{n/d}.$$

This proves part (ii). $\square$

Lemma K2. *Let $\overline{K}_2$ be a set of m distinct points of the unit circle $S^1 \subset \mathbb{R}^2$. If t is a number such that there exists an isometry $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^n$ with $\|\varphi(\overline{\mathbf{p}})\|_\infty = t$ for all $\overline{\mathbf{p}} \in \overline{K}_2$, then $t \leq \sqrt{8/m}$.*

Proof. We begin in the same way as in the proof of Lemma K1, this time setting $d = 2$: A is the matrix of φ and $\mathbf{a}_i \in \mathbb{R}^2$ is its ith row. By (32) we have $\sum_{i=1}^n \|\mathbf{a}_i\|^2 = 2$. We are going to bound the left-hand side from below in terms of m and t.

Since the ith coordinate of $\varphi(\overline{\mathbf{p}})$ equals $\langle \mathbf{a}_i, \overline{\mathbf{p}} \rangle$, the condition $\|\varphi(\overline{\mathbf{p}})\|_\infty = t$ for all $\overline{\mathbf{p}} \in \overline{K}_2$ can be reformulated as follows:

(C1) For every $\overline{\mathbf{p}} \in \overline{K}_2$ there exists an $i = i(\overline{\mathbf{p}})$ with $|\langle \mathbf{a}_i, \overline{\mathbf{p}} \rangle| = t$.

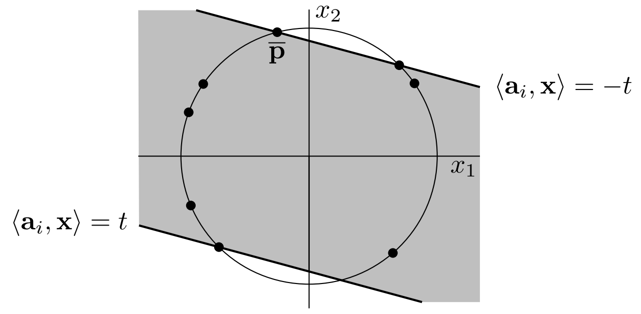
(C2) For all $\overline{\mathbf{p}} \in \overline{K}_2$ and all i we have $|\langle \mathbf{a}_i, \overline{\mathbf{p}} \rangle| \leq t$.

From (C1) we can infer that

$$\text{if } i = i(\overline{p}) \text{ for some } \overline{\mathbf{p}} \in \overline{K}_2, \text{ then } \|\mathbf{a}_i\| \geq t. \tag{33}$$

Indeed, $\overline{\mathbf{p}}$ is a unit vector, so $|\langle \mathbf{y}, \overline{\mathbf{p}} \rangle| \leq \|\mathbf{y}\|$ for all $\mathbf{y}$, and thus $|\langle \mathbf{a}_i, \overline{\mathbf{p}} \rangle| = t$ implies $\|\mathbf{a}_i\| \geq t$.

It remains to show that there are *many* distinct i with $i = i(\overline{\mathbf{p}})$ for some $\overline{\mathbf{p}} \in \overline{K}_2$. To this end, we observe that any given i can serve as $i(\overline{\mathbf{p}})$ for at most 4 distinct points $\overline{\mathbf{p}}$. This can be seen from the following geometric picture:


x_2
$\overline{\mathbf{p}}$
$\langle \mathbf{a}_i, \mathbf{x} \rangle = -t$
x_1
$\langle \mathbf{a}_i, \mathbf{x} \rangle = t$

The condition $i = i(\overline{\mathbf{p}})$ means that the point $\overline{\mathbf{p}}$ lies on one of the lines $\{\mathbf{x} \in \mathbb{R}^2 : \langle \mathbf{a}_i, \mathbf{x}\rangle = t\}$ and $\{\mathbf{x} \in \mathbb{R}^2 : \langle \mathbf{a}_i, \mathbf{x}\rangle = -t\}$, and (C2) implies that all points of $\overline{K}_2$ lie within the parallel strip between these two lines. In this situation, the boundary of such a parallel strip can contain at most 4 points of $\overline{K}_2$ (actually, at most 2 points provided that $\overline{K}_2$ is chosen in a suitably general position).

Consequently, there are at least $m/4$ distinct vectors of Euclidean norm at least t among the $\mathbf{a}_i$, and so $\sum_{i=1}^{n} \|\mathbf{a}_i\|^2 \geq t^2 m/4$. Since we already know that the left-hand side equals 2, we arrive at the claim of Lemma K2. □

Two ways of making $\boldsymbol{\delta}$-dense sets. The last missing ingredient for the proof of the theorem is a way of making a $\frac{1}{2}$-dense set $\overline{K}_1$ in S^{d-1}, as in Lemma K1(ii), that is not too large. More precisely, it will be enough to know that for every $d \geq 1$ such a $\overline{K}_1$ exists of size at most $g(d)$, for an arbitrary function g.

This is a well-known geometric result. One somewhat sloppy but quick way of proving it starts by observing that the integer grid $\mathbb{Z}^d$ is $\sqrt{d}$-dense in $\mathbb{R}^d$ (actually $\frac{1}{2}\sqrt{d}$-dense). If we re-scale it by $1/(4\sqrt{d})$ and intersect it with the cube $[-1,1]^d$, we have a $\frac{1}{4}$-dense set N_0 in that cube, of size at most $(8\sqrt{d}+1)^d$. Finally, for every point $\mathbf{x} \in N_0$ that has distance at most $\frac{1}{4}$ to S^{d-1}, we choose a point $\mathbf{y} \in S^{d-1}$ at most $\frac{1}{4}$ apart from $\mathbf{x}$, and we let $N \subset S^{d-1}$ consist of all these $\mathbf{y}$. It is easily checked that N is $\frac{1}{2}$-dense in S^{d-1}. This yields $g(d)$ of order $d^{O(d)}$.

Another proof, the standard "textbook" one, uses a greedy algorithm and a volume argument. We place the first point $\mathbf{p}_1$ to S^{d-1} arbitrarily, and having already chosen $\mathbf{p}_1, \ldots, \mathbf{p}_{i-1}$, we place $\mathbf{p}_i$ to S^{d-1} so that it has distance at least $\frac{1}{2}$ from $\mathbf{p}_1, \ldots, \mathbf{p}_{i-1}$. This process finishes as soon as we can no longer place the next point, i.e., the resulting set is $\frac{1}{2}$-dense. To estimate the number m of points produced in this way, we observe that the balls of radius $\frac{1}{4}$ around the $\mathbf{p}_i$ are all disjoint and contained in the ball of radius $\frac{5}{4}$ around $\mathbf{0}$. Thus, the total volume of the small balls is at most the volume of the large ball, and this gives $m \leq 5^d$, a better estimate than for the grid-based argument.

Proof of the theorem. We choose an even $n \geq 2g(100)$, we let $\overline{K}_1$ be a $\frac{1}{2}$-dense set in S^{99} of size at most $\frac{n}{2}$, and $\overline{K}_2$ is a set of $\frac{n}{2}$ points in S^1. We let $K := K_1 \cup K_2$, where $K_1, K_2 \subset S^{n-1}$ are isometric images of $\overline{K}_1$ and $\overline{K}_2$, respectively.

Lemma K1(ii) shows that for every rotation ρ there is a point $\mathbf{p} \in \rho K_1$ with $\|\mathbf{p}\|_\infty \geq \frac{1}{2}\sqrt{100/n} > 4n^{-1/2}$. On the other hand, if ρ is a rotation such that $\|\mathbf{p}\|_\infty$ equals the same number t for all $\mathbf{p} \in \rho K_2$, then $t \leq \sqrt{16/n} = 4n^{-1/2}$ by Lemma K2. This proves that $K = K_1 \cup K_2$ cannot be rotated so that all of its points have the same $\|.\|_\infty$ norm. □

Source. B. S. Kashin and S. J. Szarek, *The Knaster problem and the geometry of high-dimensional cubes*, C. R. Acad. Sci. Paris, Ser. I **336** (2003), 931–936.

Miniature 33

Set Pairs and Exterior Products

We prove yet another theorem about intersection properties of sets.

Theorem. *Let* $A_1, A_2, \ldots, A_n$ *be* k*-element sets, let* $B_1, B_2, \ldots, B_n$ *be* ℓ*-element sets, and let*

(i) $A_i \cap B_i = \emptyset$ *for all* $i = 1, 2, \ldots, n$*, while*

(ii) $A_i \cap B_j \neq \emptyset$ *for all* i, j *with* $1 \leq i < j \leq n$.

Then $n \leq \binom{k+\ell}{k}$.

It is easy to understand where $\binom{k+\ell}{k}$ comes from: Let $X := \{1, 2, \ldots, k+\ell\}$, let $A_1, A_2, \ldots, A_n$ be a list of all k-element subsets of X, and let us set $B_i := X \setminus A_i$ for every i. Then the A_i and B_i meet the conditions of the theorem and $n = \binom{k+\ell}{k}$.

The perhaps surprising thing is that we can't produce more sets satisfying (i) and (ii) even if we use a much larger ground set (note that the theorem doesn't put any restrictions on the number of elements in the union of the A_i and B_i; it only limits their size and intersection pattern).

The above theorem and similar ones have been used in the proofs of numerous interesting results in graph and hypergraph theory, combinatorial geometry, and theoretical computer science; one even speaks of the *set-pair method*. We won't discuss these applications here, though. The theorem is included mainly because of the proof method, where we briefly meet a remarkable mathematical object, the exterior algebra of a vector space.

The theorem is known in the literature as the **skew Bollobás theorem**. Bollobás originally proved a weaker (non-skew) version, where condition (ii) is strengthened to

(ii′) $A_i \cap B_j \neq \emptyset$ for all $i, j = 1, 2, \ldots, n$, $i \neq j$.

That version has a short probabilistic (or, if you prefer, double-counting) proof. However, for the skew version only linear-algebraic proofs are known. One of them uses the polynomial method (which we encountered in various forms in Miniatures 15, 16, 17), and another one, shown next, is a simple instance of a different and powerful method.

We begin with a simple claim asserting the existence of arbitrarily many vectors "in general position".

Claim. *For every $d \geq 1$ and every $m \geq 1$ there exist vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in \mathbb{R}^d$ such that every d or fewer among them are linearly independent.*

Proof. We fix m distinct and nonzero real numbers $t_1, t_2, \ldots, t_m$ arbitrarily and set $\mathbf{v}_i := (t_i, t_i^2, \ldots, t_i^d)$ (these are points on the so-called **moment curve** in $\mathbb{R}^d$).

Since this construction is symmetric, it suffices to check linear independence of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d$ (we assume $m \geq d$, for otherwise, the result is trivial). So let $\sum_{j=1}^{d} \alpha_j \mathbf{v}_j = \mathbf{0}$. This means $\sum_{j=1}^{d} \alpha_j t_i^j = 0$ for all i, i.e., $t_1, \ldots, t_d$ are roots of the polynomial $p(x) := \alpha_d x^d + \alpha_{d-1} x^{d-1} + \cdots + \alpha_1 x$. But 0 is another root, so we have $d+1$ distinct roots altogether, and since $p(x)$ has degree at most d, it cannot have $d+1$ distinct roots unless it is the zero polynomial. So $\alpha_1 = \alpha_2 = \cdots = \alpha_d = 0$.

Alternatively, one can prove the linear independence of the $\mathbf{v}_i$ using the Vandermonde determinant (usually computed in introductory courses of linear algebra).

Yet another proof follows easily by induction if one believes that $\mathbb{R}^d$ is not the union of finitely many $(d-1)$-dimensional linear subspaces. (But proving this rigorously is probably as complicated as the proof above.) □

On permutations and signs. We recall that the sign of a permutation $\pi\colon \{1,2,\ldots,d\} \to \{1,2,\ldots,d\}$ can be defined as $\operatorname{sgn}(\pi) = (-1)^{\operatorname{inv}(\pi)}$, where $\operatorname{inv}(\pi) = |\{(i,j) : 1 \le i < j \le d \text{ and } \pi(i) > \pi(j)\}|$ is the number of **inversions** of π.

Let d be a fixed integer and let $\mathbf{s} = (s_1, s_2, \ldots, s_k)$ be a sequence of integers from $\{1,2,\ldots,d\}$. We analogously define the sign of $\mathbf{s}$ as

$$\operatorname{sgn}(\mathbf{s}) := \begin{cases} (-1)^{\operatorname{inv}(\mathbf{s})} & \text{if all terms in } \mathbf{s} \text{ are distinct,} \\ 0 & \text{otherwise,} \end{cases}$$

where $\operatorname{inv}(\mathbf{s}) = |\{(i,j) : 1 \le i < j \le k \text{ and } s_i > s_j\}|$.

If we regard a permutation π as the sequence $(\pi(1), \ldots, \pi(d))$, then both definitions of the sign agree, of course.

The exterior algebra of a finite-dimensional vector space. In 1844 Hermann Grassmann, a high-school teacher in Stettin (a city in Prussia at that time, then in Germany, and nowadays in Poland spelled Szczecin), published a book proposing a new algebraic foundation for geometry. He developed foundations of linear algebra more or less as we know it today, and went on to introduce "exterior product" of vectors, providing a unified and coordinate-free treatment of lengths, areas, and volumes. His revolutionary mathematical discoveries were not appreciated during his lifetime (he became famous as a linguist), but later on, they were completed and partially re-developed by others. They belong among the fundamental concepts of modern mathematics, with many applications e.g. in differential geometry, algebraic geometry, and physics.

Here we will build the **exterior algebra** (also called the **Grassmann algebra**) of a finite-dimensional space in a minimalistic way (which is not the most conceptual one), checking only the properties we need for the proof of the above theorem.

Proposition. *Let V be a d-dimensional vector space.*[1] *Then there is a countable sequence $W_0, W_1, W_2, \ldots$ of vector spaces (among with only $W_0, \ldots, W_d$ really matter) and a binary operation $\wedge$ ("exterior product" or "wedge product") on $W_0 \cup W_1 \cup W_2 \cup \cdots$ with the following properties:*

(EA1) $\dim W_k = \binom{d}{k}$. *In particular, W_1 is isomorphic to V, while $W_k = \{\mathbf{0}\}$ for $k > d$.*

(EA2) *If $\mathbf{u} \in W_k$ and $\mathbf{v} \in W_\ell$, then $\mathbf{u} \wedge \mathbf{v} \in W_{k+\ell}$.*

(EA3) *The exterior product is* **associative**, *i.e.,* $(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{w} = \mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w})$.

(EA4) *The exterior product is* **bilinear**, *i.e.,* $(\alpha\mathbf{u} + \beta\mathbf{v}) \wedge \mathbf{w} = \alpha(\mathbf{u}\wedge\mathbf{w}) + \beta(\mathbf{v}\wedge\mathbf{w})$ *and* $\mathbf{u}\wedge(\alpha\mathbf{v}+\beta\mathbf{w}) = \alpha(\mathbf{u}\wedge\mathbf{v})+\beta(\mathbf{u}\wedge\mathbf{w})$.

(EA5) *The exterior product reflects linear dependence in the following way: For any $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d \in W_1$, we have $\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \cdots \wedge \mathbf{v}_d = \mathbf{0}$ if and only if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d$ are linearly dependent.*

Proof. Let $\mathcal{F}_k$ denote the set of all k-element subsets of $\{1, 2, \ldots, k\}$. For each $k = 0, 1, \ldots, d$ we fix some $\binom{d}{k}$-dimensional vector space W_k, and let us fix a basis $(\mathbf{b}_K : K \in \mathcal{F}_k)$ of W_k. Here $\mathbf{b}_K$ is just a *name* for a vector in the basis, which will be notationally more convenient than the usual indexing of a basis by integers $1, 2, \ldots$. We set, trivially, $W_{d+1} = W_{d+2} = \cdots = \{\mathbf{0}\}$.

We first define the exterior product on the basis vectors. Let $K, L \subseteq \{1, 2, \ldots, d\}$, where $s_1 < s_2 < \cdots < s_k$ are the elements of K in increasing order and $t_1 < \cdots < t_\ell$ are the elements of L in increasing order. Then we set

$$\mathbf{b}_K \wedge \mathbf{b}_L := \begin{cases} \operatorname{sgn}((s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_\ell))\mathbf{b}_{K \cup L} & \text{if } k + \ell \le d, \\ \mathbf{0} \in W_{k+\ell} & \text{if } k + \ell > d. \end{cases}$$

We note that, in particular, for $K \cap L \neq \emptyset$ we have $\mathbf{b}_K \wedge \mathbf{b}_L = \mathbf{0}$, since then the sequence $(s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_\ell)$ has a repeated term and thus its sign is 0. The signs are a bit tricky, but they are crucial for the good behavior of the exterior product with respect to linear independence, i.e., (EA5).

---

[1] Over any field, but we will use only the real case.

We extend $\wedge$ to all vectors bilinearly: If $\mathbf{u} \in W_k$ and $\mathbf{v} \in W_\ell$, we write them in the appropriate bases as $\mathbf{u} = \sum_{K \in \mathcal{F}_k} \alpha_K \mathbf{b}_K$, $\mathbf{v} = \sum_{L \in \mathcal{F}_\ell} \beta_L \mathbf{b}_L$, and we put

$$\mathbf{u} \wedge \mathbf{w} := \sum_{K \in \mathcal{F}_k, L \in \mathcal{F}_\ell} \alpha_K \beta_L (\mathbf{b}_K \wedge \mathbf{b}_L).$$

Now (EA1), (EA2), and (EA4) (bilinearity) are clear.

As for the associativity (EA3), it suffices to check it for basis vectors, i.e., to verify

$$(\mathbf{b}_K \wedge \mathbf{b}_L) \wedge \mathbf{b}_M = \mathbf{b}_K \wedge (\mathbf{b}_L \wedge \mathbf{b}_M) \tag{34}$$

for all K, L, M. The interesting case is when K, L, M are pairwise disjoint and $|K| + |L| + |M| \leq d$. Then, obviously, both sides of (34) are $\pm \mathbf{b}_{K \cup L \cup M}$, and it suffices to check that the signs match.

To this end, we let $s_1 < \cdots < s_k$ be the elements of K in increasing order, and similarly for $t_1 < \cdots < t_\ell$ and L and for $z_1 < \cdots < z_m$ and M. By counting the inversions of the appropriate sequences, we find that $(\mathbf{b}_K \wedge \mathbf{b}_L) \wedge \mathbf{b}_M = (-1)^N \mathbf{b}_{K \cup L \cup M}$, where $N = \mathrm{inv}((s_1, \ldots, s_k, t_1, \ldots, t_\ell)) + \mathrm{inv}((s_1, \ldots, s_k, z_1, \ldots, z_m)) + \mathrm{inv}((t_1, \ldots, t_\ell, z_1, \ldots, z_m))$, and the right-hand side of (34) comes out the same.

Next, if K, L, M are not pairwise disjoint or $k + \ell + m > d$, it is easily checked that both sides of (34) are $\mathbf{0} \in W_{k+\ell+m}$. Finally, having checked (34), it is routine to verify associativity in general—one just writes out the three participating vectors in the respective bases, expands both sides using bilinearity, and uses (34).

It remains to prove (EA5), which is the most interesting part where, finally, the choice of the sign turns from a hassle into a blessing.

Let $\mathbf{v}_1, \ldots, \mathbf{v}_d \in W_1$ be arbitrary, and let us write them in the basis $\mathbf{b}_{\{1\}}, \ldots, \mathbf{b}_{\{d\}}$ of W_1 as

$$\mathbf{v}_i = \sum_{j=1}^{d} a_{ij} \mathbf{b}_{\{j\}}.$$

Then, using bilinearity and associativity, we have

$$\mathbf{v}_1 \wedge \mathbf{v}_2 \wedge \cdots \wedge \mathbf{v}_d = \sum_{j_1, j_2, \ldots, j_d = 1}^{n} a_{1j_1} a_{2j_2} \cdots a_{dj_d} \mathbf{b}_{j_1} \wedge \mathbf{b}_{j_2} \wedge \cdots \wedge \mathbf{b}_{j_d}.$$

By the definition of the exterior product of basis vectors, all terms on the right-hand side where some two j_i coincide are $\mathbf{0}$. What remains is a sum over all d-tuples of distinct j_i's, in other words, over all permutations of $\{1, 2, \ldots, d\}$:

$$\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_d = \sum_{\pi} a_{1\pi(1)} a_{2\pi(2)} \cdots a_{d\pi(d)} \mathbf{b}_{\pi(1)} \wedge \mathbf{b}_{\pi(2)} \wedge \cdots \wedge \mathbf{b}_{\pi(d)}.$$

By considerations very similar to those in checking the associativity, we find that $\mathbf{b}_{\pi(1)} \wedge \mathbf{b}_{\pi(2)} \wedge \cdots \wedge \mathbf{b}_{\pi(d)} = \operatorname{sgn}(\pi)\mathbf{b}_{\{1,2,\ldots,d\}}$. Then the last sum transforms into $\det(A)\mathbf{b}_{\{1,2,\ldots,d\}}$, which is $\mathbf{0}$ exactly if the $\mathbf{v}_i$ are linearly dependent. The proposition is proved. □

With just a little more effort, (EA5) can be extended to any number of vectors; i.e., $\mathbf{v}_1, \ldots \mathbf{v}_n \in W_1$ are linearly dependent exactly if their exterior product is $\mathbf{0}$ (we won't need this but not mentioning it seems inappropriate).

Proof of the theorem. Let $d := k + \ell$ and let us consider the exterior algebra of $\mathbb{R}^d$ as in the proposition, with the vector spaces $W_0, W_1, \ldots$ and the operation $\wedge$. Let us assume, without loss of generality, that $A_1 \cup \cdots \cup A_n \cup B_1 \cup \cdots \cup B_n = \{1, 2, \ldots, m\}$ for some integer m, and let us fix m vectors $\mathbf{v}_1, \ldots \mathbf{v}_m \in W_1 \cong \mathbb{R}^d$ in general position according to the claim above (every d or fewer linearly independent). Note that m may be considerably larger than d.

Let $A \subseteq \{1, 2, \ldots, m\}$ be an arbitrary subset, and let us write its elements in increasing order as $i_1 < i_2 < \cdots < i_r$, where $r = |A|$. Then we define

$$\mathbf{w}_A := \mathbf{v}_{i_1} \wedge \mathbf{v}_{i_2} \wedge \cdots \wedge \mathbf{v}_{i_r}.$$

Thus, $\mathbf{w}_A \in W_r$.

For $A, B \subseteq \{1, 2, \ldots, m\}$ with $|A| + |B| = d$, (EA3) and (EA5) yield

$$\mathbf{w}_A \wedge \mathbf{w}_B = \begin{cases} \pm\mathbf{w}_{A\cup B} \neq \mathbf{0} & \text{for } A \cap B = \emptyset, \\ \mathbf{0} & \text{for } A \cap B \neq \emptyset. \end{cases}$$

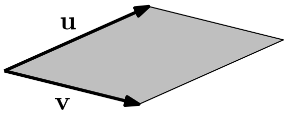
We claim that the n vectors $\mathbf{w}_{A_1}, \mathbf{w}_{A_2}, \ldots, \mathbf{w}_{A_n} \in W_k$ are linearly independent. This will prove the theorem, since $\dim(W_k) = \binom{d}{k} = \binom{k+\ell}{k}$.

So let $\sum_{i=1}^{n} \alpha_i \mathbf{w}_{A_i} = \mathbf{0}$. Assuming that, for some j, we already know that $\alpha_i = 0$ for all $i > j$ (for $j = n$ this is a void assumption), we show that $\alpha_j = 0$ as well. To this end, we consider the exterior product

$$\begin{aligned} \mathbf{0} \wedge \mathbf{w}_{B_j} &= \mathbf{0} = \left(\sum_{i=1}^{n} \alpha_i \mathbf{w}_{A_i}\right) \wedge \mathbf{w}_{B_j} \\ &= \sum_{i=1}^{n} \alpha_i (\mathbf{w}_{A_i} \wedge \mathbf{w}_{B_j}) = \alpha_j (\mathbf{w}_{A_j} \wedge \mathbf{w}_{B_j}), \end{aligned}$$

since $\mathbf{w}_{A_i} \wedge \mathbf{w}_{B_j} = 0$ for $i < j$ (using $A_i \cap B_j \neq \emptyset$), $\alpha_i = 0$ for $i > j$ by the inductive assumption, and $\mathbf{w}_{A_i} \wedge \mathbf{w}_{B_i} \neq \mathbf{0}$ since $A_i \cap B_i = \emptyset$. Thus, $\alpha_j = 0$, and the theorem is proved. □

The geometry of the exterior product at a glance. Some low-dimensional instances of the exterior product correspond to familiar concepts. First let $d = 2$ and let us identify W_1 with $\mathbb{R}^d$ so that $(\mathbf{b}_{\{1\}}, \mathbf{b}_{\{2\}})$ corresponds to the standard orthonormal basis $(\mathbf{e}_1, \mathbf{e}_2)$. Then it can be shown that $\mathbf{u} \wedge \mathbf{v} = \pm a \cdot \mathbf{e}_1 \wedge \mathbf{e}_2$, where a is the area of the parallelogram spanned by $\mathbf{u}$ and $\mathbf{v}$.



In $\mathbb{R}^3$, again making a similar identification of W_1 with $\mathbb{R}^3$, it turns out that $\mathbf{u} \wedge \mathbf{v}$ is closely related to the *cross product* of $\mathbf{u}$ and $\mathbf{v}$ (often used in physics), and $\mathbf{u} \wedge \mathbf{v} \wedge \mathbf{w} = \pm a \cdot \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$, where a is the volume of the parallelepiped spanned by $\mathbf{u}, \mathbf{v}$, and $\mathbf{w}$. The latter, of course, is an instance of a general rule; in $\mathbb{R}^d$, the volume of the parallelepiped spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_d \in \mathbb{R}^d$ is $|\det(A)|$, where A is the matrix with the $\mathbf{v}_i$ as columns, and we've already verified that $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_d = \det(A) \cdot \mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_d$.

These are only the first indications that the exterior algebra has a very rich geometric meaning. Generally, one can think of $\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k \in W_k$ as representing, uniquely up to a scalar multiple, the k-dimensional subspace of $\mathbb{R}^d$ spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_k$. However, by far not all vectors in W_k correspond to k-dimensional subspaces in this way; W_k can be thought of as a "closure" that completes the set of all k-dimensional subspaces into a vector space.

Sources. Bollobás' theorem was proved in

B. Bollobás, *On generalized graphs*, Acta Math. Acad. Sci. Hung. **16** (1965), 447–452.

The first use of exterior algebra in combinatorics is due to Lovász:

L. Lovász, *Flats in matroids and geometric graphs*, in Combinatorial surveys (Proc. Sixth British Combinatorial Conf., Royal Holloway Coll., Egham, 1977), Academic Press, London, 1977, 45–86.

This paper contains a version of the Bollobás theorem for vector subspaces, and the proof implies the skew Bollobás theorem easily, but explicitly that theorem seems to appear first in

P. Frankl, *An extremal problem for two families of sets*, European J. Combin. **3,2** (1982), 125–127,

where it is proved via *symmetric* tensor products (while the exterior product can be interpreted as an *antisymmetric* tensor product). The method with exterior products was also discovered independently by Kalai and used with great success in the study of convex polytopes and geometrically defined simplicial complexes:

G. Kalai, *Intersection patterns of convex sets*, Israel J. Math. **48** (1984), 161–174.

Applications of the set-pair method are surveyed in two papers of Tuza, among which the second one

Zs. Tuza, *Applications of the set-pair method in extremal problems, II*, in *Combinatorics, Paul Erdős is eighty, Vol. 2*, J. Bolyai Math. Soc., Budapest, 1996, 459–490

has a somewhat wider scope.

Index

$\equiv$ (congruence), 18
$\|\cdot\|$ (Euclidean norm), 1
$\|\cdot\|_1$ (ℓ_1 norm), 148
$\|\cdot\|_\infty$ (ℓ_∞ norm), 162
$\langle\cdot,\cdot\rangle$ (standard scalar product), 1
A^T (transposed matrix), 1
$\mathbf{u}\wedge\mathbf{v}$ (exterior product), 172
$\overline{G}$ (graph complement), 139
$G\cdot H$ (strong product), 131
$\alpha(G)$ (independece number), 131
$\vartheta(G)$ (Lovász theta function), 134
$\Theta(G)$ (Shannon capacity), 131

adjacency matrix, 32, 42, 46
  bipartite, 84, 100
algebra
  exterior, 171
  Grassmann, 171
algorithm
  probabilistic, 35, 36, 108, 119, 124
  Strassen, 32
alphabet, 12
arctic circle, 92
associativity, 123

Bertrand's postulate, 107
binary operation, 123
Binet's formula, 6
bipartite adjacency matrix, 84, 100
bipartite graph, 84, 99, 105
bits, parity check, 14
Borsuk's conjecture, 60
Borsuk's question, 59

capacity, Shannon, 131, 137
Cauchy–Schwarz inequality, 146, 157
characteristic vector, 57, 60
checking matrix multiplication, 35
checking, probabilistic, 105, 124
Cheeger–Alon–Milman inequality, 154
Cholesky factorization, 21
chromatic number, 134
code, 12
  error-correcting, 11
  generalized Hamming, 15
  Hamming, 12
  linear, 14
color class, 23
complement (of a graph), 139
complete bipartite graph, 23
congruence, 17
conjecture
  Borsuk's, 60
  Kakeya's, 113
corrects t errors, 13
cosine theorem, 17, 21
covering, 53

  of edges of K_n, 41
cube, 53
curve, moment, 170
cut, 152
  sparsest, 152
cycle
  evenly placed, 87
  properly signed, 87

decoding, 13
degree, 76
  minimum, 43
δ-dense set, 164
density, 152
determinant, 18, 75, 83, 105, 174
  Vandermonde, 170
diagonalizable matrix, 21
diagram, Ferrers, 95
diameter, 59
diameter-reducing partition, 59
digraph, 77
  functional, 80
dimension, 140
  Hausdorff, 113
dimer model, 92
directed graph, 77
discrepancy theory, 65
disjoint union (of graphs), 138
distance
  Euclidean, 19
  Hamming, 13
  ℓ_1, 148
  minimum (of a code), 13
  odd, 17
  only two, 49
divide and conquer, 151

$E(G)$, 1
eigenvalue, 146
eigenvalue (of a graph), 41, 43, 47
eigenvector, 153
encoding, 13
equiangular lines, 27
equilateral set, 145
Erdős–Ko–Rado theorem, 55
error-correcting code, 11
Euclidean distance, 19
Euclidean norm, 1
Euler's formula, 90
evenly placed cycle, 87
exponent of matrix multiplication, 32
exterior algebra, 171
exterior product, 169, 172
extremal set theory, 169

$\mathbb{F}_q$, 1
factorization, Cholesky, 21
fast matrix multiplication, 32, 35, 108, 119
Ferrers diagram, 95
Fibonacci number, 3, 5
Fiedler value, 154
finite field, 1, 57, 114
Fisher inequality, generalized, 9
formula
  Binet's, 6
  Euler's, 90
Frankl–Wilson inequality, 58
function, Lovász theta, 134
functional digraph, 80
functional representation, 139

general position, 170
generalized Fisher inequality, 9
generalized Hamming code, 15
generalized polygon, 44
generating matrix (of a code), 14
girth, 43
Gottlieb's theorem, 101
Gram matrix, 21, 145
graph, 1
  bipartite, 84, 99, 105
  complete bipartite, 23
  directed, 77
  Hoffman–Singleton, 45
  honeycomb, 84
  Moore, 44
  Petersen, 41, 45
  Pfaffian, 87
  planar, 86
  square grid, 83
  2-connected, 86
graph isomorphism, 102
Grassmann algebra, 171
group

  - action, 98
  - symmetric, 84, 117
- groupoid, 123

- Hamming code, 12
- Hamming distance, 13
- Hausdorff dimension, 113
- Hoffman–Singleton graph, 45
- honeycomb graph, 84
- hyperplane, 53

- I_n, 24
- icosahedron, regular, 27
- independence number (of a graph), 131
- independent set, 131
- inequality
  - Cauchy–Schwarz, 146, 157
  - Cheeger–Alon–Milman, 154
  - Frankl–Wilson, 58
  - generalized Fisher, 9
  - triangle, 19
- integer partition, 95
- inversion, 171
- isometry, 163
- isomorphism, graph, 102

- J_n, 24

- K_n (complete graph), 24
- Kakeya needle problem, 111
- Kakeya set, 112
- Kakeya's conjecture, 113
- Kasteleyn signing, 86
- Knaster's question, 161

- ℓ_1 distance, 148
- Laplace matrix, 75, 153
- lemma
  - rank, 146
  - Sperner, 55
  - Steinitz, 72
- linear code, 14
- Lovász theta function, 134
- Lovász umbrella, 132
- lozenge tiling, 84

- matching, 105
  - perfect, 83, 105
    - random, 91
- matrix
  - adjacency, 32, 42, 46
  - adjacency, bipartite, 84, 100
  - diagonalizable, 21
  - generating (of a code), 14
  - Gram, 21, 145
  - Laplace, 75, 153
  - multiplication
    - checking, 35
    - fast, 32, 35, 108, 119
  - orthogonal, 21
  - parity check, 15
  - positive semidefinite, 20, 154
- matrix-tree theorem, 75
- minimum degree, 43
- minimum distance (of a code), 13
- model, dimer, 92
- moment curve, 170
- Moore graphs, 44

- norm
  - Euclidean, 1
  - ℓ_1, 148
  - ℓ_∞, 162
- number
  - chromatic, 134
  - Fibonacci, 3, 5

- odd distances, 17
- Oddtown, 7
- operation, binary, 123
- orthogonal matrix, 21
- orthogonal representation, 131

- parity check bits, 14
- parity check matrix, 15
- partition
  - diameter-reducing, 59
  - integer, 95
- partitioning, spectral, 153
- PCP theorem, 37
- perfect matching, 83, 105
  - random, 91
- permanent, 85
- permutation, 117
- Petersen graph, 41, 45
- Pfaffian graph, 87
- planar graph, 86
- polygon, generalized, 44

polynomial, 50, 57, 106, 114, 127, 170
polynomials, vector space, 50, 54
positive definite matrix, 10
positive semidefinite matrix, 20, 154
postulate, Bertrand's, 107
probabilistic algorithm, 35, 36, 108, 119, 124
probabilistic checking, 37, 105, 124
problem, Kakeya needle, 111
product
  exterior, 169, 172
  standard scalar, 1
  strong, 131, 137
  tensor, 61, 134, 141
  wedge, 172
properly signed cycle, 87

question
  Borsuk's, 59
  Knaster's, 161

random perfect matching, 91
rank, 7, 10, 24, 97, 145
rank lemma, 146
recurrence, 4
representation
  functional, 139
  orthogonal, 131
rhombic tiling, 84

S^n, 161
S_n, 84, 117
scalar product, standard, 1
Schwartz–Zippel theorem, 107
  application, 115, 119, 127
semigroup, 123
set
  δ-dense, 164
  equilateral, 145
  independent, 131
  Kakeya, 112
set-pair method, 169, 176
Shannon capacity, 131, 137
sign (of a permutation), 76, 171
signing, Kasteleyn, 86
skew Bollobás theorem, 170
spanning tree, 75
sparsest cut, 152
spectral partitioning, 153
Sperner lemma, 55
square grid graph, 83
Steinitz lemma, 72
Strassen algorithm, 32
strong product, 131, 137
sunflower theorem, 55
symmetric group, 84, 117

tensor product, 61, 134, 141
theorem
  cosine, 17, 21
  Erdős–Ko–Rado, 55
  Gottlieb's, 101
  matrix-tree, 75
  PCP, 37
  Schwartz–Zippel, 107
    application, 115, 119, 127
  skew Bollobás, 170
  sunflower, 55
theta function, Lovász, 134
thinning, 112
tiling
  lozenge, 84
  of a board, 83
  of a rectangle, 39
  rhombic, 84
trace, 47, 146
tree, spanning, 75
triangle, 31
triangle inequality, 19
2-connected graph, 86

umbrella, Lovász, 132
unimodal, 97

$V(G)$, 1
value, Fiedler, 154
Vandermonde determinant, 170
vector, characteristic, 57, 60
vector space of polynomials, 50, 54

wall-equivalence, 98
wedge product, 172
word, 12