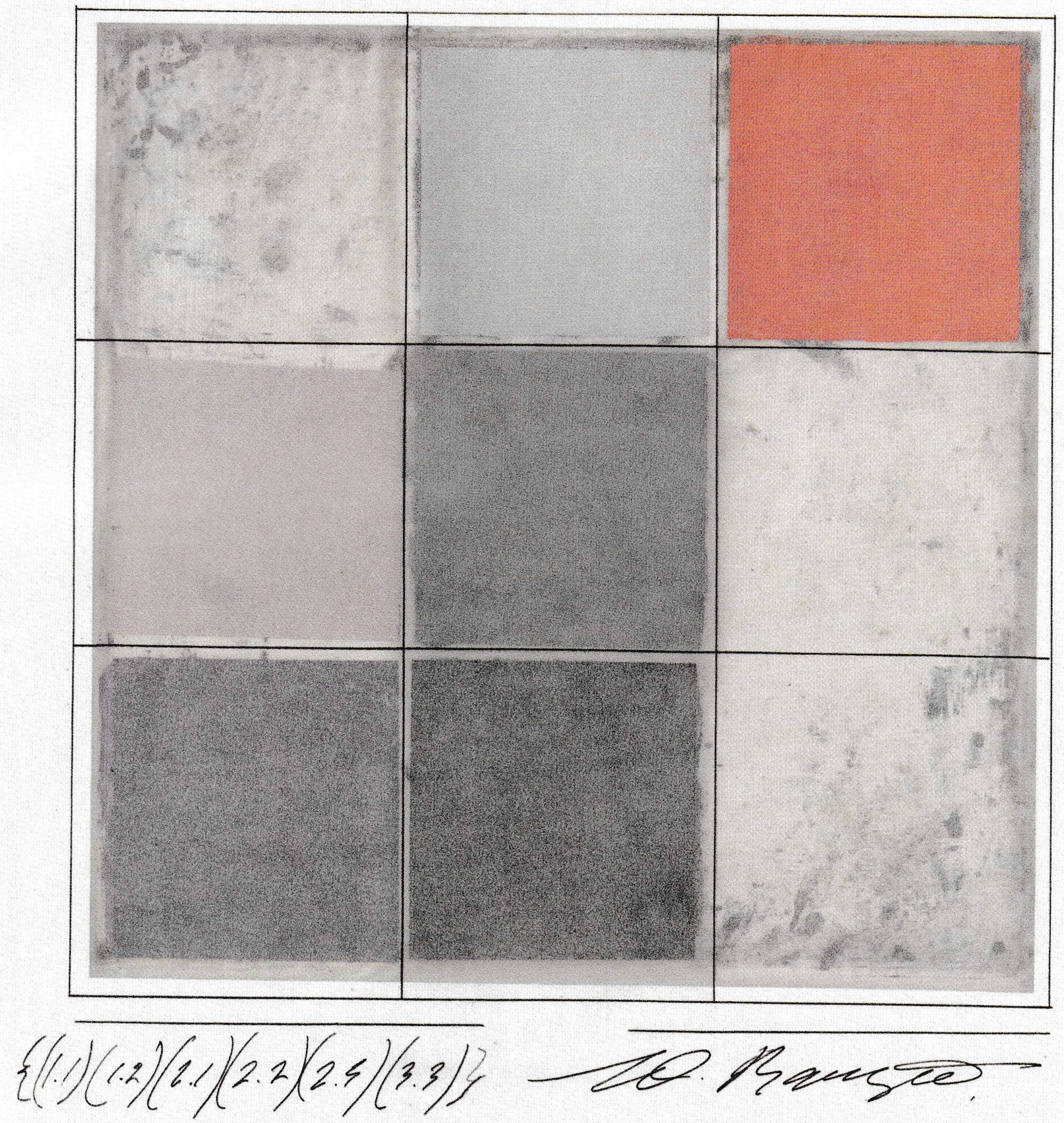
QUANTUM

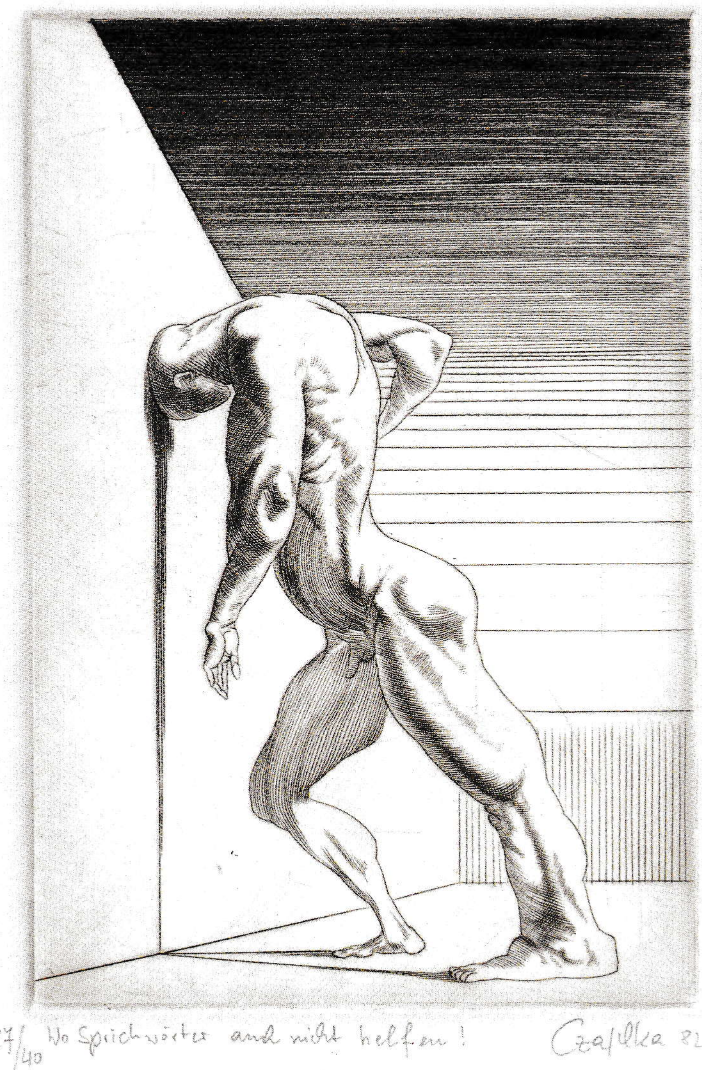
NOVEMBER/DECEMBER 1993

$5.00



SPRINGER INTERNATIONAL

GALLERY Q



From a private collection

"Wo Sprichwörter auch nicht helfen!" (1982) by Jurgen Czaschka

WE'VE ALL BEEN IN PLACES "WHERE CLEVER SAYINGS are no help at all" (*Sprichwort* means proverb or adage). When you're up against the wall—whether it's a tight deadline or a methodological impasse—a bystander's glib reference to "mind over matter" might just drive you up it.

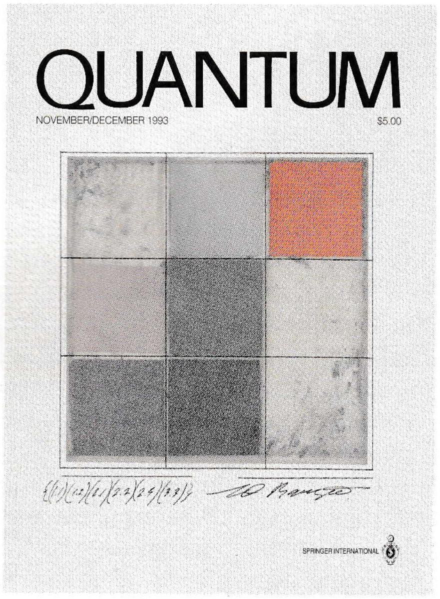
Austrian-born Jurgen Czaschka came to art by way of philosophy and literature. Before picking up his engraver's tools, he earned a doctorate from the University of Vienna and worked as a journalist and lecturer. So it's not surprising to find philosophical (one might say existential) overtones in this and many other of Czaschka's works. For instance, his engraving of Sisyphus resting—the eternal boulder bearing down on him from one direction, a resigned but resolute Sisyphus providing equal pressure from the other—owes its inspiration to the writings of Albert Camus.

A distinctive feature of Czaschka's graphic work is the asymmetry of the limbs on his human figures. This emotionally charged exaggeration of natural perspective seems to say: symmetry may be beautful, but it's not human (recall Blake's tiger with its "fearful symmetry").

Perhaps you also noticed something funny about the shadow. If the light beams striking the figure are parallel, we would expect the shadows of his legs to be parallel. There can't be two light sources—otherwise each leg would cast two shadows. Maybe the article "Late Light from Mercury" on page 40 will shed some you-know-what on the matter.

QUANTUM

NOVEMBER/DECEMBER 1993 VOLUME 4, NUMBER 2



Cover art by Yury Vashchenko

In his cover illustration Yury Vashchenko again uses the technique of covering the picture with tissue paper. (Recall the cover of the May/June issue and the portrait of Kepler therein.) This softens the colors and creates an illusion of fog, lending the image a dreaminess, or a sense of the proverbial "mists of time."

Here the tissue has the effect of taking the edge off a rather severe and abstract notion. The cover points to the lead article by Andrey N. Kolmogorov, another in the series of "primers" by the great Soviet mathematician. Not surprisingly, Vashchenko's graph turned out more colorful than its model, the purely functional graph on page 9. After all, color manipulation is yet another technique in the graphic artist's bag of tricks. But we can't help noticing that the title—"Bushels of Pairs"—might be the most colorful thing of all. (It was supplied by a *Quantum* editor—perhaps Kolmogorov would have found the notion of bushels of pairs preserving functions a bit . . . jarring.)

FEATURES

DEPARTMENTS

PUBLISHER'S PAGE

Savoring science

It's spicier than we're led to believe

AS A PERSON INTERESTED IN science, you have learned or are now learning science in a very compressed mode. Sadly, in the rush to convey as much subject matter as possible in the least amount of time, textbooks filter out the "extraneous," leaving you with only the essential content. What is more often lost is the essence of science.

Learning science would be far better if it were modeled on the way you would *write* a textbook, rather than on how you read one. To write such a book, you need to go to the library, search the literature for each topic related to your book, translate the research findings into simpler, explanatory language, and prepare the manuscript. Your book becomes expository, offering facts, terms, derivations and deductions, and detailed descriptions and explanations of models or theories. In effect, the finished book deprives the reader of the most important aspects of learning: a search of the literature, analysis, interpretation, and synthesis, along with coherent exposition. These benefits accrue only to the author, and the reader—usually a student—is thereby deprived of a real opportunity to learn. Not only that, the reader is forced to do something difficult and perhaps useful, but often boring and uninspiring.

One of the most interesting and satisfying things about doing research for a textbook is coming face to face with the history of science and mathematics. It's very revealing to see how hard it was for scientists and mathematicians of the past to carry out the original research for a particular discovery or the creation of a theory. For example, consider Isaac Newton and Robert Hooke, scientists who were contemporaries and who did most of their work in the mid-1600s.

Sir Isaac Newton was clearly one of the world's greatest scientists and mathematicians. Newton invented the differential calculus so that he could develop appropriate equations involving the rates of change of distance and velocity with time, leading to what are now called Newton's Laws of Motion. He invented the integral calculus in order to establish the points from which distances were to be measured for use in his Law of Universal Gravitation. But it is not to these monumental works that I shall now refer. Instead, let me draw your attention to Newton's "New Theory of Light and Colors."

Newton's theory of light and colors appeared in the February 19, 1671/72, issue of *Philosophical Transactions*, a publication of the Royal Society. The ambiguity in the year is due to the fact that England still used the Julian calendar, while the rest of Europe was on the Gregorian calendar (England did not change to the Gregorian calendar until September 1752).

The series of experiments that led Newton to his theory of colors is most interesting. His detailed lab work, the reasoning associated with it, and the way he used observation and measurement to lead to hypothesis and new experiment teach us more about science than any text that might refer to his work. Two aspects of this paper stand out.

Science textbooks often state that Newton used one prism to split light into its component colors and another prism to recombine the light to form "white" light. Although Newton did at one point use two prisms in a way similar to this, this was not what he deduced from the research described in the journal article. That experiment simply convinced him that the effect was not caused by irregularities in the prisms themselves. Newton actually used a prism and a lens, and it was in this way that he found conclusive evidence that different colors are refracted differently—that, for instance, blue light has a higher refractive index than red light.

Perhaps as important as Newton's theory of colors was his observation that the diameter of a circle of color at the focal point of a lens necessarily has to be of the order of 1/50 the diameter of the lens. So it would make no difference how skillfully you grind and polish a lens: the color effect is a defect linked with the refractive index of the glass for that color. He then deduced that to make a telescope for which the objective would not have color defects, he would need something that focused light, but not a lens. He realized that a parabolic surface of revolution would produce such a focus, because all colors reflect equally. This led him, after a two-year delay due to the Plague, to invent the reflecting telescope. He also suggested—but to my knowledge no one has ever made—a reflecting microscope. So if someone out there wants to do something creative, make one!

Robert Hooke, in developing the law $F = -kx$, which later came to bear his name, actually came up with the idea in 1671, but he didn't have time to prepare a paper for publication. So he did something quite common for that period. He published a 14-letter Latin anagram—a jumbled series of letters in which his discovery was concealed. He indicated that he would subsequently publish the details. Two years later he did publish his famous paper, and he translated the anagram. Nowadays students often must learn this law in a 20-minute class or lab period, yet the original science took years, and the experimental work alone consumed several months.

This history shows much more than can be summarized in a few lines of a textbook. There are ancillary benefits that involve technology and culture. These off-the-beaten-path intellectual sojourns should be a central part of everyone's education.

This is the philosophy behind *Quantum*'s Anthology department. We believe it's valuable to read important scientific writings as they originally appeared. The Anthology installment in the next issue—an autobiographical sketch by the great mathematician and teacher Sofya Kovalevskaya—will be somewhat different from those that have appeared in the past, but it will give a picture of mathematics as lived in the flesh and blood by one of its leading practitioners. Her struggles as a woman in science resonate to the present day. We think her story in the January/February issue will enrich your sense of the scientific endeavor. In the meantime, on page 35 we offer a biographical sketch of the Swiss mathematician Jacob Steiner, as promised in the May/June issue.

—Bill G. Aldridge

QUANTUM

THE MAGAZINE OF MATH AND SCIENCE

A publication of the National Science Teachers Association (NSTA)
& Quantum Bureau of the Russian Academy of Sciences
in conjunction with
the American Association of Physics Teachers (AAPT)
& the National Council of Teachers of Mathematics (NCTM)

The National Science Teachers Association is an organization of science education professionals and has as its purpose the stimulation, improvement, and coordination of science teaching and learning.

Quantum (ISSN 1048-8820) is published bimonthly by the National Science Teachers Association in cooperation with Springer-Verlag New York Inc. Volume 4 (6 issues) will be published in 1993–1994. *Quantum* contains authorized English-language translations from *Kvant*, a physics and mathematics magazine published by Quantum Bureau of the Russian Academy of Sciences, as well as original material in English. **Editorial offices:** NSTA, 1840 Wilson Boulevard, Arlington, VA 22201, telephone (703) 243-7100. **Production offices:** Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010. Supervising Production Editor: Madeline R. Kraner.

Second class postage paid at New York, NY, and additional mailing ofices. **Postmaster:** send address changes to: *Quantum* Springer-Verlag New York, Inc., Journal Fulfillment Services Department, 44 Hartz Way, Secaucus, NJ 07096-2491.

Subscription Information: North America

Student rate: $15; Personal rate (nonstudent): $20; Institutional rate: $34; Single Issue Price: $5. Rates include postage and handling. (Canadian customers please add 7% GST to subscription price. Springer-Verlag GST registration number is 123394918.) Subscriptions begin with next published issue (backstarts may be requested). Bulk rates for students are available. Send orders to *Quantum*, Springer-Verlag New York, Inc., P.O. Box 2485, Secaucus, NJ 07096; or call 1-800-SPRINGER (in New York, call [201] 348-4033).

All Countries Outside North America

Subscription rates in U.S. currency as above (all rates calculated in DM at the exchange rate current at the time of purchase) plus postage and handling. SAL (Surface Airmail Listed) is mandatory for Japan, India, Australia, and New Zealand. Customers should ask for the appropriate price list. Air mail delivery to all other countries is available upon request. Orders may be placed through your bookseller or directly through Springer-Verlag, Postfach 31 13 40, D-10643 Berlin, FRG.

Bushels of pairs

And each pair preserves a function

by Andrey N. Kolmogorov

EDITOR'S NOTE: THIS ARticle continues the elementary presentation of our modern understanding of functions and graphs that was begun in the last issue of *Quantum*.

A brief review and a clarification

In "Home on the Range" you got acquainted with the general understanding of the term "function." A function is an arbitrary mapping of a certain set E onto another set M. The set E is called the *domain* and the set M the *range* of the function. To define a function with domain E, for each element x of E we must specify a distinct "object"[1]

$$y = f(x).$$

However we do it, we get a function with domain E. For instance, the set E may comprise the students of your class, and $y = f(x)$, for any student x, may be the second letter of the student's name (here we're assuming that none of the students in the class have a name consisting of a single letter—though I once knew a girl named Olga who was usually called simply "O").

When a function is given in this way, its range M is defined automatically: it's the set of all objects y for which there is at least one x from E such that $f(x) = y$. Therefore, in describing the meaning of the term "function," we need not explicitly describe the range. It will be correct, for instance, to say simply that "a function is a law that assigns to any element x of a set E a certain object $y = f(x)$." As you may recall from the previous article, however, we shouldn't consider any of these descriptions *definitions*. If we really wanted to define the concept of a function in terms of the concept of a "law," we'd be asked to give an accurate definition of what a "law" means, and so forth. So we'll think of the concept of function as one of the basic notions of mathematics, whose sense is only explained rather than defined formally.

In school, you usually deal only with *number* functions, whose domain consists of numbers and whose values are numbers—*real numbers*, by and large. You can graph real-valued functions of real arguments on the "number plane."

Some textbooks say that a number plane is a plane onto which coordinates have been introduced in some prescribed way. Taking this literally, we find that there are a lot of number planes. Every time your teacher draws coordinate axes on the blackboard, the surface of the blackboard becomes a number plane, and you create new number planes in your notebooks—sometimes several planes on a single page!

In the third section of this article you'll get to know the sort of number plane mathematicians actually use. But first I want to make one additional comment regarding "Home on the Range."

The functions in high school algebra are usually given "analytically" by means of formulas. The domain of such a function, unless otherwise stated, is taken to be the set of all those values of the argument for which the operations on numbers prescribed by the formula can be carried out. For instance, if the sign "$\sqrt{\ }$" is understood as the "arithmetic square root," as is customary in high school, then the formula

$$y = f(x) = \left(\sqrt{x}\right)^2 \qquad (1)$$

allows us to compute the value of y that corresponds to a given x only if x is nonnegative (otherwise, the root can't be extracted).

For any nonnegative x,

$$y = f(x) = x. \qquad (2)$$

This formula is simpler than formula (1), and we'd like to look upon it as defining our function. However,

[1]You know from the previous article that the values of a function can be not only numbers but days of the week, boys or girls—anything at all.

Art by Yury Vashchenko


Sasha
Kolya
Sasha
Kolya

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Petya	✓				✓		
Kolya		✓				✓	
Sasha			✓				✓
Volodya				✓			

Figure 1

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Kolya		✓				✓	
Petya	✓				✓		
Sasha			✓				✓
Volodya				✓			

Figure 2

the domain of the function given by formula (2) consists not only of nonnegative numbers x but of *all* numbers x. So if we want to give a new definition of exactly the function defined by formula (1), we have to write

$$y = f(x) = \begin{cases} x & \text{for } x \geq 0, \\ \text{undefined} & \text{for } x < 0. \end{cases} \quad (3)$$

Similarly, the function $g(x) = (x^3 - 1)/(x - 1)$ can be written as

$$g(x) = \begin{cases} x^2 + x + 1 & \text{for } x \neq 0, \\ \text{undefined} & \text{for } x = 0. \end{cases} \quad (4)$$

One has to be precise with such algebraic transformations (especially on examinations!).

The graph of a function

Figure 1 shows a "duty chart" similar to the one we discussed in the previous article. We already know that it's the graph of a function: the name of a boy can be considered a function of a day of the week. Since there are seven days in a week and four boys, we've drawn $7 \times 4 = 28$ squares, but check marks appear in only seven of them.

If the boys had decided to arrange their names in alphabetical order, they would get the table in figure 2. It looks different, but it depicts the same distribution of jobs—that is, the same function. In both tables, 28 squares correspond to 28 possible *pairs* (day of week, boy). Of these 28 pairs, *even* pairs are singled out:

(Sun, Sasha), (Mon, Petya),
(Tue, Kolya), (Wed, Sasha),
(Thu, Volodya), (Fri, Petya),
(Sat, Kolya)

—that is, all the pairs of the form

(day of week, boy on duty that day),

or, formally, the pairs $(x, f(x))$. Only the selection of these pairs is essential in defining the function.

After seeing this example, you probably won't be surprised by the following *definition*:

The graph of a function f is the set of all pairs[2] *(x, y), such that (1) the first element x of a pair belongs to the domain of the function and (2) the second element of the pair y = f(x).*

In our example the graph of function f is

$$\Gamma_f = \{(\text{Sun, Sasha}), (\text{Mon, Petya}), (\text{Tue, Kolya}), (\text{Wed, Sasha}), (\text{Thu, Volodya}), (\text{Fri, Petya}), (\text{Sat, Kolya})\}.$$[3]

For the functions f_1, f_2, f_3, f_4 given by the following table

x	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
A	A	B	A	B
B	A	B	B	A

this definition yields the graphs

$$\Gamma_1 = \{(A, A), (B, A)\},$$
$$\Gamma_2 = \{(A, B), (B, B)\},$$
$$\Gamma_3 = \{(A, A), (B, B)\},$$
$$\Gamma_4 = \{(A, B), (B, A)\}.$$

It's clear that for functions with a finite domain, the number of elements in the graph (that is, of the pairs constituting the graph) is equal to the number of elements in the domain. For functions with an infinite domain, it's impossible to write out all the pairs $(x, f(x))$. So we have to describe these pairs by means of their properties.

For instance, the graph of the function

$$y = f(x) = \sqrt{1 - x^2}$$

consists of all number pairs of the form $(x, \sqrt{1 - x^2})$ (fig. 3)—that is, of all the pairs (x, y) satisfying two conditions: $x^2 + y^2 = 1$ and $y \geq 0$. This definition for the graph of this function can be written as

$$\Gamma_f = \{(x, y) \mid x^2 + y^2 = 1, y \geq 0\}.$$

The most general definition for the graph of a function can be written in the following form:[4]

$$\Gamma_f = \{(x, y) \mid y = f(x)\}.$$

By defining the graph of a function

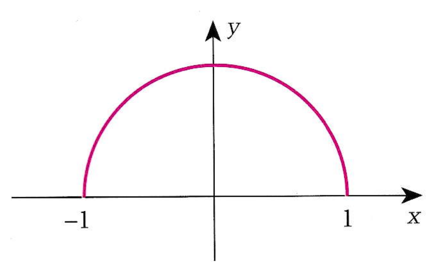


Figure 3

[2]All the "pairs" in this article are "ordered pairs": the pair (1, 2) differs from (2, 1). The first and second elements of a pair may coincide: (1, 1) and (2, 2) are pairs, too.

[3]Braces are commonly used to denote arbitrary sets.—*Ed.*

[4]We use the standard set-theory notation $\{x \mid A(x)\}$ for the set of all objects x satisfying the condition $A(x)$. For instance, $\{x \mid x^2 = 1\}$ is the set of all numbers x such that $x^2 = 1$—that is, the set of two numbers $\{1, -1\}$.

as the set of the pairs each of which consists of a value of the argument and the corresponding value of the function, we've cleared the notion of graph of all incidental details. In this abstract understanding, every function has a unique graph.

The number plane

Let's turn to real functions of the real variable, which you encounter most frequently in school. The graph of such a function is usually defined as the set of points $P(x, y)$ on the number plane with coordinates (x, y) satisfying the equation $y = f(x)$. This formulation and the general definition of a graph given in the previous section are similar but slightly different. There we talked about the set of *pairs* (x, y), while the usual "textbook" definition deals with the set of *points P* with coordinates x and y. But we can try to bring the two formulations to a full agreement, can't we?

It turns out to be quite easy. And it is this simple solution that gained ground throughout the modern scientific literature. *The number plane* is *defined* as *the set of all pairs of real numbers.* The number plane is denoted by $\mathbf{R}^2$. Symbolically, this definition is written as

$$\mathbf{R}^2 = \{(x, y) \mid x \in \mathbf{R}, y \in \mathbf{R}\}.$$

If you think it over for a moment, you'll see that with this definition of the number plane, the usual textbook definition of the graph of a real function of the real variable becomes a special case of the general definition given in the previous section.

Now the notation $P(x, y)$ for a point with coordinates x and y becomes redundant. *Points of the number plane* are thought of simply as *pairs of numbers* (x, y) by themselves. And we can simply say "point (0, 0)" (the origin), "points (1, 2), (–1, –2)," and so on.

It's worthwhile to note that the term "number line" must also take on a new meaning: the *number line* is simply the *set of real numbers* $\mathbf{R}$ itself. Then the *points of the number line* should be identified with the *real numbers*. Geometric language is often applied to numbers, though this is not always pointed out directly in high school textbooks—for instance, the set of numbers $[a, b] = \{x \mid a \le x \le b\}$ is called a "segment," "point" 2 is said to lie "on" the segment [1, 3], and so on.

Let's define a plane *geometric figure* as any set of points in the number plane. An example is the circle with center (0, 0) and radius 1 (fig. 4). This is the set of points—that is, pairs of numbers (x, y)—such that $x^2 + y^2 = 1$. Naturally, points and geometric figures in the number plane can be presented pictorially in a diagram. To this end, coordinate axes are chosen on a physical plane (like a sheet of paper or blackboard), and a point (x, y) of the number plane is represented by a "physical point" with coordinates x and y. Of course, this can be only an approximate representation. Graphs drawn on paper or on the blackboard are also only approximate images of "real" graphs of functions, which we now identify simply with subsets of the number plane, from our new point of view. It is these "real" graphs that are meant when we say that a function is fully determined by its graph.

Suppose a set of pairs $M = \{(x, y)\}$ is given. It may be, for instance, any "figure" in the number plane. What should be required additionally to ensure that this set is the graph of a certain function?

The answer is clear: the necessary and sufficient condition for a set M to be a graph is that it not contain two pairs (x, y_1) and (x, y_2) with a common first element x and different second elements y_1 and y_2. (Give a proof yourself.) The red curve in figure 5 is the graph of a function, while the black one is not.

A set of pairs (x, y) that doesn't contain two pairs of the form (x, y_1), (x, y_2), $y_1 \ne y_2$, may be called a *functional graph*. Notice that we've defined this concept without resort to the notion of "function." Isn't it possible then to take it as a starting point for a formal definition of the very notion of function, which we've been considering so far a fundamental one—that is, not subject to a formal definition? The answer to this question is not at all simple, so I don't want to go into it here.

Geometric transformations

To help you get used to the breadth of the general notion of "function," let's consider the simplest *geometric transformations*.

To turn a plane figure about a point O (fig. 6), we can place a sheet

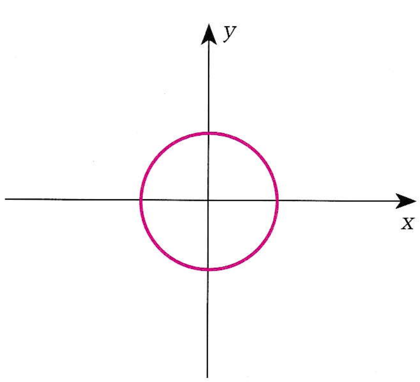


Figure 4

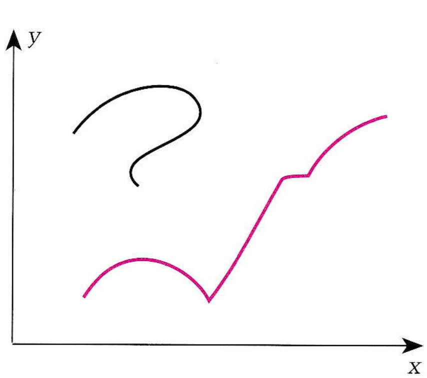


Figure 5

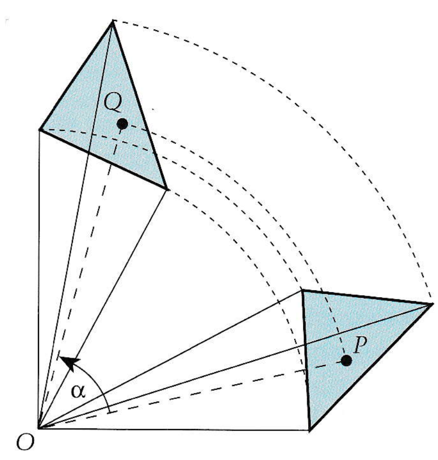


Figure 6

of tracing paper on the plane, trace the figure, pin the paper at point O, turn the paper over, copy the figure's copy from the tracing paper back onto the plane (using, say, carbon paper). Under this operation all points of our figure are rotated about point O in the same direction and through the same angle.

Let

$$Q = R_O^\alpha(P) \qquad (3)$$

be the position of point P after a counterclockwise rotation through angle α about O. If the point O and the angle α are fixed, formula (3) relates a uniquely defined point Q to each point P. According to our general definition, R_O^α is clearly a function. Its domain is the set of all points P of the plane.

The angle of rotation must be given with a sign. In figure 7, point Q_2 is obtained from point P by a rotation through 120°, while point Q_1 is the rotation of P through −120° (or 240°). If Q is obtained from P by a rotation through α degrees, then P can be obtained by rotating Q through $-\alpha$ degrees: that is, if $Q = R_O^\alpha(P)$, then $P = R_O^{-\alpha}(Q)$. Thus, a rotation R_O^α is always an *invertible* function.

It's more common to refer to rotations as *mappings* rather than functions. *The inverse mapping of the rotation R_O^α is the rotation $R_O^{-\alpha}$*. Symbolically, we write this as

$$R_O^{-\alpha}(R_O^\alpha(P)) = P,$$

$$(R_O^\alpha)^{-1} = R_O^{-\alpha}.$$

A rotation maps the set of points of the plane onto itself. If the plane is viewed as the set of all its points (which is the case in the modern presentation of geometry), we may say that *a rotation is an invertible mapping of the plane onto itself.*

Invertible mappings of the plane onto itself are called *geometric transformations of the plane.* Geometric transformations have appeared in our magazine in the past,

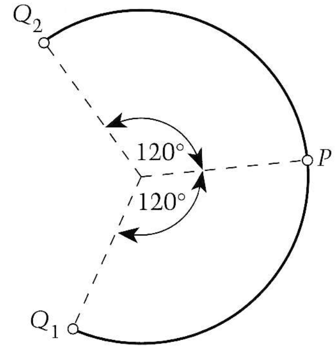


Figure 7

and will certainly appear repeatedly in the future.[5] For the time being, here's just one more example of a geometric transformation of the plane. *A mapping*

$$P \to Q = T(P)$$

of the plane onto itself is called a translation if all points P are displaced the same distance and in the same direction (fig. 8).

Vectors

Maybe you're tired of getting acquainted with new notions and unusual interpretations of notions you already know. But let's make one last effort. Let's try to understand what's meant by the *graph* of a translation $P \to Q = T(P)$. According to the general definition, it's the set of all pairs of points (A, B) such that

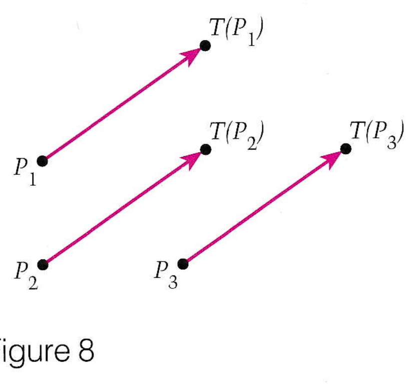


Figure 8

[5]Kolmogorov is referring to *Kvant*, not having lived long enough to participate in the creation of *Quantum*. But such articles have appeared in these pages as well—see, for instance, "[Getting to Know] Inversion" in the September/October 1992 issue.—*Ed.*

$B = T(A)$. Choose one such pair of points (A_0, B_0). How can the other pairs be characterized? For any of them the segment AB has the same length and the same direction as the segment A_0B_0 (fig. 9). The graph of the translation T is, by definition, the set of all these pairs (A, B).

It's conventional to assume that any pair (A, B) defines a "bound vector" $\overrightarrow{AB}$, and bound vectors $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ define the same "free vector" if the segments AB and $A'B'$ are equal in length and have the same direction.

More simply, a bound vector is just a pair of points (A, B) itself, and a "free vector" $\overrightarrow{AB}$ is the set of all bound vectors (A', B') equal to (A, B) in length and direction. But according to the general definition of a graph, this set is nothing but the graph of the translation T defined by the condition $T(A) = B$.

If $T(A_1) = B_1$, $T(A_2) = B_2$, $T(A_3) = B_3$, ..., we write

$$\overrightarrow{A_1B_1} = \overrightarrow{A_2B_2} = \overrightarrow{A_3B_3} = \ldots = \mathbf{a}$$

and

$$T_\mathbf{a} = T_{\overrightarrow{A_1B_1}} = T_{\overrightarrow{A_2B_2}} = T_{\overrightarrow{A_3B_3}} = \ldots .$$

The logic involved in creating general notions has led us to a somewhat unusual statement: *a free vector* **a** *is nothing other than the graph of the corresponding translation $T_\mathbf{a}$ by this vector.* You should ponder this statement well and make sure you understand that this conclusion is an inevitable consequence of our definitions of a graph, free vector (a set of bound vectors equal in length

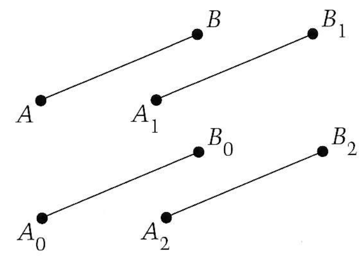


Figure 9

and direction to one another), and bound vector (a pair of points). I should point out, however, that these definitions of bound and free vectors are not universally accepted—they simply strike me as the most convenient.

Problems

Here are some problems to help you check your understanding of this article.

A brief review

1. State the domain of the following functions:

$$\text{(a)}\ f_1(x) = \frac{x}{x - |x|},$$

$$\text{(b)}\ f_2(x) = \frac{\sqrt{1-x}}{\sqrt{1+x^2}},$$

$$\text{(c)}\ f_3(x) = \frac{x^4 - 1}{x^2 - 1}.$$

2. What condition should be added to the formula $f(x) = x^2 + 1$ so as to obtain a definition of the function $f_3(x)$ from problem 1?

3. What condition should be added to the formula $f(x) = 1$ so as to get a definition of the function $f_4(x) = \left(\sqrt{x}\right)^2 + \left(\sqrt{1-x}\right)^2$?

NOTE. In problems 1 and 3 the sign $\sqrt{}$ denotes the "arithmetic square root"—that is, a nonnegative number.

The graph of a function

4. How many functions with the domain {1, 2, 3} whose graphs are subsets of the set {(1, 1), (1, 2), (2, 1), (2, 2), (2, 3), (3, 3)} are there? (See figure 10.) How many of these functions are invertible?

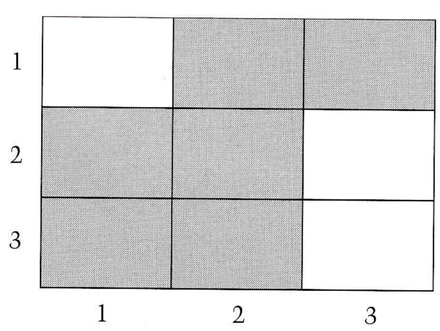


Figure 10

5. Show that the graph of the inverse function f^{-1} is defined by the formula $\Gamma_{f^{-1}} = \{(x, y) \mid (y, x) \in \Gamma_f\}$. (Naturally, we assume that the function f has an inverse.)

The number plane

6. Describe the graph of the Dirichlet function

$$D(x) = \begin{cases} 1, & \text{if } x \text{ is rational,} \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$

7. A number x in the interval [0, 1] is expanded into an infinite *ternary* (that is, to the base three) fraction $x = 0.x_1x_2x_3\ldots$ (x_i = 0, 1, or 2). The function $y = C(x)$ is defined as the number y in [0, 1] whose expansion into a *binary* fraction $y = 0.y_1y_2y_3\ldots$ is given by the condition

$$y_n = \begin{cases} 1, & \text{if } x_n \neq 0 \text{ and there are no ones among the digits } x_1, x_2, \ldots, x_{n-1}, \\ 0 & \text{in any other case.} \end{cases}$$

Try to draw the graph of this function. Show that it contains an infinite number of horizontal line segments. If you're familiar with the notion of a continuous function, try to prove that function $y = C(x)$ is continuous. (In this problem we do not avoid ternary fractions whose digits, starting from a certain place, are all twos, and binary fractions all of whose digits, starting from a certain place, are ones. For example, we assume that in the ternary notation 0.222... = 1, 0.1222... = 0.200..., and in binary notation 0.111... = 1, 0.010111... = 0.011000... .)[6]

Geometric transformations

8. Describe in geometric terms the transformations of the number plane given analytically by the following formulas: (a) $(x, y) \to (-y, x)$, (b) $(x, y) \to (x, -y)$, (c) $(x, y) \to (y, -x)$, (d) $(x, y) \to (-x, y)$, (e) $(x, y) \to (x + 1, y)$, (f) $(x, y) \to (x + a, y + a)$.

9. For rotations about a common center O, prove the formula

$$F(P) = R^{\alpha}_{O_1}\left[R^{-\alpha}_{O_2}(P)\right] \qquad (4)$$

10. Prove that for any two centers O_1 and O_2 the transformation

$$F(P) = R^{\alpha}_{O_1}\left[R^{-\alpha}_{O_2}(P)\right]$$

is a translation. By what distance and in what direction?

Vectors

11. Prove the formula

$$T_{\mathbf{a}}(T_{\mathbf{b}}(P)) = T_{\mathbf{a}+\mathbf{b}}(P). \qquad (5)$$

12. Prove that the transformation

$$F(P) = T_{\mathbf{a}}[R^{\alpha}_{O}(P)]$$

is a rotation through angle α. What point is its center?

NOTE. Formulas (4) and (5) can be written more concisely as $R^{\alpha}_{O}R^{\beta}_{O} = R^{\alpha+\beta}_{O}$ and $T_{\mathbf{a}}T_{\mathbf{b}} = T_{\mathbf{a}+\mathbf{b}}$. Taking a function of a function is in many respects similar to multiplication. But this is a special subject that can't be developed within the scope of this article. I'll use the above short notation for a function of a function (a composition of mappings) in problems 13 and 14.

13. Prove that $T_{\mathbf{a}}T_{\mathbf{b}} = T_{\mathbf{b}}T_{\mathbf{a}}$ for any two translations, and that $R^{\alpha}_{O}R^{\beta}_{O} = R^{\beta}_{O}R^{\alpha}_{O}$ for any two rotations about a common center. Give an example showing that, in general, $R^{\alpha}_{O_1}R^{\beta}_{O_2} \neq R^{\beta}_{O_2}R^{\alpha}_{O_1}$ if the centers of rotations O_1 and O_2 are different.

14. Give an exhaustive explanation of the case in which $R^{\alpha}_{O_1}R^{\beta}_{O_2} = R^{\beta}_{O_2}R^{\alpha}_{O_1}$. ■

[6]It might be well to remind you that the ternary fraction $0.x_1x_2x_3\ldots$ (x_n = 0, 1, 2), by definition, is equal to $x = x_1/3 + x_2/3^2 + x_3/3^3 + \ldots$, and the binary fraction $0.y_1y_2y_3\ldots$ (y_n = 0, 1) equals $y = y_1/2 + y_2/2^2 + y_3/2^3 + \ldots$. This problem is the most difficult one in the article, but also the most interesting. While solving it, you may discover, in addition to the aforementioned properties of $C(x)$, other peculiar features of this function. For instance, it is nondecreasing and maps the interval [0, 1] onto itself. (We offer a solution to this problem in the answer section.)—*Ed.*

The case of the mythical beast

Sherlock Holmes unravels a "diabolical" mystery

by Roman Vinokur

"IT HAPPENED IN THE fifth century A.D. on the rocky shores of a deep lake in what is now northern England. A wandering monk, Brother George, had come here to convert the local inhabitants to Christianity. The path Brother George had chosen was long and difficult, but he was a man of courage. To his surprise, the locals told him that a mighty god lived in the lake. They took the monk to the very edge of a high cliff that drops down to the still, cold water. The hills around the lake were thickly forested, so the smooth surface of the water seemed to be painted green.

"Suddenly a giant beast rose from the depths. Its head looked like that of an enormous seal, and it had a single white horn in the middle of its shiny black forehead. The locals looked frightened, but they told the monk that the water god ate only plants. The creature looked at them, its head almost reaching the top of the cliff. The locals dropped to their knees, bowed low, and asked the beast not to punish them. Brother George lifted his crucifix and ordered the creature in God's name to return to the nether world. But the creature paid no heed.

"Now, in those days, missionaries could handle the sword as well as the Word—it was a dangerous occupation. Grabbing a spear that one of the locals had dropped, Brother George threw it and hit the beast right in the eye. The creature howled, fell back, and disappeared into the waters forever . . .

"Since then, no one has seen it for certain, although a few people have said they've seen it from afar."

After a short pause the young marine engineer, John Turner by name, continued his story.

"I think the beast was a dinosaur who came from a distant sea to live in this lake, and that its relatives stayed behind. There probably weren't many of them, but enough to survive as a species . . ."

Sherlock Holmes appeared profoundly interested. "I wish you much luck in finding that creature in the lake," he said, leaning back in his chair, "or anywhere else, for that matter. I suppose you have spent a great deal of your time trying to discover it for science. I see some impediments to any such expectations, but it's not impossible . . . And what is your opinion, Dr. Watson?"

I said that such a discovery would be wonderful, and that there was a chance that some breeds of dinosaur could survive—for example, creatures like the crocodiles and lizards of our own time. Moreover, one of my patients, a well-known geographer, told me about mysterious giant beasts inhabiting the wild jungles of Cameroon. He didn't happen to see these creatures himself. Nevertheless, a few of the local hunters met up with them, and their reports of the *mokele-mbebve* (as they called the "dinosaurs") seemed perfectly credible.

Suddenly Holmes laughed in the hearty, noiseless manner that was peculiar to him and said, addressing our fair guest, "By the way, I believe your little black poodle would be pleased to help its master hunt for unknown animals, wouldn't it? Pray, do not be surprised, sir. That is my line of work—to know things."

Nonetheless our brave mariner was flabbergasted. "Sir," he said, "I've read about your talents, but please refresh my memory—where exactly did we meet before? My dog Judy is indeed a little black poodle."

"I saw bite marks on the heels of your shoes," replied Holmes, laughing softly. "That usually happens when one has a frisky puppy at home. Besides, I chanced to see through the window that you stopped in Baker Street to watch a black poodle that was out for a walk with its master. At that moment you looked like someone recalling something very familiar and pleasant. You piqued my curiosity, and so I took a look at your heels when you sat down next to me. That's all there is to it. So now you can see that I am not a magician. But pray, explain to us how we may be of service."

"I really want to find that beast in the lake," Turner said decisively. "I

Art by Sergey Ivanov

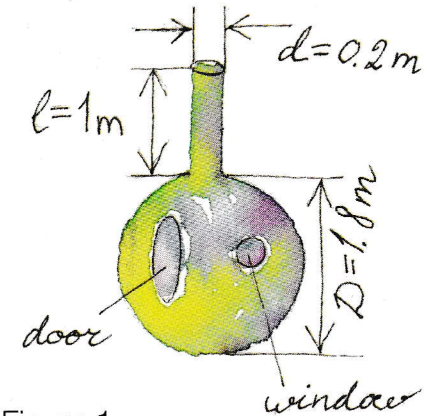


Figure 1

designed a vessel to carry me under the surface so I could explore the depths of the lake. Here's a sketch." He produced the drawing shown in figure 1. "It's like a big bottle: a hollow sphere with a cylindrical tube coming out of it. A hermetic door and a window are built into the side. The window is made of plate glass, and everything else is steel. I had the vessel built according to my instructions and brought to the shore of the lake. I planned to tow it to the middle of the lake and sink it with special anchors, which are not shown in the drawing.

"But here I came across a serious problem. I had hired some workers to do a few final tasks on site—installing ventilation hoses and telephone lines and sealing the opening at the top of the tube. These workers became ill, and they tried to convince me that there was something amiss with my underwater vessel—that the devil had taken up residence there to thwart my plans. They wouldn't listen to reason, so I proposed to spend one hour in the vessel to show them that nothing would happen.

"It was a clear, warm day. The sun shone encouragingly, a light breeze blew from the lake. But to my astonishment, Judy tried to keep me from getting into the vessel. She clamped her jaws on my pant cuffs and wouldn't let go. But somehow I got loose and entered the vessel, closing the door behind me. Almost immediately I felt as if my insides were shaking. Everything went black. An inexplicable fear welled up inside me. I couldn't take it any longer and jumped out of the vessel.

"Afterwards, I tried again—several times, in fact—with the same result. The workers believe that the devil is making fools of us to prevent our investigation of the lake. They think he lives in the depths and sometimes transforms himself into the legendary unicorn. Sir, I am relying on you and your deductive method to explain this to me—unless, of course, the workers are right and there are unclean forces at work here."

"Neither God nor the devil has ever made a personal appearance in my life," Holmes said thoughtfully. "I suspect that the marine beast has nothing to do with what you experienced either, and that the dreadful cause you seek lies in inanimate nature. I came across a similar case once, when I was investigating the death of a hunter who had taken shelter in a small cave to escape bad weather. There were no signs of struggle with man or beast—no wounds, nothing to indicate how he had died. To tell you the truth, I could not have solved the mystery without the help of my former teacher, a well-known professor of physics . . ."

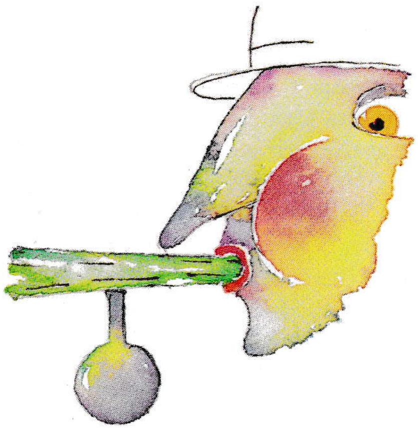


Figure 2

Holmes took a small toy out of his pocket. It was made of multicolored glass and looked very much like Turner's underwater vessel. He raised it to his lips and blew across the opening (fig. 2). After several tries he managed to produce a clear, sustained whistling sound. I caught a confused glance from our guest, and I felt a little awkward on my friend's account. I thought he was showing signs of entering his second childhood.

"This whistling toy," Holmes began to explain, "is what a physicist would call a Helmholtz resonator. It's named after the scientist who first used resonators to analyze sounds according to the frequency of their oscillations. The device consists of two basic elements: a thin tube open at both ends and a chamber that is much bigger in volume. A change in the air pressure at the outer end of the tube causes the air inside the resonator to move. Let's imagine that some air flows into the chamber. This causes an increase in the pressure in the chamber, which inhibits any further flow of air into the chamber. Similarly, a flow of air out of the chamber causes a decrease in the air pressure in the chamber, which inhibits any further outward flow of air. This means that the air inside the chamber can be viewed as a kind of spring, and the air in the tube plays the role of a mass attached to the spring. In this sense the Helmholtz resonator is like a simple oscillator (fig. 3).

"I should point out that this analogy is not always valid. It is valid only when the speed of the air particles is approximately the same along the entire length of the tube. This holds true at sufficiently low frequencies, when the length of the

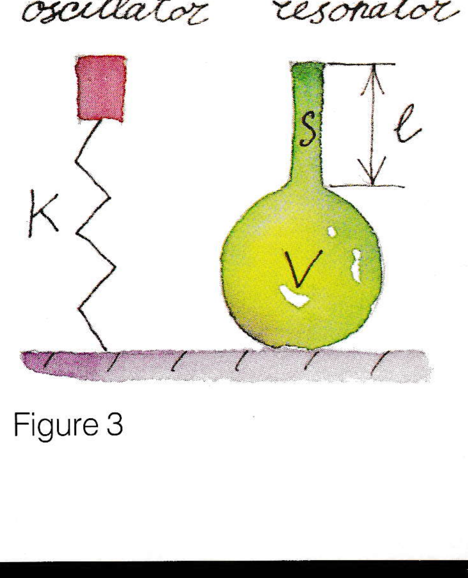


Figure 3

sound wave in the air is much larger than the dimensions of the resonator. Otherwise, several simple oscillators would be needed to model the vibrations of the air in the vessel."

"As I recall," the engineer offered, "the natural frequency of a simple oscillator is

$$f = \frac{\sqrt{K/M}}{2\pi}, \qquad (1)$$

where K is the spring constant and M is the mass. What would you say about the natural frequency of the Helmholtz resonator, Mr. Holmes?"

"If the length of the tube is much greater than its diameter, the frequency is equal to

$$f_r = \frac{v_s}{2\pi}\sqrt{\frac{S}{Vl}}, \qquad (2)$$

where $v_s \cong 340$ m/s is the speed of sound in air, S is the cross-sectional area of the tube, V is the volume of the chamber, and l is the length of the tube. Substituting the dimensions of the underwater vessel into equation (2) . . ." Holmes took a pencil and his notebook and looked at the engineer's sketch (fig. 1). "We find that

$$f_r \cong 5 \text{ Hz}.$$

Agreed?"

"I think I follow you," Turner said in an unsure voice. "The toy whistled when you blew across its neck. In my case, the wind did the blowing. But the wind blowing off the lake wasn't that strong . . ."

"You are correct," Holmes agreed, "at least as far as the wind's role is concerned. When air flows past an object, the wake behind it is not regular. There one can find so-called 'vortices' (fig. 4), which alternately leave the object in such a way that a definite period of time elapses between the formation of successive vortices.

"This 'vortex stream' was investigated theoretically by Theodore Karman not so long ago. Air pressure in a vortex is less than in the undisturbed atmosphere (this is why tornados act like huge vacuum cleaners, sweeping up everything they come across). Likewise, the air pressure in vortices is less than in the spaces between them. If the object is symmetrical, the total air pressure behind it changes harmonically. The frequency of the alternating pressure can be expressed by using only the dimensions of the actual physical parameters:

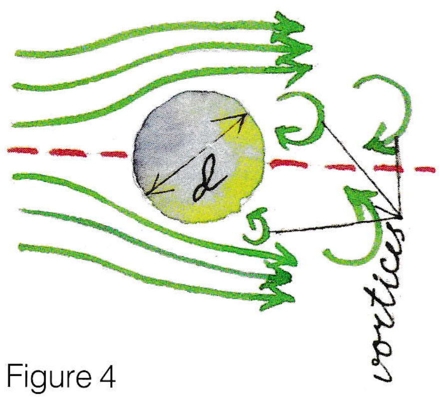


Figure 4

$$f_k = k\frac{v}{a}, \qquad (3)$$

where v is the speed of the air flow, a is the effective size of the object, and k is a dimensionless coefficient that depends on the object's shape and orientation. The value of k is usually found experimentally, but theoretical predictions can also be made. For instance, if the air flow is perpendicular to a long cylinder, the value of k is found to equal approximately 0.2, provided that $a = D$ (where D is the diameter of the cylinder). Considering that the wind speed at the time of your incidents at the lake shore was about 5 m/s (as you said, the wind was not blowing strongly), and substituting all values into equation (3) (I know that $D = 0.2$ m and $k = 0.2$), we get

$$f_k = 5 \text{ Hz}.$$

"I should point out that, in the case we are investigating, the role of the object can be played not only by the tube as a whole but also by its upper rim. Nevertheless, we can again use equation (3), supposing that $a = h$, where h is the wall thickness. All that remains is to measure the coefficient k. In the present case, however, I believe the role of the tube as a whole is more significant, since both frequencies are the same."

"So $f_r = f_k$," cried Turner, "which means there was acoustical resonance! And so the amplitude of the oscillations in air pressure in the underwater vessel could be quite large. The same phenomenon causes the whistle to produce a noise. But—"

"Of course you didn't hear anything," interrupted Holmes, finishing the engineer's thought. "Sound with a frequency of less than 16 Hz is inaudible. It's called infrasound, and its effect on human beings is not completely understood. We do know, however, that high-intensity ultrasound causes headache, fatigue, and anxiety. Moreover, powerful infrasound can cause more serious problems. Our internal organs (heart, liver, stomach, kidneys) are attached to the bones by elastic connective tissue, and at low frequencies they may be considered simple oscillators. The natural frequencies of most of them are below 12 Hz (which is in the infrasonic range). Thus, the organs may resonate.

"Of course, the amplitude of any resonance vibrations depends significantly on damping, which transforms mechanical energy into thermal energy. In the ideal case of zero damping, the resonance amplitude would increase to infinity. In real cases, however, this amplitude decreases as the damping increases. Also, the amplitude is proportional to the amplitude of the harmonic force causing the vibrations. Such a force produced by Karman vortices is approximately proportional to ρv^2, where ρ is the air density and v is the wind speed. Your troubles, sir, were relatively harmless because the wind happened to be weak. In the case I mentioned earlier—the dead man in the cave—the winds were of gale force . . ."

"But what do you make of Judy's behavior?" I asked my friend. "She tried to save her master. So, it seems, dogs can sense infrasound as well as smells that we cannot perceive?"

"Well, at least they do it better than people," Holmes replied, laughing. "By the way, the wavelength of infrasound with a frequency of 5 Hz equals

$$\lambda = \frac{v_s}{f} = \frac{340 \text{ m/s}}{5 \text{ Hz}} = 68 \text{ m}.$$

This value is much greater than the maximum size of the vessel (about 3 m), so our model is correct. However, I had one more reason to calculate the infrasound wavelength for this case. It's a well-known fact in acoustics that small sources of sonic waves—small, that is, in comparison with the wavelength of the sound being radiated—cannot be very powerful because of their limited ability to radiate the sound. This is why you felt the infrasound only when you were in your underwater vessel or very close by—most of the acoustical energy was kept inside.

"I trust, sir, after our conversation, you can explain to your workers that the devil was not to blame for the discomfort they experienced in your underwater vessel."

"Most certainly," the engineer replied gratefully. "I will do so first thing tomorrow."

. . . As it turned out, "tomorrow" was the day the First World War began. An isolated British destroyer was torpedoed by a German submarine and sank with all its crew aboard—except for officer John Turner. He found himself swimming in the endless sea. Time passed, but no one came to save him. He was tired and cold. He had all but lost hope when he felt something large swimming nearby. "Shark!" he thought with horror. He looked around . . . It wasn't a shark, but a giant beast with a white horn on its broad, seal-like forehead. For a while they floated side by side. Then the beast swam off at great speed and dove into the depths of the ocean. Half an hour later, Turner was being hoisted into a boat in the strong arms of an American sailor.

"Do you know, sir, that you are a lucky man?" The voice Turner heard was cheerful and husky. "Our captain hasn't slept for two days running. He thought he saw a huge, horned seal or something in his binoculars. He ordered us to change course to get a closer look, and we found you instead! If it hadn't been for that mirage . . . why, it saved your life, didn't it, sir?" ◘

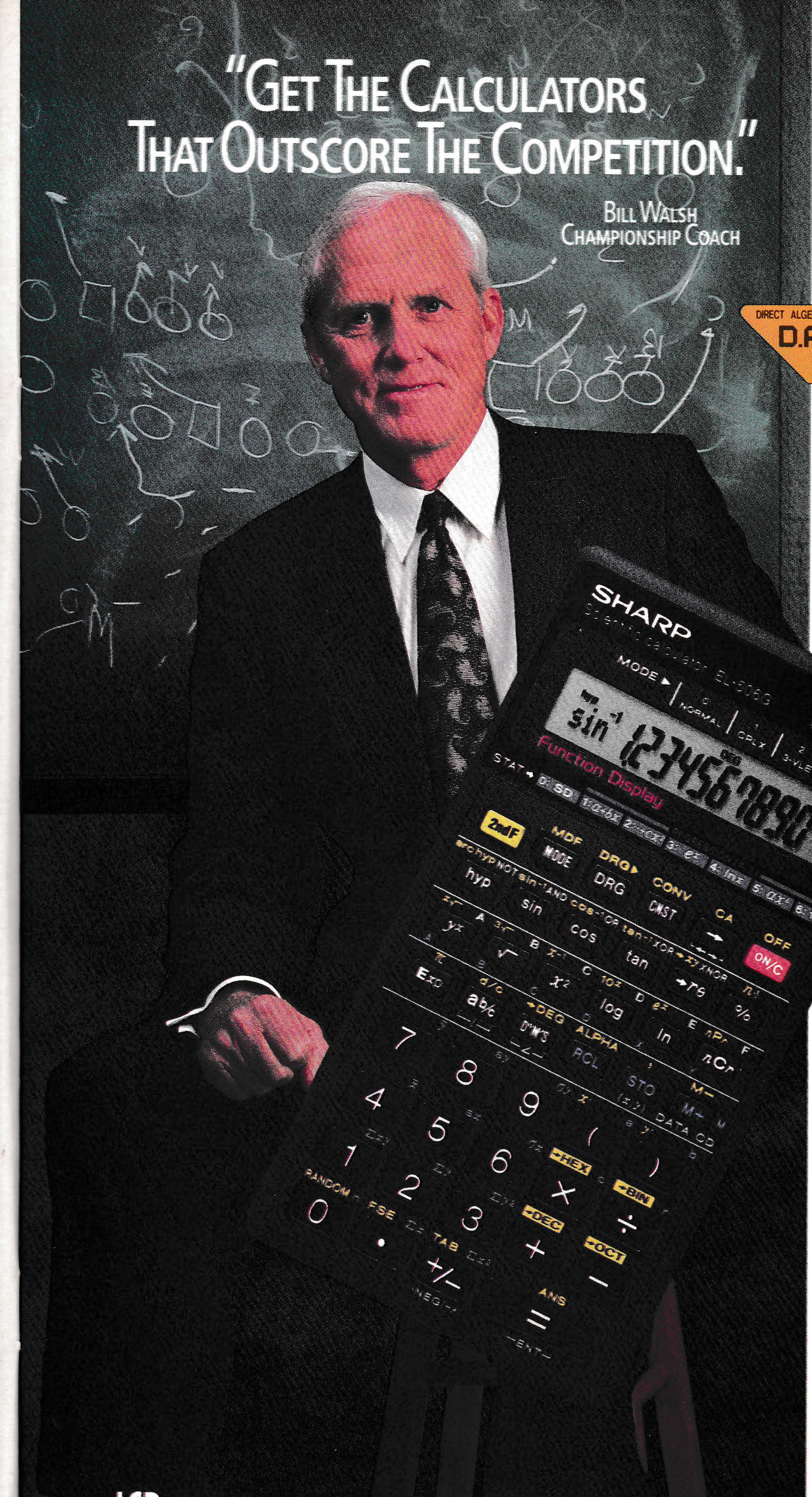
Talk back to us

We actually like it! We want to know what you think about *Quantum*. So

drop us a line at *Quantum*/NSTA, 1840 Wilson Blvd., Arlington VA 22201

or

zip us an electron or two at 72030.3162@compuserve.com.



ONLY SHARP'S D.A.L. SCIENTIFIC CALCULATORS GIVE YOU SUCH AN EASY WAY TO SIMPLIFY MATH.

Sharp introduces the most revolutionary scientific calculator advance in over a decade. D.A.L. It stands for Direct Algebraic Logic. And, it signifies a vast improvement in the algebraic logic currently used in scientific calculators.

DIRECT ALGEBRAIC LOGIC D.A.L.

Only Sharp makes this exciting innovation available on its five non-programmable models.

Developed in response to requests by educators, D.A.L. allows the user to enter equations in the exact order in which they appear on paper. This greatly simplifies entries and reduces errors.

EXAMPLE	D.A.L. KEYSTROKES	WITHOUT D.A.L.
2 log (3 + 2) =	2 log (3 + 2) =	2 × (3 + 2) log =

But most importantly, D.A.L. substantially reduces the time needed to learn a scientific calculator. It also makes teaching calculations faster and easier. These remarkable calculators provide up to 276* functions which frees students from tedious computations. And gives teachers an easier way to make math concepts come alive. What's more, the ETS has even approved usage of these calculators on several standardized tests (including PSAT and SAT).

All of Sharp's D.A.L. models have Sharp's leading LCD technology including a large 10-digit LCD display with 2-digit exponent. This high-contrast display is easy to read and includes 3 or 4 character function display to indicate type of calculation in progress. You get highly durable plastic keys on all models plus protective hard cases on most. Two even offer Twin Power solar/battery operation.

D.A.L. When it comes to simplifying math, nothing scores higher. Which puts Sharp's D.A.L. Scientific Calculators in a class by themselves.

For a free booklet on Sharp's D.A.L. Calculator line, call: 1-800-BE-SHARP.

Or see the following models at your Sharp retailer:

EL-509G, EL-531G, EL-506G, EL-520G, EL-546G

SHARP®
FROM SHARP MINDS COME SHARP PRODUCTS™

LCD BANKING SYSTEMS • CALCULATORS • CASH REGISTERS • COLOR SCANNERS • COMPUTERS • COPIERS • FACSIMILE • LASER PRINTERS • TYPEWRITERS • WIZARD® ORGANIZERS

*Function counts are for Sharp models only. Other manufacturers may use different counting methods.

Circle No. 5 on Reader Service Card

Ones up front in powers of two

And equal rights for pigeons

by Vladimir Boltyansky

RAISING THE NUMBER TWO to integer powers, we notice that ones keep emerging again and again at the beginning of the numbers obtained. How often do they crop up? In other words, *what's the probability that a power of two chosen at random begins with the digit 1?*

Let's make the statement of this problem more exact. In the illustration on the facing page, fifteen numbers are written out, and four of them begin with one. Write these numbers on fifteen cards, shuffle the cards, and choose one card at random. The probability that the number on this card begins with one is 4/15.

Now take the first n powers of two—that is, the numbers $2^1, 2^2, \ldots, 2^n$. Suppose a_n of them begin with one (in decimal notation). Then the probability that a number chosen at random from the numbers $2^1, 2^2, \ldots, 2^n$ begins with the digit 1 is equal to a_n/n. This probability depends on n, the number of the powers we took. We'll see below that, as n grows, the ratio a_n/n approaches a certain number p_1—in other words, a_n/n has a limit. We write

$$a_n/n \to p_1 \text{ as } n \to \infty. \qquad (1)$$

It is this limit that is called *the probability that a power of 2 chosen at random begins with the digit 1*.

You'll understand this article better if you solve the following problems, which illustrate what it means when we say that the "probability that a term of some *infinite* sequence taken at random will satisfy a certain property." In general, this probability is equal to the limit as $n \to \infty$ of the ratio of the number of terms satisfying the property, among the first n terms, to the number n. Notice that this probability is not always well defined, because the limit may not exist.

Problems

1. What is the probability that a random positive integer is divisible by 3?

2. What is the probability that a digit in the decimal expansion of the number 161/222 taken at random is a five? A two?

3. What's the probability that a positive integer taken at random is a perfect square?

Calculating the limit

Clearly, there is exactly *one* two-digit number among the powers of two that begin with one (namely, 16). And there's exactly *one* three-digit number among the powers of two that begin with one (namely, 128). The same is true for four-digit numbers, too.

In general, for any $k > 1$, there is exactly one k-digit power of two that begins with one. Indeed, the *smallest* k-digit number that is the power of two must necessarily begin with one—otherwise, dividing it by 2, we would get a smaller k-digit power of two. And no other k-digit power of two begins with one, because the next power of two has 2 or 3 in the first place, the following one begins with an even greater digit, and so on, until we run into a $(k + 1)$-digit number.

From this it follows that if 2^n has m digits, then there are exactly $m - 1$ numbers beginning with one among the powers $2^1, 2^2, \ldots, 2^n$—that is, $a_n = m - 1$. But the number 2^n has m digits if and only if

$$10^{m-1} \le 2^n < 10^m,$$

or

$$m - 1 \le n \log 2 < m.$$

Substituting $a_n + 1$ for m, and rearranging these inequalities, we get

$$\log 2 - \frac{1}{n} < \frac{a_n}{n} \le \log 2.$$

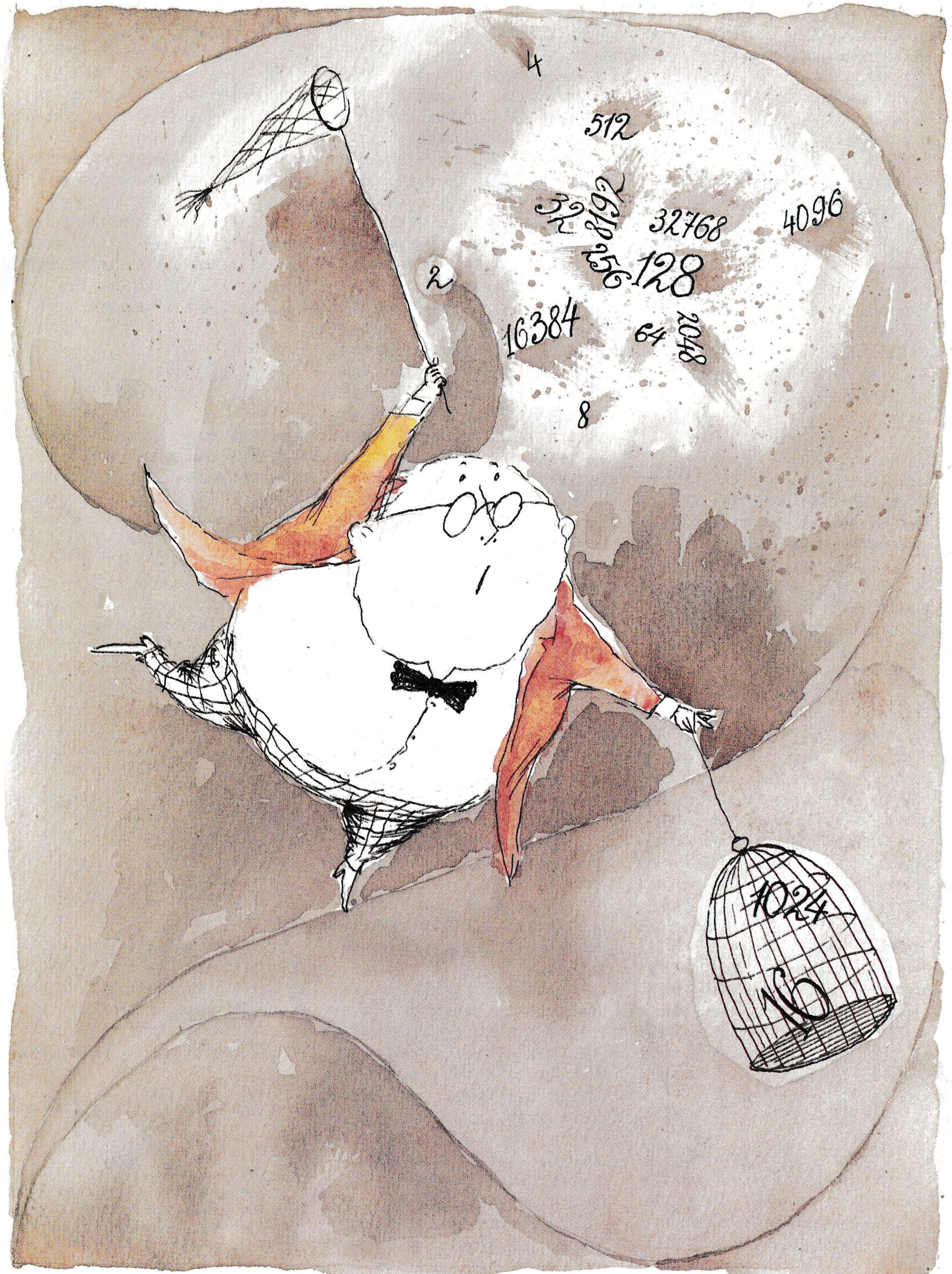
Now it's clear that

$$\frac{a_n}{n} \to p_1 = \log 2 = 0.30103\ldots$$

Thus, a little more than 30 percent of the powers of two begin with one.

Problems

4. Denote by p_q the probability that a random power of two begins with the digit q. Assuming that


4
512
32
8192
32768
4096
2
256
128
16384
64
2048
8
1024
16

these probabilities are well defined, prove the equalities $p_2 + p_3 = p_1$; $p_4 + p_5 = p_2$; $p_6 + p_7 = p_3$; $p_8 + p_9 = p_4$.

5. Prove that $p_4 = 1 - 3\log 2 = 0.096\ldots$.

The last problem shows that the powers of two begin with 1 more than three times as often as with 4.

The general problem

Now let's try to answer a more general question: *what is the probability that a power of a given positive integer l taken at random begins with a digit q (in decimal notation)?*

Suppose the decimal notation of the number l^n begins with q:

$$q \cdot 10^m \le l^n < (q+1)10^m$$

for a certain integer m. Dividing these inequalities by $q \cdot 10^m$, and taking the logarithm to the base 10, we arrive at

$$0 \le (n\log l - \log q) - m < \log\frac{q+1}{q}. \qquad (2)$$

Now, $(q + 1)/1 = 1 + 1/q \le 2 < 10$, so the number on the right-hand side is less than 1. Inequalities (2) show that the fractional part of the number $n\log l - \log q$ is less than $\log(q + 1)/q$. (Let me remind you that the *fractional part* $\{x\}$ of a number x is the difference between x and its integer part $[x]$: $\{x\} = x - [x]$, and $[x]$ is the largest integer not exceeding x.) Conversely, if the fractional part of the number $n\log l - \log q$ is less than $\log(q + 1)/q$—that is, if inequalities (2) are valid for some integer m—then the decimal notation of l^n begins with q. So our problem is equivalent to this: *what is the probability that for a positive integer n taken at random,*

$$\{n\log l - \log q\} < \log\frac{q+1}{q}?$$

The following theorem will help us solve this problem.

Fractional Parts Theorem. *Let α be an irrational number and β an arbitrary real number; let I be an interval of length h contained in the segment [0, 1]. Consider the infinite arithmetic sequence $\alpha + \beta$, $2\alpha + \beta$, ..., $n\alpha + \beta$, ... Then the probability that the fractional part of an arbitrary term of this sequence belongs to the interval I is equal to h.*

I'll give a proof of this theorem in the last section. Now I'll show how it's applied to our problem.

First of all we notice that if l is a power of ten, then all the powers of l begin with one: in this case the problem is trivial.

So let's assume that l is not a power of ten. Then the number $\alpha = \log l$ is *irrational* (see problem 6 below). Set $\alpha = n\log l$, $\beta = -\log q$ in the statement of the Fractional Parts Theorem. The theorem then tells us that the probability that a term of the sequence $n\log l - \log q$ ($n = 1, 2, \ldots$) has a fractional part lying in the interval $[0, \log(q + 1)/q)$ is equal to $\log(q + 1)/q$. And this yields the solution we seek: *the probability that a randomly chosen power of a number l (not a power of ten) begins with the digit q is equal to $\log(q + 1)/q$.*

Unexpectedly, this probability does not depend on l! For instance, powers of two and powers of three begin with 1 equally often (namely, with a probability of $\log 2$).

Problems

6. Prove that if a positive integer l is not a power of 10, then $\log l$ is irrational.

7. Compute the probabilities p_2, ..., p_9 defined in problem 4. Give a new proof of the equations in that problem.

8. Find the probability that a randomly chosen power of two begins with the combination of digits 1000.

9. Find the probability that a randomly chosen power of a number l (distinct from a power of ten) begins with a given combination of digits $q_1q_2\ldots q_r$, where $q_1 \ne 0$. Derive, in particular, the fact that a power of l can begin with *any* combination of digits $q_1q_2\ldots q_r$ ($q_1 \ne 0$).

10. (a) Given a randomly selected power of a number l, $l \ne 10^n$, prove that the probability that its *second* digit is 0 equals

$$p_0^{(2)} = \log(11 \cdot 21 \cdot 31 \cdot \ldots \cdot 91) - \log(9!) - 9.$$

Hint: this probability is the sum of the probabilities that a power of l begins with 10, 20, ..., 90.

(b) Given a randomly selected power of a number l, $l \ne 10^n$, prove that the probability that its kth digit (starting from the left) is q (for any $k > 1$, $q = 0, 1, \ldots, 9$) equals

$$p_q^{(k)} = \sum \log\left(1 + \frac{1}{q + 10i}\right),$$

where the sum is taken over all i from 10^{k-2} through $10^{k-1} - 1$.

11. Using the result from problem 10(b), prove that $p_q^{(k)} \to 1/10$ as $k \to \infty$. Hint: use the estimate $\ln(1 + x) < x$ for $x > 0$,[1] and the relations

$$\log\left(1+\frac{1}{10i}\right) - \log\left(1+\frac{1}{10i+q}\right) < \log\left(1+\frac{q}{100i(i-1)}\right) = \frac{1}{\ln 10}\cdot\ln\left(1+\frac{q}{100i(i-1)}\right) < \frac{1}{\ln 10}\frac{q}{100i(i-1)} = \frac{q}{100\ln 10}\left(\frac{1}{i-1} - \frac{1}{i}\right).$$

12. Generalize problems 9 and 10(b) to the case where the powers of l are written in the number system with a base $b > 1$. In particular, the formula of problem 10(b) in the binary system for $q = 0$ takes the form

$$p_0^{(k)} = \log_2\left(\frac{2^{k-1}+1}{2^{k-1}}\cdot\frac{2^{k-1}+3}{2^{k-1}+2}\cdot\ldots\cdot\frac{2^k-1}{2^k-2}\right).$$

[1]This inequality can be proven, for instance, by using the methods from "Derivatives in Algebraic Problems" in this issue.—*Ed.*

So near and yet so far

As another illustration of how the Fractional Parts Theorem works, consider this problem:

Let a number α be irrational. Prove that $\cos n\alpha\pi > 0.999$ for a certain positive integer n.

Notice that the inequality $\cos n\alpha\pi > 0.999$ is equivalent to these:

$$2m\pi - \varepsilon < n\alpha\pi < 2m\pi + \varepsilon,$$

where $\varepsilon = \cos^{-1} 0.999$ and m is an integer, or

$$m < \frac{n\alpha}{2} + \frac{\varepsilon}{2\pi} < m + \frac{\varepsilon}{\pi}.$$

In other words, the inequality $\cos n\alpha\pi > 0.999$ is valid if and only if the number $\{n\alpha/2 + \varepsilon/2\pi\}$ belongs to the interval $(0, \varepsilon/\pi)$. According to the Fractional Parts Theorem, a positive integer n taken at random satisfies this condition with a probability of $\varepsilon/\pi = \cos^{-1} 0.999/\pi \cong 0.014$. So for a sufficiently large N approximately 1.4 percent of the numbers $1, 2, \ldots, N$ satisfy this condition.

Of course, here we could take any number less than 1 instead of 0.999. This means that the number $\cos n\alpha\pi$, for any fixed irrational α, approaches 1 arbitrarily close, though it never becomes exactly equal to 1.

Problems

13. Arcs of length 1 are marked off one after another starting from an arbitrary point A_0 on a circle of radius 1. Let $A_1, A_2, \ldots$ be the successive endpoints of these arcs. Prove that any arc of this circle contains a point A_i.

14. Prove that the function $f(x) = \sin x + \sin \alpha x$ is not periodic for any irrational α.

15. Consider two infinite arithmetic sequences $a_1, a_1 + d_1, a_1 + 2d_1, \ldots$ and $a_2, a_2 + d_2, a_2 + 2d_2, \ldots$. The numbers d_1 and d_2 are positive and their ratio d_1/d_2 is irrational. Does there exist a term in one sequence and a term in the other sequence such that the absolute value of their difference is less than 0.000001?

16. Consider the set of circles of radius ε whose centers are all the points with integer coordinates—a "forest of radius ε." Draw a line that makes an angle ϕ with the x-axis such that $\tan \phi$ is an irrational number. Prove that this line will intersect the forest no matter how small the radius ε is (see figure 1).

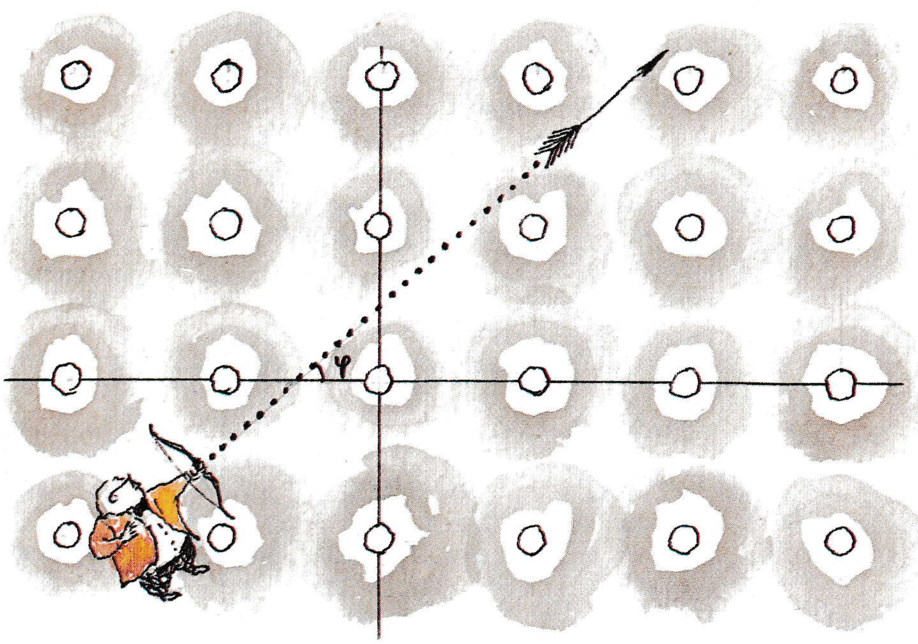


Figure 1

A proof

Our proof of the Fractional Parts Theorem is based on the following two assertions:

1. *For any irrational number α and any integer $l > 0$ there exists a positive integer p such that $p\alpha$ differs from the nearest integer m by no more than $1/l$:*

$$|p\alpha - m| < \frac{1}{l}.$$

2. *As before, let l be an integer > 0. Consider an arithmetic sequence with a difference δ such that $|\delta| < 1/l$. Take its first n terms and suppose that f_n of them have fractional parts in the interval I of length h, from the statement of the theorem (we shall call such terms "favorable"). Then for all sufficiently large n,*

$$\left|\frac{f_n}{n} - h\right| < \frac{2}{l}.$$

Now I'll show how the Fractional Parts Theorem is deduced from these two statements. Recall that the Fractional Parts Theorem concerns a sequence of the form $x_1 = \alpha + \beta$, $x_2 = 2\alpha + \beta, \ldots, x_n = n\alpha + \beta$. Write out the first n terms of the sequence and circle every pth of them starting with x_1, where p is the number from statement (1) chosen for a certain $l > 0$:

$$x_1, x_2, \ldots, x_p, x_{p+1}, x_{p+2}, \ldots, x_{2p+1}, \ldots.$$

The circled numbers form an arithmetic sequence with a difference $p\alpha$. Since we're interested only in their fractional parts, it doesn't matter is we add an integer to the terms of the sequence or subtract an integer. We can therefore replace this sequence by the sequence with the same first term x_1 and the difference $\delta = p\alpha - m$, where m is the integer from statement (1). By statement (1), $|\delta| < 1/l$, so statement (2) implies, for all sufficiently large n, that

$$n^{(1)}\left(h - \frac{2}{l}\right) < f^{(1)} < n^{(1)}\left(h + \frac{2}{l}\right),$$

where $n^{(1)}$ is the number of all circled terms, and $f^{(1)}$ the number of favorable circled terms, among $x_1, x_2, \ldots, x_n$. Now let's do the same thing

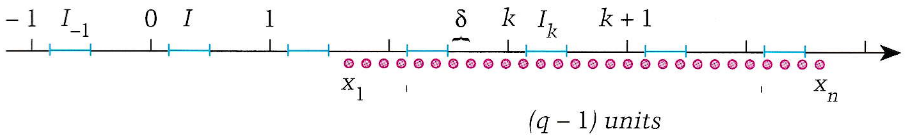


Figure 2

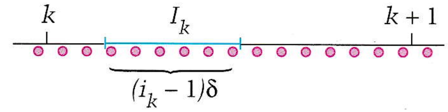


Figure 3

starting from x_2, then from x_3, and so on. We'll get p sequences, each satisfying similar inequalities (with $n^{(k)}$ and $f^{(k)}$ for the sequence starting with x_k, $k = 1, 2, \ldots, p$). Summing all these inequalities, and taking n large enough to make them all true, we have

$$n\left(h-\frac{2}{l}\right)<f_n<n\left(h+\frac{2}{l}\right),$$

where $f_n = f^{(1)} + \ldots + f^{(p)}$ is the number of favorable terms in the given sequence $x_1, x_2, \ldots, x_n$. Therefore, this sequence satisfies the expression $|f_n/n - h| < 2/l$ for any $l > 0$ and all sufficiently large n, or $f_n/n \to h$ as $n \to \infty$. This is what the Fractional Parts Theorem states. It remains to prove statements (1) and (2) to complete the proof of our theorem.

Proof of statement (1). Divide the segment [0, 1] into l equal parts and consider the numbers $\{\alpha\}, \{2\alpha\}, \ldots, \{(l+1)\alpha\}$. Since α is irrational, these are all different. Using the "pigeonhole principle" (see, for instance, "Pigeons in Every Pigeonhole" in the January 1990 issue of *Quantum*), we can think of these $l + 1$ numbers as "pigeons," and the l sections of the segment [0, 1] as "pigeonholes." By this very useful (though obvious) principle, two pigeons must sit in the same pigeonhole—that is, there are two numbers $p_1\alpha$ and $p_2\alpha$, $p_1 > p_2$, whose fractional parts differ by no more than $1/l$. So for $p = p_1 - p_2$ there is an integer m such that $|p\alpha - m| = |(p_1\alpha - p_2\alpha)| - m < 1/l$, and we're done. (In terms of pigeons, our theorem may be viewed as the "pigeonhole principle for infinitely many pigeons," and it says that our "fractional-part pigeons" are distributed uniformly in their pigeonholes.)

Proof of statement (2). Suppose that the first n terms of the given sequence are $x_1, x_2, x_3, \ldots, x_n$ (see figure 2). For any integer k, let I_k be the image of the interval I translated k units, so that $I_k \subset [k, k+1]$. Let q be the number of intervals I_k that lie entirely between x_1 and x_n (that is, q does not count any interval I_k that x_1 and x_n may fall in). Then a case-by-case analysis will show that no matter where x_1 and x_n fall, $q - 1 < |x_n - x_1| = (n-1)\delta < q + 2$. It follows that

$$n\delta - 3 < (n-1)\delta - 2 < q < (n-1)\delta + 1 < n\delta + 1. \quad (3)$$

Let i_k be the number of the terms x_i in the interval I_k. Then (see figure 3) $(i_k - 1)\delta \le h$, or $i_k \le h/\delta + 1$. Summing these inequalities over all (at most $q + 2$) intervals I_k that contain the terms x_i, $i = 1, \ldots, n$, we get the following upper bound for the number f_n of favorable terms:

$$\begin{aligned} f_n = \sum i_k &\le (q+2)\left(\frac{h}{\delta}+1\right) \\ &< (n\delta+3)\frac{h+\delta}{\delta} \\ &< n(h+\delta)+\frac{6}{\delta} \end{aligned}$$

(here we use the right-hand portion of inequality (3), $h \le 1$ and $\delta < 1/l \le 1$). On the other hand, for each of the q intervals I_k between x_1 and x_n, $(i_k + 1)\delta > h$, so $i_k > h/\delta - 1$. Summing these yields

$$\begin{aligned} f_n > q\left(\frac{h}{\delta}-1\right) &> (n\delta-3)\frac{h-\delta}{\delta} \\ &> n(h-\delta)-\frac{3}{\delta} \end{aligned}$$

(here we use the left-hand portion of inequality (3)). Combine the two bounds:

$$\left|f_n - nh\right| < n\delta + \frac{6}{\delta} < n\left(\frac{1}{l}+\frac{6}{\delta n}\right).$$

Now it remains to divide by n and choose n such that $6/\delta n < 1/l$ (or $n > 6l/\delta$).

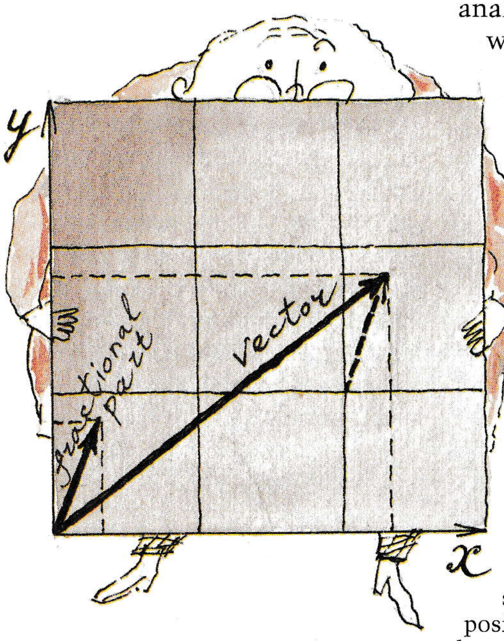


Figure 4

[2]In this figure, and throughout the proof, the difference δ of the sequence is assumed to be positive. The only change that should be made in the case of $\delta < 0$ is to replace δ with $|\delta|$ in all formulas.

Problems

17. Prove the following two-dimensional generalization of the Fractional Parts Theorem:

Fix a coordinate frame on the plane. Define the *fractional part* of a vector **v** with coordinates (x, y) as the vector $\{\mathbf{v}\}$ with coordinates $(\{x\}, \{y\})$ (fig. 4). Let M be a polygon contained in the square S with the vertices (0, 0), (0, 1), (1, 0), (1, 1). A vector **u** will be

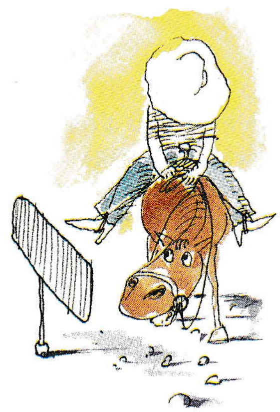
CONTINUED ON PAGE 45

Just for the fun of it!

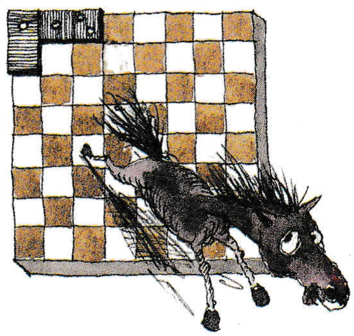
B96

Logic behind coincidences. I've thought of a three-digit number such that each of the numbers 543, 142, and 562 coincides with it in exactly one decimal location. Guess what this number is. (V. Proizvolov)

B97



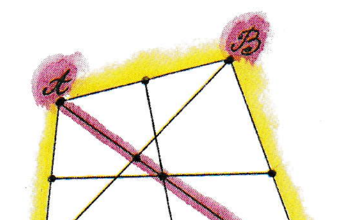
Entropy and Tesseract. While driving down an unfamiliar road, I noticed a sign that said: "Entropy—150 ents, Tesseract—110 tesses." Apparently the residents of Entropy measure distance in units called "ents," and the folks in Tesseract measure distance in "tesses." I drove further down the road. Before I came to either town, I saw another sign: "Entropy—10 ents, Tesseract—26 tesses." Find the point between Entropy and Tesseract where the distance from Entropy, measured in ents, equals the distance from Tesseract, measured in tesses. (T. Stickels)



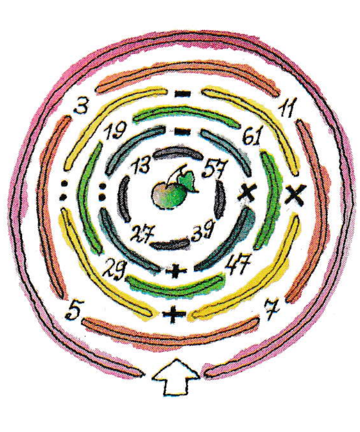
B98

Tiling with dominoes. A chessboard is covered with 32 dominoes so that each domino covers exactly two squares. After counting the dominoes oriented horizontally and vertically, it was found that there are evenly many dominoes with each orientation. Will this be true for any covering of the chessboard with 32 dominoes? (V. Proizvolov)

B99



Halving it all (cont'd). Three line segments are drawn in a convex quadrilateral: a diagonal and both midlines (the segments that join the midpoints of opposite sides). The other diagonal divides one of these segments in half. Prove that it bisects the other two segments as well. (N. Netsvetayev, V. Dubrovsky)



B100

Reaching one hundred. Find a path to the center of the maze in the figure such that you get 100 by performing the operations along this path. (A. Larionov)

ANSWERS, HINTS & SOLUTIONS ON PAGE 58

Xe
N2
Ar
He
H2
O2
CO2
N2
NH3
CH4
He
H2

A simple capacity for heat

As usual, it's not as simple as it seems

by Valeryan Edelman

I'M SURE THAT MANY OF you will look at the title of this article and shrug your shoulders: "What's so interesting about *that*? Yes, we need to know heat capacities in order to calculate the thermal energy required to raise the temperature of an object. It's certainly essential for technology, so people were found who were willing to spend gobs of time measuring the specific heats for various materials.[1] (These measurements aren't so difficult in principle—most of us have even studied a bit of calorimetry in school.) Then they compiled tables and reference books—anyone can use them, and there's nothing more to think about. So, you've come up with a trite and rather boring topic." And yet . . .

If you look through the most weighty scientific journals, you can always find papers in which specific heat is studied, and not always that of newly created materials. Often quite ordinary materials are investigated, often under unusual conditions. So, what's the big secret? Why—in this age of lasers, high-energy physics, microelectronics, thermonuclear synthesis, and so on—hasn't the interest of physicists in this apparently routine subject faded? The answer is that the specific heat is closely related to the structure of matter and the dynamics of the motion of subatomic particles. Sometimes it's a measurement of specific heats that makes it possible to know at least something about the nature of things when the most modern methods are useless.

If you want to see this connection with your own eyes, you won't have far to look. Figure 1 will help convince you. It shows how the specific heat of ordinary water changes with temperature. Of course, the first thing that impresses anyone is the jump at 0°C—the temperature at which water changes from solid to liquid. But that's not the whole story: in both the solid and liquid phases, the specific heat depends on temperature in a complicated way (see the blowup of the water portion of figure 1). Mind you, this is the case with *water*, which not so long ago served as the standard for specific heats!

Modern science can explain what occurs with water, but we'll not study this phenomenon. It's better to start with the simplest thing we know: ideal gases.

Ideal gases

Strictly speaking, this subhead isn't quite accurate, since we're going to be talking about the actual gases that are ideal in one sense only—at

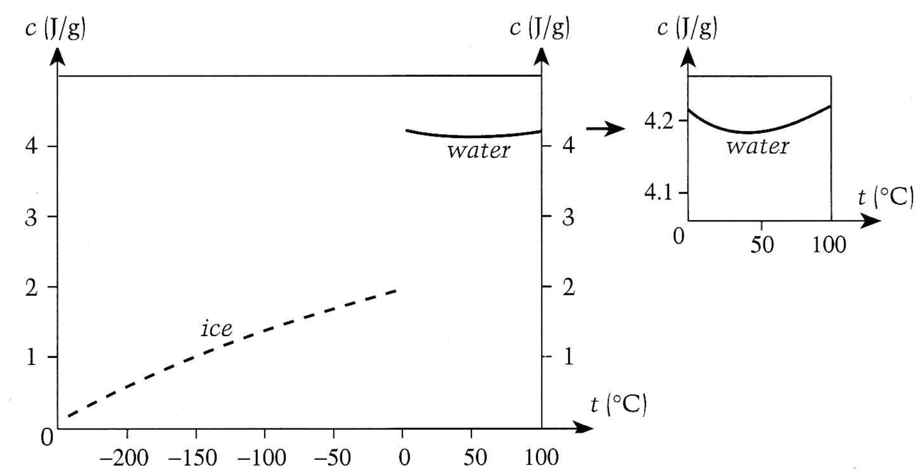


Figure 1

[1]The heat capacity of an object is the amount of heat required to raise the temperature of the object by one degree and is a property of the particular object. The specific heat is the heat capacity per unit mass and is an intrinsic property of the material from which the object is made.—*Ed.*

Art by Dmitry Krymov

Table 1

gas	He	Ar	Xe	H_2	N_2	O_2	CO_2	NH_3	CH_4
c_{sp} (J/g · K)	3.15	0.31	0.096	10.26	0.74	0.66	0.65	1.62	1.68

Specific heat of gases at room temperature and constant volume.

room temperature the following law holds with great precision:

$$PV = NkT.$$

Here, as usual, P, V, and T are pressure, volume, and absolute gas temperature; N is the number of molecules; and $k = 1.38 \cdot 10^{-23}$ J/K is the Boltzmann constant. The average energy of translational motion of a gas molecule is equal to

$$E_{tr} = \frac{3}{2}kT.$$

Let's look up the values for the specific heat at constant volume for several gases in a standard reference book (see table 1). At first glance it's difficult to see any regularity in these numbers. But let's take our time and put these values into another form. We'll be guided by the fact that one gram of different gases contains different numbers of molecules. The formula for E_{tr} shows that we need to compare the values for one molecule. It's not hard to recalculate the numbers in table 1. Remember that a mole of any substance has the same number of molecules (Avogadro's number: $N_A = 6.02 \cdot 10^{23}$). It's easy to find the heat capacity of one mole—that is, the molar heat capacity: $c_\mu = c_{sp}\mu$, where μ is the molar mass of the substance. So the heat capacity for one molecule is

$$c_{mol} = c_{sp}\frac{\mu}{N_A}.$$

Naturally this value is very small, and it will be convenient for us to compare the values of $c'_{mol} = c_{mol}/k$ (it's easy to convince ourselves that c'_{mol} is dimensionless—it's just a number).

Let's calculate the values of c'_{mol} for the gases in table 1 and see what happens. What is striking about table 2 is that c'_{mol} is the same for all the monatomic gases. In other words, the molecular heat capacity of all the single atomic gases is the same and equal to $1.5k$—that is, $\frac{3}{2}k$. But this coefficient $\frac{3}{2}$ is a very familiar number: it's the same coefficient that appears in the formula describing the average energy of translational motion for molecules in an ideal gas. Since the molecular heat capacity $c_{mol} = \Delta E/\Delta T$, we get

$$c_{mol} = \frac{3}{2}k\frac{\Delta T}{\Delta T} = \frac{3}{2}k.$$

The result is remarkable: for helium, neon, and argon, all the heat is completely transformed into the kinetic energy of translational motion of the atoms. One might imagine that the atoms could rotate, but it's evident that no heat goes into this motion, and so there is no heat capacity associated with this rotation. Generally speaking, this conclusion holds at moderate temperatures only (the data in tables 1 and 2 were obtained at such temperatures). At very high temperatures (thousands of degrees), things get a little complicated. Experiments and theory both show that rotations can be induced. Still, we won't make life more difficult at this point. Even so, questions crop up: what about the molecular heat capacity of gases whose molecules consist of two or more atoms? Their molecular heat capacities are somewhat greater than $\frac{3}{2}k$. It's curious that for diatomic gases—hydrogen, oxygen, nitrogen—the extra amount per atom is very close to $\frac{1}{2}k$. However, for multiatomic gases the situation is more complicated: the extra amounts are equal to ~$0.65k$/atom for CO_2 and only ~$0.34k$/atom for CH_4. So maybe it's not just a matter of the number of atoms per molecule.

Let's approach this problem from another direction and see what different types of motion are possible for the molecules. A diatomic molecule can be represented as in figure 2: atoms connected by a spring. Translational motion of the molecule can be described as the motion of its center of mass along three mutually perpendicular axes x, y, z (see figure 2). A molecule can be rotated about the y-axis and the z-

Table 2

gas	He	Ar	Xe	H_2	N_2	O_2	CO_2	NH_3	CH_4
c_{sp} (J/g · K)	3.15	0.31	0.096	10.26	0.74	0.66	0.65	1.62	1.68
$c'_{mol} = c_{mol}/k$	1.50	1.50	1.50	2.45	2.49	2.53	3.42	3.30	3.23

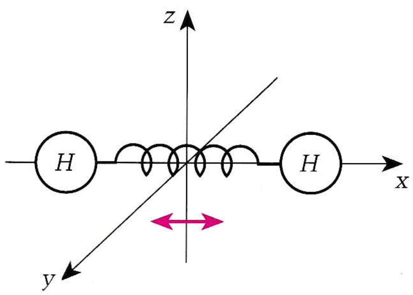


Figure 2

axis—that is, perpendicular to the spring. Molecular rotation about the x-axis (along the spring) is excluded. This rotation is analogous to rotating a monatomic molecule, and we've seen that energy does not go into such rotations. Finally, the atoms themselves can oscillate along the x-axis toward each other.

Thus, for a diatomic molecule, taking translational motion along three coordinate axes into account, we find that there are six possible types of motion (they are also called "degrees of freedom"). If a molecule is made of n atoms and $n \geq 3$, it becomes difficult to paint a similar picture. But there's a simple rule that allows us to calculate the number of degrees of freedom: the total number is equal to three coordinates for each atom times the number of atoms, or $3n$. This includes three degrees for the translational motion and three degrees for the rotations about three mutually perpendicular axes.

Using this recipe, let's calculate the number of possible motions for different gas molecules and add three lines to table 2 (see table 3). Now we'll look at the multiatomic gases CO_2, NH_3, and CH_4. Each of these molecules can be rotated about the three axes, but the number of possible types of oscillation for these molecules is different: there are three types of oscillation for CO_2, six for NH_3, and nine for CH_4. Yet the molecular heat capacities of these gases are almost identical! It's reasonable to assume that the energy added to the molecule is expended on translational and rotational motions, not atomic oscillations. In other words, oscillations do not contribute to the molecular heat capacity. But then it's also reasonable to exclude the oscillations in diatomic molecules, arguing that the additional (as compared to monatomic molecules) molecular heat capacity is exclusively related to the rotations.

As we can see from table 3, for the diatomic molecules at room temperature, this addition is very close to $\frac{1}{2}k$ for each rotational degree of freedom. If the same rule is applied to multiatomic molecules, then the molecular heat capacity is equal to $3k$. (Actually, it's somewhat higher, but we won't pay much attention to this for the time being.)

We've arrived at an interesting result: for each of all possible motions of a molecule as a whole (be it a displacement along one of the coordinate axes or a rotation about one of these axes), there is an addition to the molecular heat capacity of $\frac{1}{2}k$. Physicists call this conclusion the equipartition theorem.

Oscillations: theorembusters?

So—if it's a theorem, why isn't it universal? Why is an exception made for oscillations? We can certainly say that the "extra" molecular heat capacity in multiatomic molecules is related to oscillations, but the contributions from oscillations for CO_2 ($0.15k$ for each oscillation) and CH_4 ($0.025k$ per oscillation) are so different that it's not worth talking about equipartition.

The situation gets even more complicated if we look at a huge "supermolecule"—that is, a piece of a solid body. All the molecules in solids are located at the nodes of the crystal lattice and can't move translationally or rotationally. The only possible kind of motion (if we neglect motion of the object as a whole) is atomic oscillation about their equilibrium positions. Therefore, the molar heat capacity of solids is related to the oscillational excitation. This molar heat capacity is not insignificant; almost all crystals have nearly identical molar heat capacities under ordinary conditions: close to 25 J/mole · K. This is known as the Dulong and Petit law. For example, here are the molar heat capacities of some solids (in joules per mole · K):

Table 3

gas	He	Ar	Xe	H_2	N_2	O_2	CO_2	NH_3	CH_4
c_{sp} (J/g · K)	3.15	0.31	0.096	10.26	0.74	0.66	0.65	1.62	1.68
c_{mol}/k	1.50	1.50	1.50	2.45	2.49	2.53	3.42	3.30	3.23
translational degrees of freedom	3	3	3	3	3	3	3	3	3
rotational degrees of freedom	0	0	0	2	2	2	3	3	3
oscillational degrees of freedom	0	0	0	1	1	1	3	6	9

aluminum	24.4
silver	25.2
copper	24.6
gold	26.5
lead	26.6

It's easy to calculate that the heat capacity per atom in a solid is $3k$ on average—that is, for each oscillation. To understand this, we need to look at the experimental results.

First, it would be nice to know whether the molecular heat capacity of a gas depends on temperature. If it does, how? Let's look at several gases: helium, hydrogen, and oxygen (fig. 3). Right away we can see that for the monatomic gas (helium), the molecular heat capacity is constant. However, the behavior of the heat capacity for hydrogen is quite another story. At low temperatures it's equal to $1.5k$—that is, hydrogen behaves like a monatomic gas. At $T \cong$ 70 K the molecular heat capacity increases; at $T \cong 250$ K it attains a new, almost constant value: $c_{mol} = 2.5k$. But right around 1,000 K a new increase kicks in, and at 2,000 K the molecular heat capacity for hydrogen becomes greater than $3k$. With a further increase in temperature, the heat capacity continues to increase, but we won't examine this region, because too many other phenomena take place when a gas is heated to such high temperatures.

Let's now see how oxygen behaves over the same temperature range. It's impossible to measure the heat capacity for oxygen in the gas phase at temperatures markedly lower than 100 K, since it condenses to a liquid. However, if this were not so, then at low temperatures the molecular heat capacity of gaseous oxygen would be equal to $1.5k$. The experimental curve begins at $2.5k$ and increases to $3.5k$ by 2,000 K.

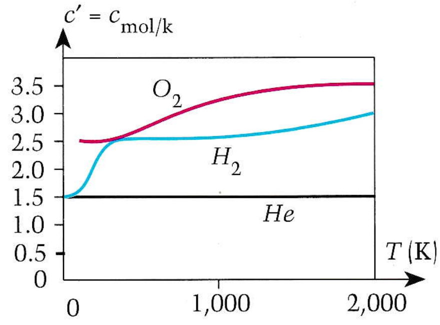


Figure 3

What conclusions can we draw from these experimental data?

1. There is always a translational motion of the gas molecules, and the heat capacity related to it does not depend on temperature.
2. Rotations and oscillations vanish at low temperatures—they're "frozen," so to speak. At room temperature rotations are present, but oscillations are still frozen. For example, rotation for hydrogen molecules is "unfrozen" at $T \cong 100$ K.
3. There is only one type of oscillation in the oxygen molecule, but the molecular heat capacity increases by $1k$, not by $\frac{1}{2}k$ as we might have guesed. Therefore, the molecular heat capacity for each unfrozen oscillation is equal to $1k$.

We've come across this number before: it's the heat capacity assigned to each oscillation in a solid! What can we make of these results?

Note that the factor $\frac{1}{2}$ appears in the translational motion of molecules, where the energy is purely kinetic energy; and that the factor of 1 appears in the oscillations, where the energy is both kinetic and potential. Since the energy in oscillations changes back and forth between kinetic energy and potential energy, the average value of the kinetic energy is equal to that of the potential energy. If the total energy of the oscillation is kT (that is, the contribution to the molecular heat capacity is $1k$), the average value of the kinetic and potential energies is $\frac{1}{2}kT$—again we've returned to $\frac{1}{2}$. These conclusions agree very well with many experiments and are confirmed by the theory based on quantum mechanics.

Of course, the case considered here is the simplest of all possible ones. We could examine it in more detail. In calculating the degrees of freedom, we regarded an atom as a single particle, but that's not the case. Each atom has its own internal degrees of freedom for each electron. However, hundreds of thousands of degrees are needed to defrost them. As a matter of fact, plasma physicists study gases in which electron motion is defrosted.[2]

We can go even further: the nuclei consists of individual neutrons and protons, but even these aren't elementary. However, to excite such thermal motion, millions and billions of degrees, or even more, are necessary. Here's where the path into our deep past begins—into the history of the birth of stars, galaxies, and the universe itself. And the first step on this path is an understanding of ideal gases and their laws. ◘

[2]See "The Fourth State of Matter" in the last issue of *Quantum*.—*Ed.*

HOW DO YOU FIGURE?

Challenges in physics and math

Math

M96

Pentagon slashing. Does there exist a (nonconvex) pentagon that can be cut into two congruent pentagons? (S. Hosid)

M97

Arcs in opposition. A circle is divided into $3k$ arcs, k of which are of unit length, k others are of length 2, and the remaining k are of length 3. Prove that at least two of the $3k$ endpoints of the arcs are diametrically opposite. (V. Proizvolov)

M98

Integral solution. Find all positive integer solutions (x, y) of the equation $x^y - y^x = x + y$. (A. Zaychik)

M99

Complete coverage. One thousand squares are drawn on the coordinate plane such that their sides are parallel to the coordinate axes. Prove that one can choose some of these squares in such a way that the center of every given square is covered by at least one and no more than four of the chosen squares. (A. Plotkin)

M100

Polygons follow rules. For what n can a regular n-gon be drawn on paper ruled with equally spaced parallel lines so that all its vertices lie on the lines? (N. Vasilyev)

Physics

P96

Snow catcher. A woman skiing across a field with a speed v = 20 km/h in a heavy snowfall observed that her mouth encountered $N_1 = 50$ snowflakes per minute. After turning back, she noticed that only $N_2 = 30$ snowflakes hit her mouth per minute when skiing with the same speed. Estimate the visibility during this time, assuming $S = 24\ \text{cm}^2$ for the area of the skier's mouth in the direction of travel and d = 1 cm for the average diameter of the snowflakes. (M. Semyonov)

P97

Electrical cube. A set of 28 identical resistors R connect all the corners of a cube. Calculate the equivalent resistance between two adjacent corners. (C. Wörner)

P98

Say "seaweed." The objective of a camera for underwater photography is a thin plano-convex lens with a diameter D = 10 mm made of glass with a refractive index n = 1.8. Its convex surface has a radius of curvature R = 7.5 cm and is on the water side of the lens. Estimate the distance F from the lens to the photographic film needed to shoot distant objects underwater. The refractive index of water is $n_w = 1.3$. The camera is filled with air with a refractive index of 1. (V. Pogozhev)

P99

Draining experience. What energy is dissipated in the circuit shown in figure 1 when the switch K is toggled? The values of all the components are known. (S. Zhuravlyov, V. Peterson, V. Pogozhev, M. Semyonov)

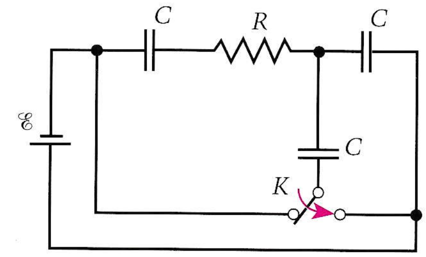


Figure 1

P100

One ring, three strings. A thin homogenous ring of radius $R = L/2$ is suspended by three identical vertical pieces of nonstretchable string of length L, their fixed ends forming a horizontal equilateral triangle (fig. 2). Estimate the period of small torsional oscillations of the ring. (S. Krotov)

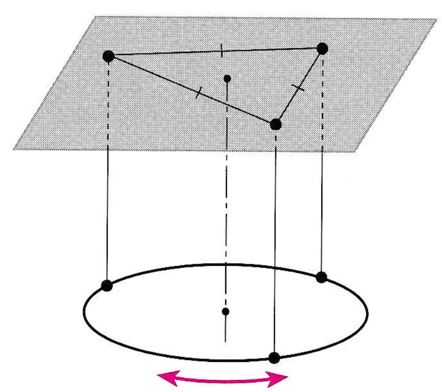


Figure 2

ANSWERS, HINTS & SOLUTIONS ON PAGE 54

Derivatives in algebraic problems

Taking the implicit route to counting roots

by Alexander Zvonkin

TRY TO ANSWER THIS QUESTION: "How many roots does the equation

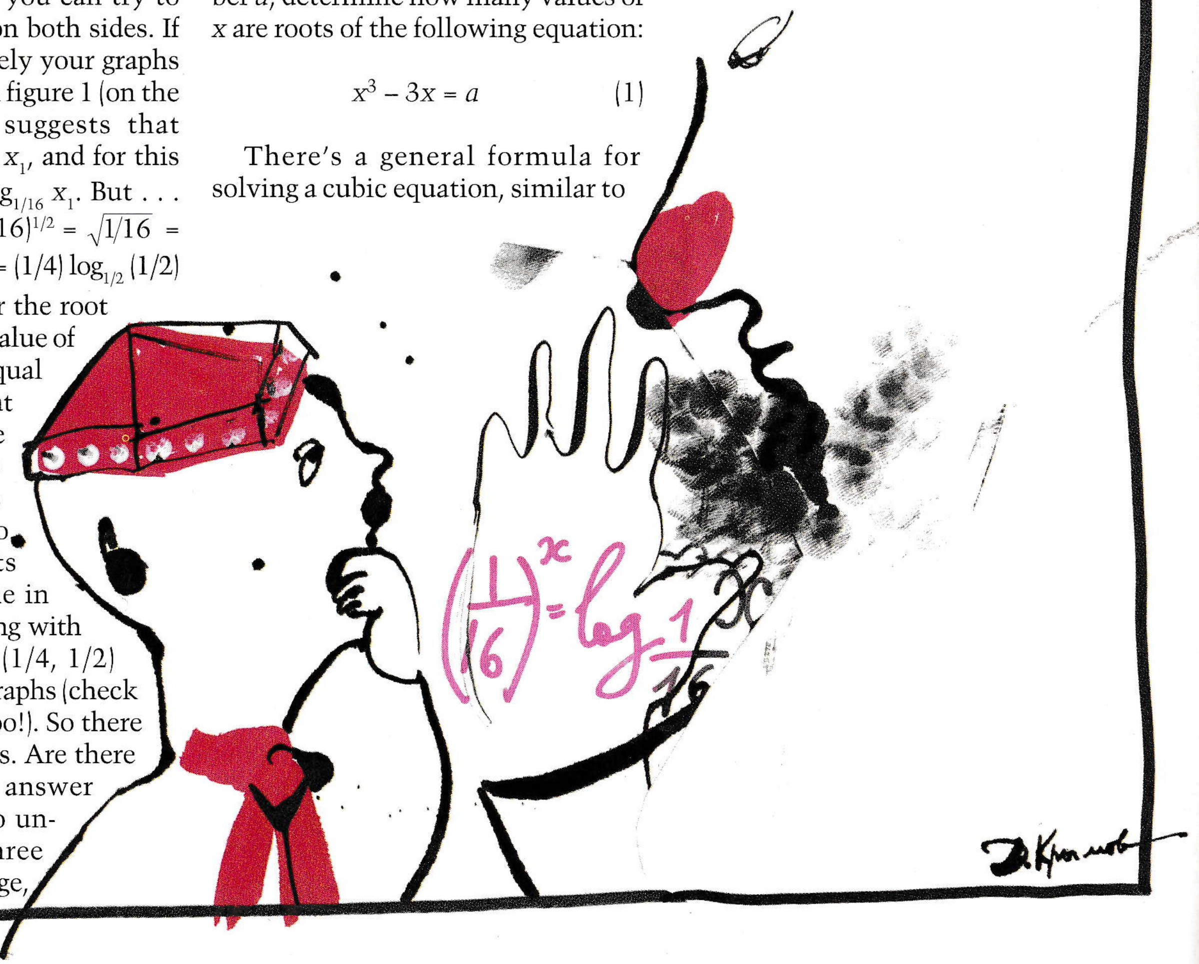
$$(1/16)^x = \log_{1/16} x$$

have?" This equation can't be solved in explicit form, but you can try to graph the functions on both sides. If you do this, most likely your graphs will look like those in figure 1 (on the facing page). This suggests that there's only one root x_1, and for this root $(1/16)^{x_1} = x_1 = \log_{1/16} x_1$. But . . . just take $x = 1/2$: $(1/16)^{1/2} = \sqrt{1/16} = 1/4$, and $\log_{1/16}(1/2) = (1/4)\log_{1/2}(1/2) = 1/4$. In addition, for the root $x = 1/2$ the common value of our functions is *not* equal to x, which means that our equation has one more root: the two graphs are symmetric about the line $y = x$, so their common points not on this line come in symmetric pairs—along with (1/2, 1/4), the point (1/4, 1/2) also belongs to both graphs (check $x = 1/4$—it's a root, too!). So there are at least three roots. Are there any other roots? To answer this question, and to understand how the three roots happen to emerge, we have to examine our functions more thoroughly. We'll do this later on. Now let's look at some simpler algebraic problems whose solution involves calculus—in particular, differentiation.

Example 1. For a given real number a, determine how many values of x are roots of the following equation:

$$x^3 - 3x = a \qquad (1)$$

There's a general formula for solving a cubic equation, similar to a well-known quadratic formula, but much more cumbersome. However, we don't need the roots themselves, we need only to find their number. Can't we find it *without* solving the equation?



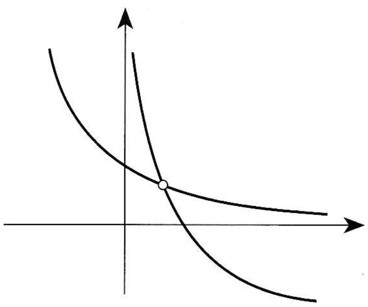


Figure 1

Let's sketch the graph of the function $f(x) = x^3 - 3x$. It's an odd function $(f(-x) = -f(x))$ with three zeros: $x = 0$, $x = -\sqrt{3}$, and $x = \sqrt{3}$; its derivative $f'(x) = 3(x^2 - 1)$ has two roots $x = \pm 1$, is positive for $x < -1$ and $x > 1$, and negative for $-1 < x < 1$. So the function increases on the interval $(-\infty, -1)$, attains its (local) maximum at $x = -1$, falls from $x = -1$ to $x = 1$, has its minimum at $x = 1$, and again rises on $(1, \infty)$; the values at extremal points are $f(-1) = 2$ and $f(1) = -2$. Finally, we get the graph shown in figure 2a. The number of the roots of our equation simply equals the number of intersections of the graph with the horizontal line $y = a$ (several lines are drawn in figure 2b). So we can "read" the answer right from the graph: equation (1) has only one root for $a < -2$ and $a > 2$ (or $|a| > 2$), three roots for $|a| < 2$, and two roots for $|a| = 2$ (the red lines in the figure).

A more rigorous proof of this result is based on a fundamental property of continuous functions, the Intermediate Value Theorem, which says that whenever a continuous function takes a value greater than a and a value less than a at some points x_1 and x_2, it necessarily takes the value a at a point x between x_1 and x_2; and on the obvious observation that a monotonic function takes any of its values only once. In particular, our function $f(x)$ has three intervals of monotonicity: $(-\infty, -1]$, $[-1,1]$, and $[1, \infty)$; so our equation can have not more than three roots—at most one root in each of the intervals. On the other hand, it does have a root in the first interval for any $a \leq 2$ (because function f takes values both less and greater than any such a on this interval), and it has a root in the second interval for any $a \in [-2, 2]$ (because $f(-1) = 2$, $f(1) = -2$), and a root for $a \geq 2$ in the third interval. Combining this statements, we get the answer. Note the special role of the red lines in figure 2b that touch the graph at the extremal points—they mark the changes in the number of roots.

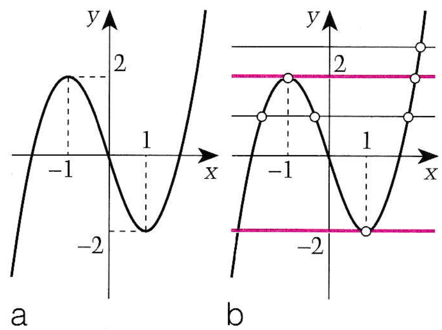


Figure 2

Similar arguments apply to the exercises below. As a rule, they are easy to reproduce, so I'll leave them to the reader.

Exercise 1. Find the number of the roots of the equations (a) $3x^5 - 50x^3 + 135x = a$; (b) $x^2e^x = a$.

Example 2. How many roots does the equation

$$a^x = x$$

have?

Sketching the graphs $y = a^x$ and $y = x$ for a varying from zero to infinity, we get the five essentially different cases shown in figure 3. Now the answer is seen with "the naked eye." In fact, the only thing left to do is find the value $a = a_0$ corresponding to figure 3d—that is, to the case when the line $y = x$ is tangent to the curve.

Let x_0 be the x-coordinate of the point of contact. Since at this point both the values of the two functions $y = a_0^x$ and $y = x$ and their slopes coincide, we can write the following two equations:

$$\begin{cases} a_0^{x_0} = x_0, \\ x_0^{x_0} \ln a_0 = 1 \end{cases}$$

(because $(a^x)' = a^x \ln a$). Substituting x_0 for $a_0^{x_0}$ in the second equation yields $x_0 = 1/\ln a_0$; plugging this into the first equation and taking the logarithm, we get

$$\frac{\ln a_0}{\ln a_0} = \ln \frac{1}{\ln a_0} = -\ln \ln a_0,$$

or $\ln \ln a_0 = -1$. It follows that $\ln a_0 = 1/e$, or

$$a_0 = e^{1/e}.$$

So equation (2) has one root when

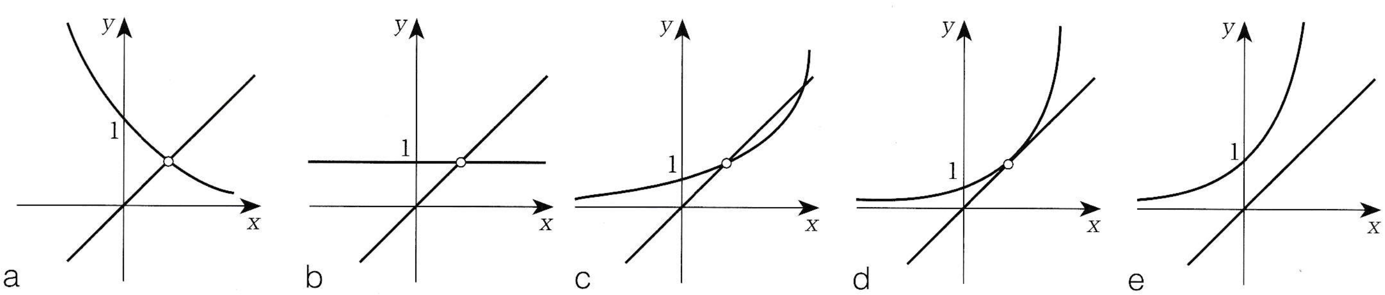


Figure 3

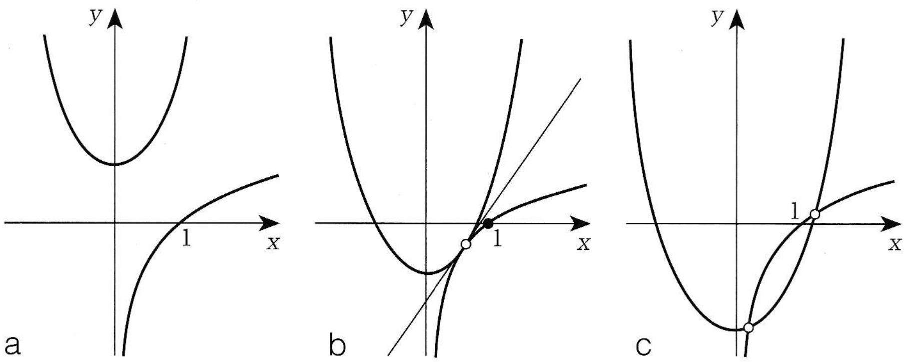


Figure 4

$0 < a \le 1$ or $a = e^{1/e}$, two roots when $1 < a < e^{1/e}$, and no roots when $a > e^{1/e}$.

Exercise 2. Find the number of roots of the equation $x/\ln x = a$.

Example 3. For what values of a does there exist a positive b such that the equation

$$x^2 + a = 2b \ln x$$

has a unique solution?

Again, let's begin with a drawing. In figure 4 you see three different cases of the relative positions of the graphs of $f(x) = x^2 + a$ and $g(x) = 2b \ln x$ (for $b > 0$). It's clear that the only case in which the graphs can have a unique common point is when they are tangent to each other (fig. 4b). Equating the values of the functions and their derivatives $f'(x) = 2x$ and $g'(x) = 2b/x$ at the point x of contact, we obtain

$$\begin{cases} x^2 + a = 2b \ln x, \\ 2x = 2b/x. \end{cases}$$

Since x must be positive (otherwise $\ln x$ is undefined), we have $x = \sqrt{b}$ from the second equation, and so

$$a = b \ln b - b. \qquad (3)$$

This condition is necessary and sufficient for the graphs to touch each other (at the point $x = \sqrt{b}$. So the problem is reduced to the following question: for what a is there a value of b such that $a = b \ln b - b$?

Question: why is the condition $b > 0$ omitted?

Let's plot the graph of function (3) in the (b, a)-plane (see figure 5). It shows that the answer has the form $a \ge a_0$, where a_0 is the minimum value $\phi(b_0)$ of the function $\phi(b) = b \ln b - b$, which can be found from the equation

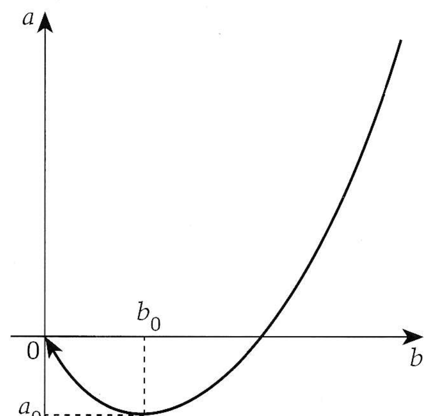


Figure 5

$\phi'(b_0) = 0$. Since $\phi'(b) = \ln b$, $b_0 = 1$ and $a_0 = \phi(b_0) = -1$.

Question: equation (3) has *two* roots for $-1 < a < 0$. What does this mean for the graphs in figure 4? Make a drawing.

So far we've only been counting the roots of equations. But sometimes knowing the number of roots of an equation helps us solve it. For instance, when you know that an equation has only one root, you can simply try to guess its value.

Example 4. Solve $\cos x = 1 - x^2/2$.

One root of this equation is quite easy to guess: $x = 0$. Are there any other roots? Look at figure 6. The graphs of functions $y = \cos x$ and $y = 1 - x^2/2$ are so close to each other near the point $x = 0$ that it's impossible to tell without a special examination which of the two figures—6a or 6b—is correct. Let's try to *prove* that $\cos x > 1 - x^2/2$ for all $x \ne 0$ (that is, figure 6a is the correct one, and the root $x = 0$ is unique).

Consider the function $f(x) = \cos x - 1 + x^2/2$. This is an even function ($f(-x) = f(x)$), so we can confine ourselves to only positive values of x. Since $f(0) = 0$, it suffices to show that $f(x)$ increases on the interval $[0, \infty)$ or that the derivative $f'(x)$ is positive for $x > 0$. But $f'(x) = -\sin x + x$, so $f'(x) > 0$ follows from the well-known inequality $\sin x < x$ (for $x > 0$).

The next example has to do with a generalization of the arithmetic–geometric mean inequality $(x + y)/2 \ge \sqrt{xy}$ for $x > 0$, $y > 0$. Rewriting it in the form

$$x^{\frac{1}{2}} \cdot y^{\frac{1}{2}} \le \tfrac{1}{2}x + \tfrac{1}{2}y$$

suggests the more general inequality in example 5.

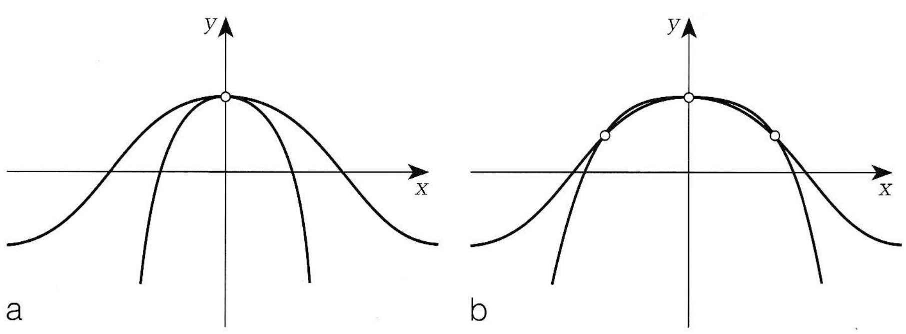


Figure 6

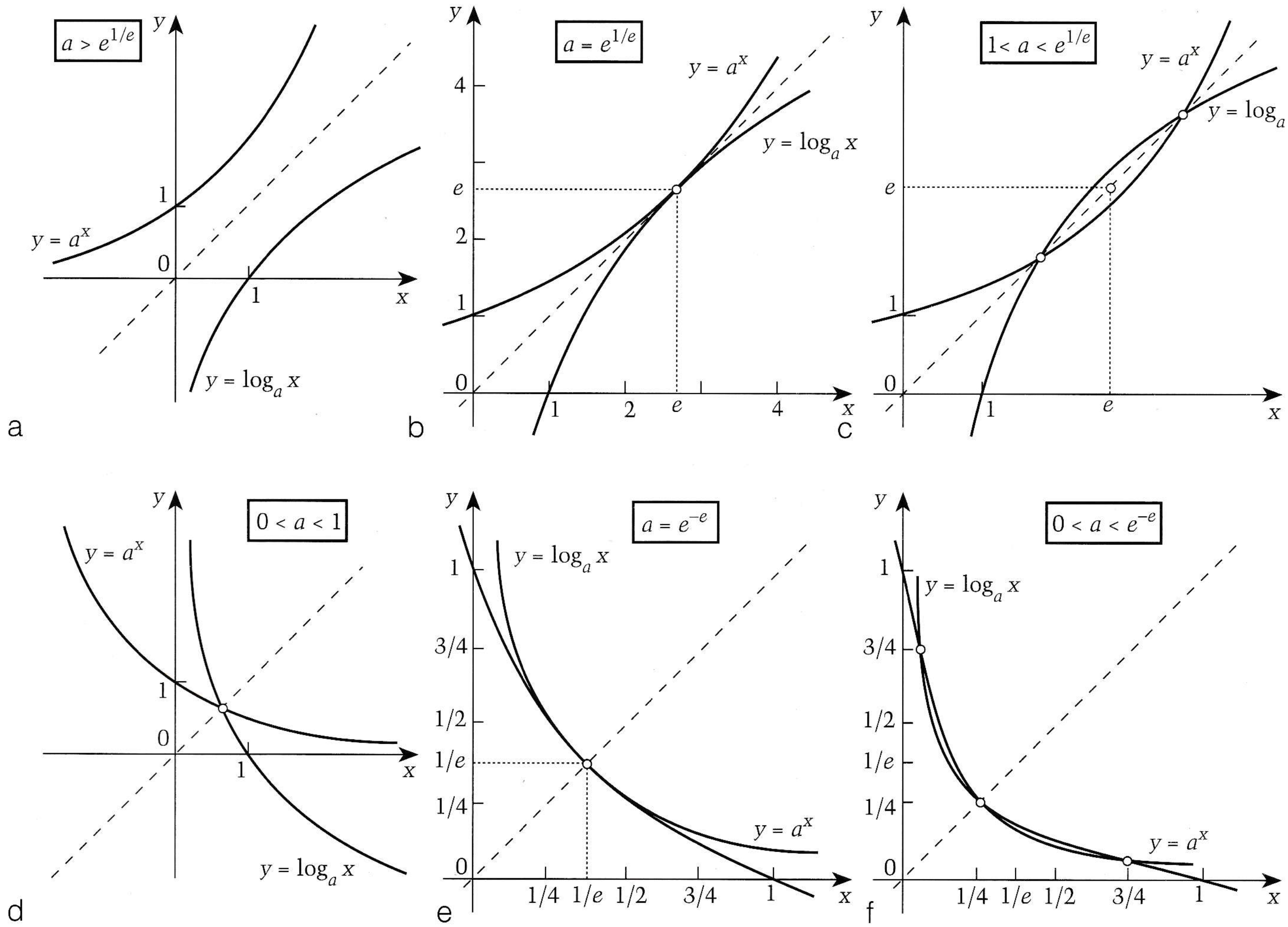


Figure 7

Example 5. For any positive a and b such that $a + b = 1$ and any positive x and y,

$$x^a y^b \le ax + by,$$

with equality holding if and only if $x = y$.[1]

First, we rework the inequality, replacing b with $1 - a$, carrying all the terms over onto one side, and dividing the inequality by $y > 0$:

$$a \cdot \frac{x}{y} - \left(\frac{x}{y}\right)^a + 1 - a \ge 0.$$

Denoting $x/y = t$, we arrive at the inequality

$$f(t) = at - t^a + 1 - a \ge 0,$$

which is to be proven for all $t > 0$ and $0 < a < 1$. I leave this proof as an exercise for the reader. (Hint: using derivatives, show that the minimal value of $f(t)$ is zero and is attained only once—at the point $t = 1$; this accounts for the case of exact equality, too.)

Exercise 3. Solve the following equations: (a) $\ln x = x - 1$; (b) $\sin x = x - (1/16)x^3$.

Now we can return to the problem posed at the beginning of this article. We'll consider an even more general question.

Example 6. How many roots does the equation

$$a^x = \log_a x$$

have?

Figure 7 presents all six possible cases of the relative positions of the graphs $y = a^x$ and $y = \log_a x$ that occur as parameter a sweeps from infinity to zero. Note that the graphs are symmetric to each other about the line $y = x$, because the functions on both sides of the given equation are mutually inverse. We see that the critical values of a, at which the number of roots changes, are (1) $a = a_0$—when the two graphs touch each other and the line $y = x$ (fig. 7b); (2) $a = 1$; and (3) $a = a_1$—when the graphs are tangent again, but cross the line $y = x$ at right angles (fig. 7e). In fact, we've

CONTINUED ON PAGE 43

[1]A particular case of this inequality was offered as math challenge M68 in the November/December 1992 issue of *Quantum*.—*Ed.*

The sines and cosines yo

And the "bows" used

THE SINE AND COSINE ARE THE BASIC ELEments of trigonometry, the science of measuring the parts of triangles. These functions are used by construction engineers, geodesists, and others.

By definition, the sine of an angle α is the y-coordinate of the point M on the unit circle centered at the origin such that the angle between the positive x-axis and the ray OM is α. The cosine of α is the x-coordinate of this point. The relation between these two functions is given by

$$\cos^2 \alpha + \sin^2 \alpha = 1$$

and

$$\cos \alpha = \sin\left(\frac{\pi}{2} - \alpha\right).$$

It's interesting that the sine was introduced not by the ancient Greeks (although they made the major contributions to the study of the geometry of the triangle) but by the Indians, whose mathematical interests were closer to practice. The term "sine" itself owes its origin to a grammatical misunderstanding. In their calculations the Indians made extensive used of half the length of the chord subtending a given arc (in figure 1, $MA = \sin \alpha = \frac{1}{2}MN$) rather than the whole chord. They called it *ardhajiva*—"half of a bowstring."[1] Later the word *ardha* ("half") was dropped, and *jiva* became the name of the "sine line" (MA in figure 1). The Arabs, who passed Greek knowledge along to us, also deliv-

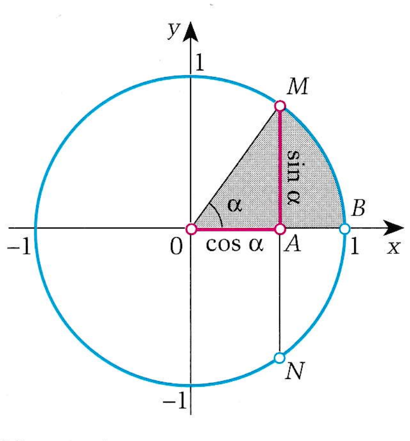


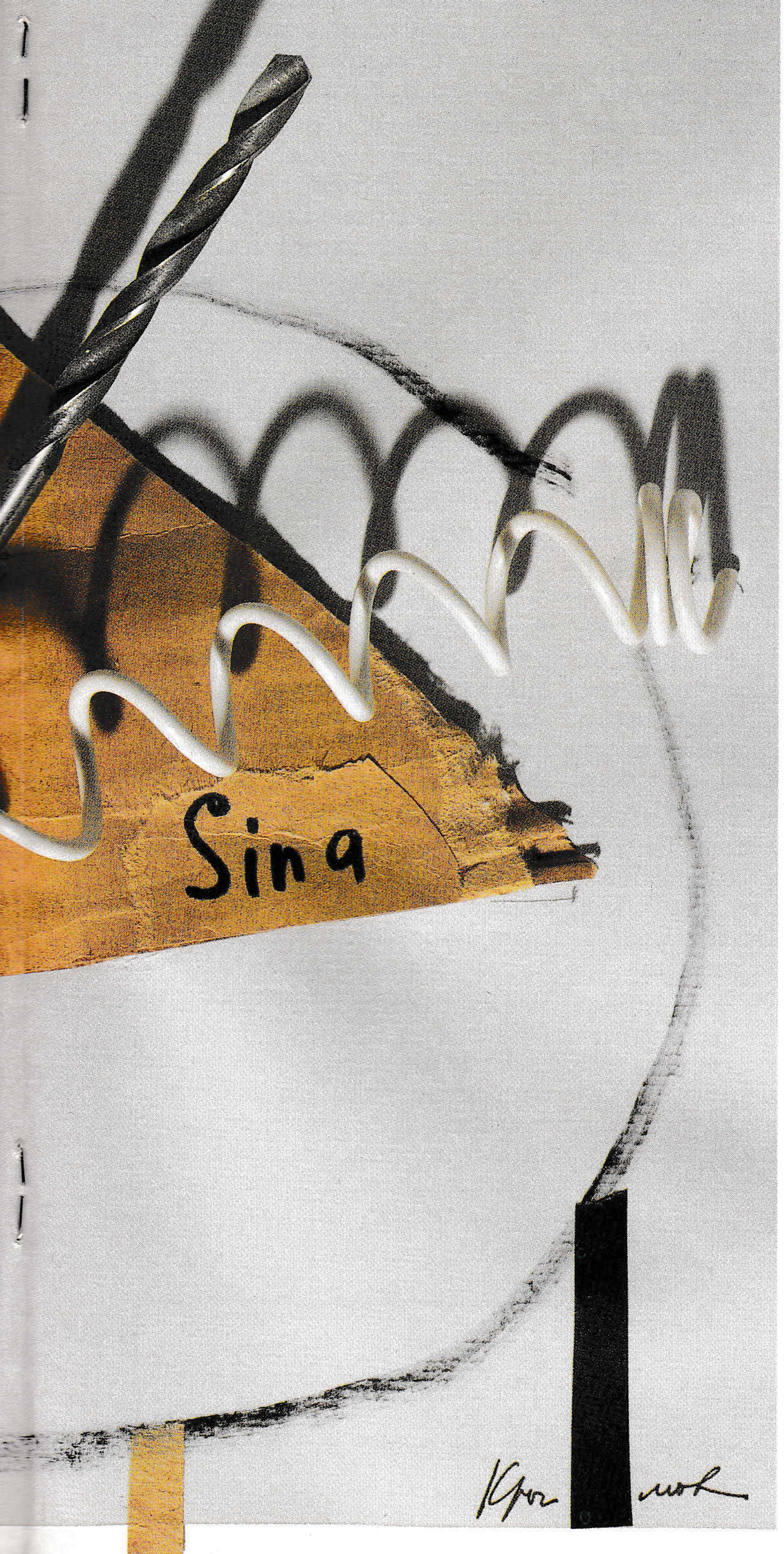
Figure 1

[1]Similarly, our "chord" and "arc" come from the Greek χορδη (string) and the Latin *arcus* (bow).

Art by Dmitry Krymov

you do and don't know

sed to define them



ered the science and culture of India to Europe. For instance, the "Arabic numerals" that we use were borrowed from the Indians. And the notion of sine also reached us through the Arabs. They transliterated the word *jiva* as *jiba*, which is written in Arabic in the same way as *jaib* (in Arabic script vowels are denoted by specials signs above or under the line and are often simply omitted). The word *jaib* means "cavity," and when in the 12th century Arabic treatises were translated into Latin, this word was rendered as *sinus*, the Latin word with the same meaning.

The Indians used the cosine, too. Their *kotijiva*— the sine of the remainder (after subtracting from 90°)—eventually turned into the Latin *sinus coplementi*—the sine of the complement, "cosine" for short. Another trigonometric function introduced by the Indians was *utkramajiva*, the difference between the radius and the "cosine line"; in Europe it was named *sinus versus*, the reversed sine. In modern notation it's defined by the formula $\operatorname{sinv} \alpha = 1 - \cos \alpha$. In figure 1, the cosine of the angle α is equal to OA, and $\operatorname{sinv} \alpha$ equals AB, the height of the circular segment MBN. It's interesting that in Russia the height of the segment used to be called the "arrow," which takes us back to the Indian bow with the bowstring MN.

Let's skip over most of the many trigonometric formulas and turn to the graph of the function $y = \sin x$. It's called the sine curve, or sinusoid, and seems very artificial, though its undulations resemble waves on water. Indeed, fluid waves, as well as radio, light, and sound waves, are directly linked to the sine function. To make a template for drawing the sine curve, wind a sheet of paper several times around a candle and cut it with a sharp knife at an angle of 45° to the candle's axis (its wick). After unrolling the paper (fig. 2), you'll get two

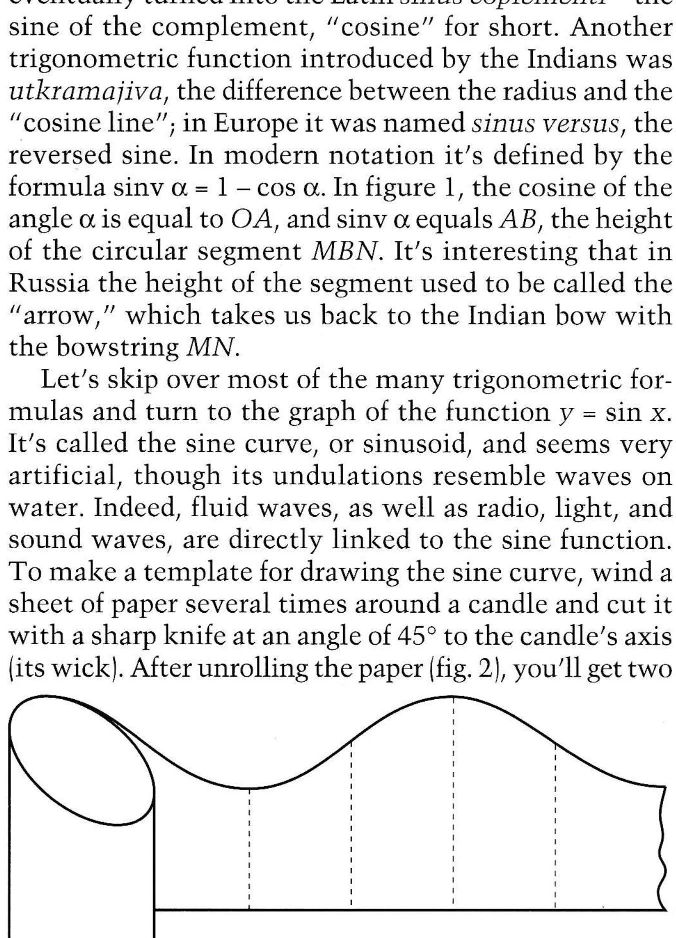


Figure 2

wonderful templates of a sinusoid with the radius of the candle taken as the unit. The graphs of all functions of the form $y = a \sin(kx + b) + c$ are also called sinusoids. They can be obtained from the standard sine curve by shrinking or stretching along the axes and by translation. So the graph of $y = \cos x = \sin(\pi/2 - x)$ is a sine curve, as is the graph of $y = \sin^2 x = \frac{1}{2}\sin(\pi/2 - 2x)$. You can also see a sine curve when you look at a spring or drill from the side.

Now let's consider the hyperbola $y = \frac{1}{2}x$ and turn it 45° clockwise about the origin (fig. 3). The equation of the curve thus obtained is $x^2 - y^2 = 1$ (why?). It will intersect the x-axis at points $B(1, 0)$ and $B'(-1, 0)$; from here

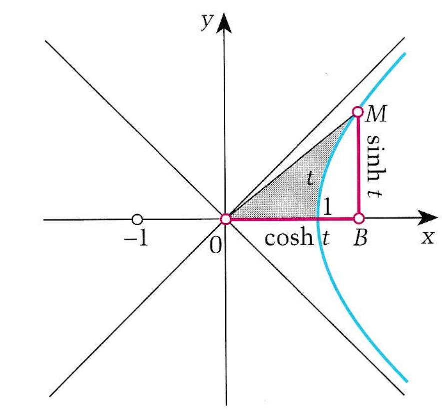


Figure 3

on we'll be considering only its right half. Turn back to figure 1 for a moment and note that the value α there can be interpreted as twice the area of the circular sector OBM.

Now take a point M on the (rotated) hyperbola and define the parameter t as twice the area of the hyperbolic sector OBM taken with a plus sign if M is in the upper half-plane and a minus sign if M is below the x-axis. Then t takes all real values from $-\infty$ to ∞, and each value of t corresponds to one and only one location of point M on the hyperbola. For every t the coordinates of the corresponding point M are called the hyperbolic cosine and sine of t; they are denoted by $\cosh t$ and $\sinh t$, so $M = (\cosh t, \sinh t)$.

Obviously, $\cosh^2 t - \sinh^2 t = 1$. The area of a hyperbolic sector can be computed by means of integration. This yields the following expressions for the hyperbolic functions in terms of the exponential function e^x:

$$\cosh t = \frac{e^t + e^{-t}}{2}, \quad \sinh t = \frac{e^t - e^{-t}}{2}.$$

By using these formulas, you can derive the addition formulas for $\cosh(t \pm s)$ and $\sinh(t \pm s)$. You'll find that they are almost the same as the addition formulas for cos and sin. (The only difference is that the signs in the trigonometric and hyperbolic formulas for the cosines are opposite.) In fact, there's a very close connection between the trigonometric and hyperbolic functions.

This becomes clear if we pass from the real to the complex numbers. Recently we've met with another hyperbolic function—the hyperbolic tangent $\tanh t = \sinh t/\cosh t$—in the context of hyperbolic geometry and relativity (see "In the Curved Space of Relativistic Velocities" in the March/April 1993 issue of *Quantum*.)

The graphs of $y = \sinh t$ and $y = \cosh t$ are shown in figure 4. We see that one of the functions is even and the other is odd. The graph of $\cosh t$ is also called the catenary or "chain line" from the Latin *catena* (chain), because it's the shape taken by a chain suspended at its ends.

Besides trigonometric and hyperbolic sines and cosines, there are other kinds as well—for instance, *lemniscatic*, which are defined via the lemniscate of Bernoulli. This curve (fig. 5) is the locus of points in

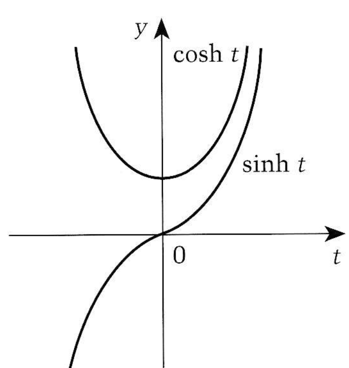


Figure 4

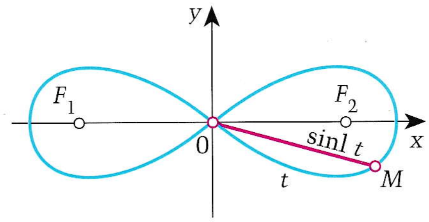


Figure 5

the plane such that the product of their distances to points F_1 and F_2 is constant and equals a quarter of the square of the distance between them. It was discovered by Jakob Bernoulli almost exactly 400 years ago, in 1694. He described it as "shaped like a figure 8, or a knot, or a ribbon bow," and used the Latin word *lemniscus* (a ribbon fastened to a victor's garland) as its name.

If $F_1F_2 = \sqrt{2}$, the Cartesian equation of the lemniscate is

$$(x^2 + y^2)^2 = 2(x^2 - y^2).$$

The equation in polar coordinates (r, ϕ) is simpler:

$$r^2 = 2 \cos 2\phi.$$

The argument of lemniscatic functions, as it was in our first definition of sine and cosine, is the arc length measured counterclockwise from the origin O in the right half of the lemniscate and clockwise in the left half. The lemniscatic sine $\operatorname{sinl} t$ is defined as the length of OM if M is on the right half of the curve and $-OM$ if it's on the left half. The lemniscatic cosine is defined by the formula $\operatorname{cosl} t = \operatorname{sinl}(\omega k - t)$, where ω is the arc length of one half of the curve.

These functions also have much in common with the trigonometric functions. Their graphs differ very little from the sine curve, and the functions themselves have proved extremely useful in modern mathematics. ◘

:-) ↔)-:

From a snowy Swiss summit to the apex of geometry

A short biography of Jacob Steiner

by I.M. Yaglom

THE REMARKABLE SWISS pedagogue and humanist Johann Heinrich Pestalozzi (1746–1827) is widely considered the first theorist of primary education and the originator of the idea of combining it with productive labor. He tested his ideas and brought them to life in an excellent boarding school for poor children that he founded. His colleagues traveled all over Switzerland, selecting pupils and persuading parents to send their children to the school (this was often the hardest part of their job).

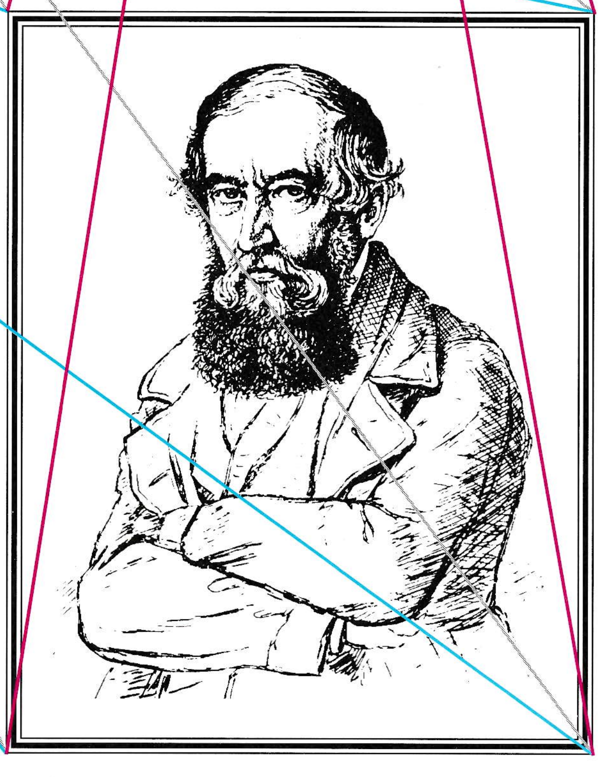
In 1814, in the mountains of Switzerland, one of Pestalozzi's colleagues met a young shepherd named Jacob Steiner, the son of a poor peasant. At that time Jacob could barely read or write, but he had taught himself some mathematics and astronomy, which especially interested him at the time. The knowledge and interests of the young peasant astounded Pestalozzi's colleague, who began urging the elder Steiner to forego the service of his valuable assistant and send him to school. It wasn't easy, but in the end the 18-year-old Steiner (born in 1796) left his native village—forever. He went to Iverden, a town near Bern, and entered Pestalozzi's boarding school free of charge. Steiner had no money for education, for food, or for lodging.

Steiner spent four years at Pestalozzi's school in Iverden. First he went to classes, then he taught mathematics there. But Pestalozzi soon realized that Steiner's talent deserved something better.

In 1818, following Pestalozzi's insistent advice, Steiner left for Heidelberg, Germany, the nearest major university center.

Pestalozzi planned that Steiner would graduate from Heidelberg University and then find his own way. But things didn't work out so smoothly. Pestalozzi gave Steiner some money to take the trip and settle down in a strange city. Still, Steiner had to earn his daily bread himself. Since he was only qualified to be a mathematics teacher, he was forced to overburden himself with private students, who paid little—after all, Steiner had no formal education. These private lessons, which Steiner detested, prevented him from regularly attending classes at the university. During the three years he spent in Heidelberg, Steiner managed to take only a few university courses (which was not insignificant, though, considering his level of preparation). Nonetheless, in Heidelberg Steiner made his first mathematical acquaintances: the outstanding talents of this Swiss were obvious to everyone he met.

In 1821 Steiner learned that a teaching position had opened up at a gymnasium (secondary school) in Berlin. He set off immediately. The position would give him a regular income. But to fill the vacancy he needed to pass examinations, and



the results were unlikely to fill the gymnasium administrators with enthusiasm.

First off, the candidate for the teaching position was asked if he was familiar with the gymnasium curriculum. "No, I'm not," Steiner replied curtly. What else could he say? In Prussian gymnasiums Latin and Greek were part of the curriculum. This son of a Swiss peasant knew these languages no better than you or I, dear reader. Steiner proved to be rather indifferent in mathematics as well: while he displayed a wide-ranging and profound understanding of geometry, his grip on algebra and trigonometry was rather feeble. Wide gaps were also discovered in the field of mathematical analysis. However, the young man's striking abilities in geometry and the flattering testimonials he brought with him did the trick: Steiner was allowed to teach mathematics for two years in all grades except the last. During this time he was required to pass all the exams in the gymnasium curriculum and an extra exam in mathematics. He successfully took this extra exam much later. As for the others, Steiner never did pass them.

Later, Steiner taught in a secondary school for 14 years. This became possible only after a vocational school opened in Berlin. There the curriculum in mathematics and the natural sciences was broadened; the ancient languages were no longer part of the curriculum (so the teachers didn't have to know them). But even though the school's organizer and director was one of Pestalozzi's students, at first Steiner was accepted only as a teaching assistant. After passing additional exams in 1829 he became a full teacher. Alas! We have to admit that the irritable and abstracted Steiner wasn't a good teacher. He worked enthusiastically with gifted students, thinking up brilliant individual challenges for them (in geometry most of all—see problems 1, 2, and 6 and the appendix). The rest of the students annoyed him: Steiner simply couldn't understand their lack of ability and interest in mathematics. From time to time, when he couldn't take it any more, he'd quit his regular job and earn a living by private lessons again. The same happened also during the fortunately short period when Steiner was barred from teaching at the gymnasium because he failed to pass an exam (the vocational school had not yet opened). However, he would invariably come back to his old school, where people were used to his idiosyncrasies and where his mathematical talents were highly valued.

Choosing problems for his students, Steiner acquired an interest in elementary geometry that never faded for the rest of his life. Let's turn to some of Steiner's results in this field. We'll begin with problems dating back to the great Euler.

Problem break

Euler established that *the three midpoints of the sides of an arbitrary triangle, three bases of its heights, and three midpoints of the segments of heights from their intersection point* (the orthocenter of the triangle) *to the vertices lie on one circle*. This circle is called an *Euler circle* or *the 9-point circle of a triangle*. It's remarkable that in any triangle the Euler circle touches the inscribed and three escribed[1] circles (fig. 1). This is often called *Feuerbach's theorem*, after the person who was the first to prove it. Few people know that Steiner, ignorant of Feuerbach's result, proved this theorem just two years later and published his result immediately, so that many mathematicians learned of it from Steiner's version rather than Feuerbach's.

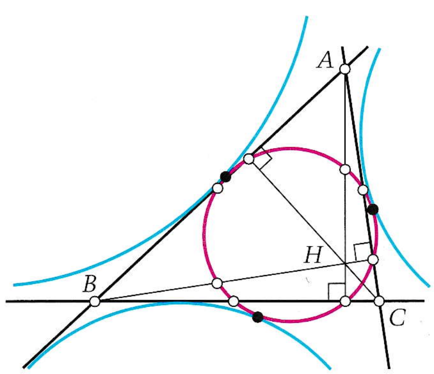


Figure 1

[1]That is, the circles that touch one side of a triangle and the extensions of the other two sides.—*Ed.*

1. Let a circle S centered at in the orthocenter H of a triangle ABC meet its midlines B_1A_1 ($\parallel AB$), C_1B_1 ($\parallel BC$), and A_1C_1 ($\parallel CA$) at points F_1 and F_2, D_1 and D_2, E_1 and E_2, respectively. Prove that $AD_1 = AD_2 = BE_1 = BE_2 = CF_1 = CF_2$ (Steiner's theorem).

Two triangles ABC and $A_1B_1C_1$ *are perspective from perspectivity center P* if lines AA_1, BB_1, and CC_1 meet at P. Two triangles ABC and $A_1B_1C_1$ are *directly similar* if they are similar and have the same orientation: tracing their perimeters in the orders $A \to B \to C$ and $A_1 \to B_1 \to C_1$, respectively, we move in the same direction—clockwise or counterclockwise. It's not hard to show that two such triangles can always be brought into coincidence by means of a *spiral similarity* (also called a *rotational dilation*)—that is, a dilation relative to some center Q combined with a rotation about Q. Point Q is called the *center of similarity* of triangles ABC and $A_1B_1C_1$.

2. Let a, b, c be lines forming a triangle T; let line l cut a, b, c at points A_0, B_0, C_0. Raise perpendiculars to the sides of triangle T at these points: $a_1 \perp a$, $b_1 \perp b$, $c_1 \perp c$. Let t denote the triangle formed by lines a_1, b_1, c_1. (a) Prove that triangles T and t are directly similar and perspective; that their circumcircles S and s intersect at right angles (that is, their tangents at either of their common points P and Q are perpendicular to each other); and that one of the points P and Q is the similarity center of the triangles and the other is their perspectivity center (all these assertions are theorems of Steiner's). (b) In what way will these theorems change if we replace the *perpendiculars* a_1, b_1, c_1 with three lines through A_0, B_0, C_0 that make the same (in absolute value and direction) angle α with a, b, c, respectively ($0 \le \alpha \le 90°$)?

The next series of problems, concerning a complete quadrilateral,

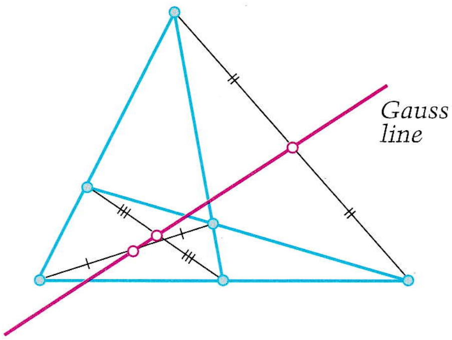


Figure 2

stems from the great Gauss. Steiner certainly knew how to choose his predecessors!

A *complete quadrilateral* Q is a figure formed by four lines in the "general position"; the four triangles formed by all triples of these lines are called the *triangles* of the quadrilateral Q. The intersection points of the lines are the *vertices* of the quadrilateral Q; the segments that join "nonadjacent" vertices (that is, those not lying on one of the given lines) are called the *diagonals* of the quadrilateral Q.

3. (Gauss's theorem) The midpoints of the three diagonals of a complete quadrilateral lie on one line (fig. 2). This line is called the *Gauss line of the quadrilateral.*

4. The circumcircles of the four triangles of a complete quadrilateral intersect at one point (fig. 3). This point C is called the *Clifford point of the quadrilateral.*

The statement of problem 4 was known before William Kingdom Clifford (1845–1879). But Clifford invented a remarkable construction (the *Clifford chain*) in which this problem is included. A *complete n-lateral N* is defined as any set of n lines in general position; it contains n complete $(n-1)$-laterals $M_1, M_2, \ldots, M_n$, each obtained by removing one of the n lines. The *Clifford point of a "complete bilateral" (a, b)* is simply the common point of a and b, and the *Clifford circle of a "complete trilateral" (a, b, c)* is the circle passing through three Clifford points of bilaterals (a, b), (b, c), (c, a)—that is, the circumcircle of the triangle with sides a, b, c. Then, for any *even n* the Clifford circles of $(n-1)$-laterals $M_1, M_2, \ldots, M_n$ meet at one point, called the *Clifford point of the complete n-lateral N* (for $n = 4$ it's the statement of problem 4). If n is *odd*, then n Clifford points of $(n-1)$-laterals $M_1, \ldots, M_n$ lie on one circle—the *Clifford circle of N.*

Auxiliary problems

5. (a) The bases of perpendiculars dropped on the sides of a triangle T from a point M of the circumcircle lie on one line (fig. 4). This line w is called the *Simpson–Wallis line* of point M relative to triangle T.

(b) Line w bisects segment MH (where H is the orthocenter of triangle T).

6. (Steiner's theorems) (a) The orthocenters of four triangles of a complete quadrilateral Q lie on one line (fig. 5). This line s is called the *Steiner line of Q*. (b) In any complete quadrilateral Q its Steiner line s is perpendicular to its Gauss line.

Biography continued

Let's get back to Steiner's life story. The greatest success of his Berlin period was his acquaintance with an amateur mathematician, rich manufacturer, and talented engineer and railway magnate by the name of August Leopold Crelle (1780–1835). Although he wasn't an outstanding scientist, he was a member of the Berlin (Prussian) Academy of Science and a corresponding member of the Petersburg (Russian) Academy of Science. He was given these honors not for his scientific activity but for his engineering achievements and organizational talents. But a successful entrepreneur must have a good understanding of people, and Crelle showed that he knew them well.

The first specialized mathematics journal in Europe was founded in 1810 by the well-known French mathematician Joseph Diez Gergonne (1771–1859) and was titled *Gergonne's Annals*. Crelle decided to found a German mathematics journal. In lining up authors for the journal, Crelle counted mostly on two persons absolutely unknown to professional mathematicians but in whose talents he believed strongly. They were a semi-educated Norwegian student, N. H. Abel, and a secondary school teacher, Jacob Steiner. The first issue of *The Jour-*

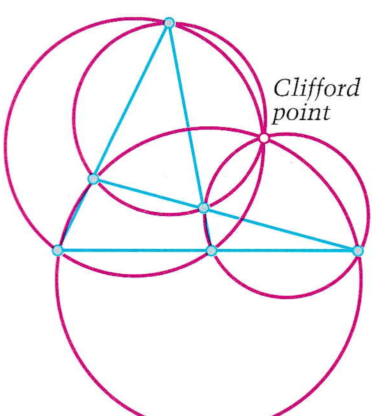


Figure 3

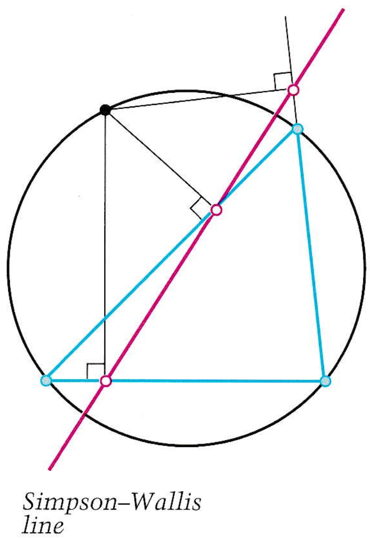


Figure 4

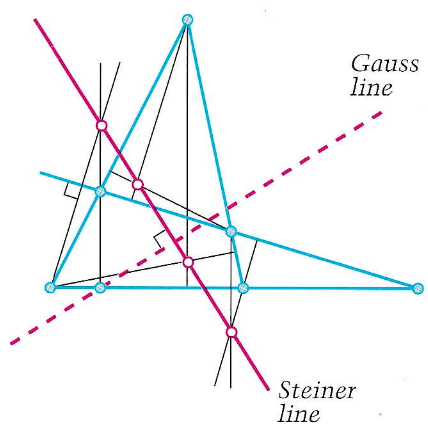


Figure 5

nal of Pure and Applied Mathematics appeared in 1826. The bulk of the issue (and of those to follow) consisted of articles by Abel and Steiner. In fact, the first three issues contained 15 articles and shorter items by Steiner! The "Crelle journal" (as mathematicians dubbed it) became "the leading mathematical journal in the world." (Gergonne stopped publishing his in 1831. It was revived, however, by another French mathematician, Joseph Liouville (1809–1882), with a title that was a direct imitation of the Crelle journal.)

The Crelle journal became Steiner's rostrum for his geometrical ideas. In addition, the influential Crelle was the force behind Steiner's election to the Berlin Academy of Sciences (1834): Steiner's outstanding scientific writings published in the Crelle journal provided a strong justification. After that it wasn't hard to secure Steiner's election as a professor. In 1835 Steiner left his secondary school and took a position in the department of natural sciences at Berlin University.

It's interesting that the indifferent school teacher Steiner turned into the outstanding university professor Steiner. In the secondary school he was irritated by students who were strangers to mathematics. But students who were enthusiastic about geometry inspired Steiner; his lectures, striking in their form and zestful in their delivery, were great successes. Even the Swiss accent of the extraordinary professor was popular with the students. They were impressed by Steiner's habit of calling students to the blackboard to solve the problems he would continually pose during his lectures. (In our own time, this was the standard practice of academician I. M. Gelfand at Moscow State University, and it was enormously popular with his students as well.)

The success of Steiner's lectures actually had a partly negative influence on the history of geometry. For instance, as late as the second half of this century, lectures in projective geometry in many of the world's universities were delivered according to the badly outdated scheme worked out by a semiliterate Swiss shepherd. Also, Steiner's archaic terminology was still used in this area, even though it had been dropped in every other field of mathematics.

Steiner felt he was the leader of German geometry. He reacted with almost pathological displeasure to any deviation from his tenets. Steiner's most illustrious contemporary in geometry was professor Julius Plücker (1801–1868). He represented the *analytical* trend in geometry, which sought to replace geometric images with coordinate notation and work with these coordinate representations of geometric objects by means of elaborate algebraic techniques. This was enough to make Steiner—a pure geometer who didn't admit analytical methods in geometry—extremely antipathetic to Plücker. But Plücker might have had two other flaws in the eyes of the son of a poor peasant family who didn't have a university education. Plücker was the scion of a family of industrial magnates in the Rhine Valley and was extraordinarily wealthy. In addition, he had graduated from *two* universities—Bonn and Paris. No wonder Plücker's writings aroused Steiner's fury.

Plücker was often sloppy in his treatises: slips of the pen and other easily removable defects abounded, so there were grounds for Steiner's attacks. Steiner's furious (and mostly unjust) criticism wore Plücker down to the point that he temporarily gave up geometry, returning to it (with great success, I might add) only after Steiner's death. In the interim he turned to experimental physics and enjoyed great success in that field. It's possible that, were it not for Steiner's attacks, the development of physical structural analysis and the discovery of cathode rays—which was made by Plücker's student, Johann Wilhelm Gittorg (1824–1914)—might have been postponed for many years.

So even the shortcomings of great scientists can sometimes work to the benefit of science!

Overexertion and malnutrition in his youth made Steiner very weak and sickly in the last two years of his life. To undergo a treatment he would often go to his native Switzerland. But in 1863 he did not come back from one of these trips. He died on April 1, 1863, in a hotel room, absolutely alone. The long period during which he lacked material means prevented him from settling down to married life. He bequeathed a sum of money to the Berlin Academy of Science—for the establishment of a prize for geometry writings (meaning pure geometry, of course). A certain sum was left to the administration of his native canton in Switzerland—to stimulate the best mathematics pupils of the primary school for poor children.

Appendix

Constructions with straightedge alone

The postulates of Euclid's *Elements* assert the possibility of indefinite extension of a given line segment, of drawing a line through two given points, and of drawing a circle with a given center (point) and radius (segment). Along with the postulates implied but not formulated by Euclid—concerning the possibility of finding the intersection of two given lines (for example, given by a pair of points of each), of two given circles (defined by their centers and radii), and a given line and circle—these postulates describe the entire range of construction problems "solvable according to Euclid." These problems boil down to the aforementioned postulates—that is, they are to be solved with straightedge and compass.

In 1797 Lorenzo Mascheroni's *Compass Geometry* was published in Italy. The book claimed that *all problems in construction solvable with a straightedge and compass are solvable with compass alone*—with one natural restriction: a line segment cannot be constructed with compass. However, using a compass, one can find any number of points of a segment if two of its points are known. Much later a book by George Mohr (1640–1697) of Denmark was discovered. It was published in 1672 (125 years before Mascheroni's treatise) in two languages (Danish and Dutch) and proved the same theorem.[2]

Jacob Steiner got interested in constructions with *straightedge* alone (see problems 1–3 below). He successively examined constructions with straightedge alone that can be performed if the following figures are drawn in the plane: (a) two parallel lines or a segment divided by a given point in a given rational ratio; (b) a parallelogram; (c) a square; (d) a circle with its center. He showed that in case (d) all the constructions "performable according to Euclid" can be carried out with straightedge alone. (Of course, we can't construct a circle with a given center and radius using only a straightedge. But we can find any number of points on this circle.)

Problems

These constructions should be done with straightedge only.

1°. Parallel lines AB and l are given. Construct (a) the midpoint C of the segment AB; (b) the point D on this segment such that $AD = AB/n$, where n is a given integer.

2°. Given are points A, B and (a) the midpoint C of AB, (b) the point D on AB such that $AD = AB/n$. Draw a line l through a given point M parallel to AB.

3°. Given a point M, a line l, and (a) a parallelogram, (b) a square, draw through M a line (a) parallel, (b) perpendicular to l.

4°. Given a circle S with center O and (a) five points A, B, Q, K, L, (b) six points Q, K, L, P, M, N, construct the points of intersection of (a) the line AB and the circle with center Q and radius KL; (b) the circles with centers Q and P and radii KL and MN, respectively.

5°. (Hilbert's problem) Prove that it's impossible to construct the center of a given circle with ruler alone (without compass!).

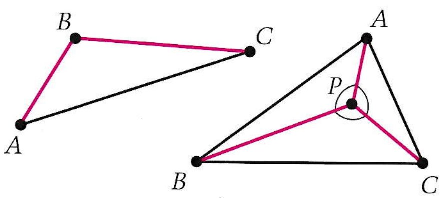
[2]See also "Constructions with Compass Alone" in the May 1990 issue of *Quantum*.—*Ed.*

The shortest network

Steiner used to give his students problems of finding the best configuration or a figure from this or that viewpoint. Here's one of his favorites.

Three villages are given. Connect them in a network of roads of minimum length.

It's more or less clear (though it remains to be proved) that the solution is given either by two sides of $\triangle ABC$ (except the longest one) or by segments AP, BP, CP, where the sum of distances from P to the vertices of the triangle is as small as possible (fig. 6a and 6b). One can prove that the solution of Steiner's problem is given in figure 6b if every angle in $\triangle ABC$ is less than 120° and in figure 6a if $\angle B \geq 120°$.



a b

Figure 6

In the case when the number of villages $n > 3$, the minimum network may be similar to figure 6a—that is, consisting of roads connecting the villages. Such a network is called a *framework*. It's always possible to find the shortest network by an exhaustive method (nowadays computers are used to find the solution of Steiner's "general" problem with a larger number of "villages"). In most cases the best network is similar to the one given in figure 6b—that is, one with extra network "nodes" where three roads meet; the roads form an angle of 120° between one another. Such nodes are called *Steiner points*, and the networks containing them are called Steiner networks (fig. 7). Alas! We have no general methods of finding minimal Steiner networks connecting n places—we don't know when "the absolutely minimum" network is a

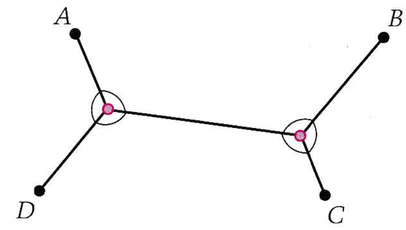


Figure 7

framework and when it's a Steiner network. It has been proposed that the minimum framework cannot be considerably longer than the minimum Steiner network. In the worst case it will be $\frac{2}{3}\sqrt{3}$ times longer (that is, 15% longer), but this hypothesis has been proven only for the case of $n \leq 5$.

As you can see, Steiner's "general" problem turned out to be not so simple. Steiner himself could come up with only a few examples of such networks for the case of $n > 3$. Today we know little more than he did![3]

Problems

6°. Prove the result formulated by Steiner (depicted in figure 6).

7°. Find the shortest network connecting four points A, B, C, D that are vertexes of (a) a square; (b) a triangular pyramid (tetrahedron).

[3]For more on the Shortest Network Problem, see the May/June 1993 issue of *Quantum*.—*Ed.*


BARKHIN

Late light from Mercury

What delayed the message from the fleet-footed god?

by Yakov Smorodinsky

ALMOST NOBODY WONDERS why it takes time for light to reach the Earth from a heavenly body. The light from the Sun travels for eight minutes before it reaches the Earth. It's easy to verify this number. The distance between the Sun and the Earth is 150 million kilometers—that is, $1.5 \cdot 10^{11}$ m. The speed of light is about $3 \cdot 10^8$ m/s. Dividing the first number by the second, we get 500 s ≅ 8 min.

However, the general theory of relativity makes some very important corrections to such reasoning. The phenomena explained by this theory are best demonstrated by Mercury. And that is the planet we'll look at.

The distance between the Earth and Mercury attains maximal or minimal values when the Sun and Mercury are in conjunction—that is, when the Earth, the Sun, and Mercury lie on the same straight line. These distances are $r_{max} = 1.38$ astronomical units (AU) at superior conjunction (fig. 1),[1] when the distance between the Earth and Mercury is maximal, and $r_{min} = 0.62$ AU at inferior conjunction (1 AU equals the average distance between the Sun and the Earth). Multiplying these numbers by 8 min/AU, we obtain the approximate time it takes light to travel from Mercury to the Earth from both positions. Of course, this calculation yields correct values if we're not interested in the details. But it's the details that we'll be examining in this article!

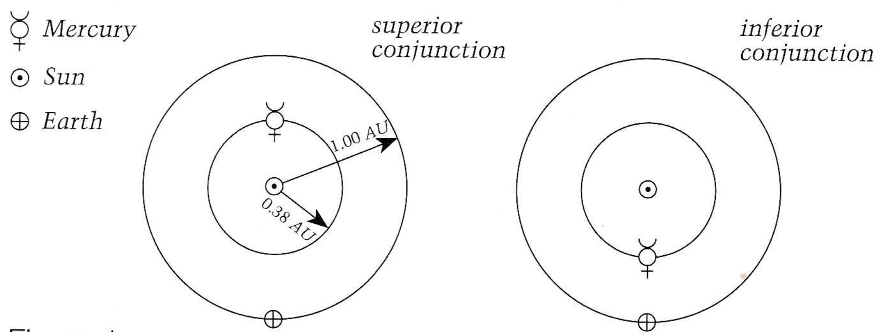


Figure 1

Let the light beam pass near the Sun when Mercury is at superior conjunction. The general theory of relativity leads to the conclusion that the speed of light is less in the Sun's gravitational field than in a vacuum (much like what happens when light propagates in transparent matter).[2] This decrease in the speed of light is very small, and calculations show that it corresponds to an increase in the light's travel time of 0.00024 s (a light beam travels 72 km during this time). Modern radar technology has made it possible to record such an exotic effect.

What are we to make of this number, 72 km? Clearly, it's hard to calculate this value. However, we can get an idea of its order of magnitude if we understand the concept of a gravitational radius and can use dimensional analysis.[3]

The quantitative characteristic of the gravitational field of a massive body is the gravitational potential energy per unit mass. According to Newton's law of universal gravitation, this gravitational potential is given by

$$\phi = -\frac{GM}{r}.$$

This formula contains two values: the product GM, which characterizes the source of the field (the Sun, in our case), and the distance r. Usually in the general theory of relativity a

[1]A reminder: the planets move almost in the same plane, which is known as the ecliptic plane (or simply the ecliptic). Here we're not taking into account that the planetary orbits are ellipses; this would lead to slight variations in r_{max} and r_{min}.

[2]This is one way of modeling the experimental observations. An alternative way would be to assume that the light travels a longer distance. This has the advantage of making the speed of light constant, in agreement with the special theory of relativity.—*Ed.*

[3]See "The Power of Dimensional Thinking" in the May/June 1992 issue of *Quantum*.—*Ed.*

Art by Sergey Barkhin

different characteristic value is used:

$$R_{gr} = \frac{2GM}{c^2}.$$

This gravitational radius is known as the Schwartzschild radius. The Sun's Schwartzschild radius is equal to 3 km, and the Earth's is only 9 mm.[4]

The gravitational potential can now be rewritten as

$$\frac{\phi}{c^2} = -\frac{1}{2}\frac{R_{gr}}{r}.$$

The left-hand side is the gravitational potential in dimensionless units—that is, the quantity does not have any dimensions of length, time, or mass. This means that its value does not change if we change our system of measurements. The quantity ϕ/c^2 is used to characterize the strength of the gravitational field in most cases.

A light beam passing near the surface of the Sun ($r \cong R_\odot = 7 \cdot 10^8$ m) can be expected to decrease in velocity by a value proportional to ϕ/c^2, since this is the only value characterizing the Sun's gravitational field:

$$\Delta v = \frac{cR_{gr}}{R_\odot}.$$

We can say that space near the Sun has the optical characteristics of a medium with a refractive index slightly greater than 1. If we assume that the gravitational field acts only near the Sun—for instance, over a distance of a few solar radii (we'll say 10 solar radii, for the sake of argument)—we can estimate that the travel time of the light increases by Δt, which is determined from the following equation:

$$t + \Delta t = \frac{10R_\odot}{v_{\text{light}}} = \frac{10R_\odot}{c - \Delta v}$$

$$\cong \frac{10R_\odot}{c}\left(1 + \frac{\Delta v}{c}\right),$$

[4]It's conventional to define R_{gr} with a factor of 2, although you may sometimes encounter formulas without this factor.

or

$$\Delta t \cong \frac{10R_\odot}{c}\frac{\Delta v}{c} = \frac{10R_\odot}{c}\frac{R_{gr}}{R_\odot} = \frac{10R_{gr}}{c}$$

$$= 10^{-4} \text{ s}.$$

This estimate yields a time during which a light beam would travel 30 km. Of course, this estimate is very approximate; in particular, the choice of the factor of 10 is extremely arbitrary. Not only that, the correct formula takes into account the distance between the Sun and the planets as well as the solar radius, since time travels more slowly along the entire path (not just near the Sun). Nevertheless, our estimate is useful for getting a feel for the problem. The more precise formula in the general theory of relativity is

$$\Delta t = \frac{2R_{gr}}{c}\left(1 + \ln\frac{R_\odot^2}{r_1 r_2}\right),$$

where the logarithm contains the ratio of the Sun's radius to the Earth–Sun distance r_1 and the ratio of the Sun's radius to the Mercury–Sun distance r_2. And it was the logarithm that we missed in our reasoning. This logarithm isn't insignificant—it's equal to 11.2. So the precise formula is

$$\Delta t = \frac{22.4R_{gr}}{c}.$$

In order to verify this formula with experimental observations, we need to know the moment corresponding to superior conjunction as if the Sun had no effect on the light beam. For this we need to know the astronomical distances with an accuracy of 1–2 km. Such requirements are at the outer limits of modern technology.

The experimentalist encounters still other problems. It's not so easy to know the location on the planet's surface where the light beam is reflected—what's being measured is the time the signal travels from the Earth to the planet and, after reflection, back to the Earth (the radar echo).

Nevertheless, such experiments were done by a group of American physicists. They measured the signals sent to Mercury, Venus, and Mars. The results corresponded to theory, but the errors were still too large (about 5–10%).

Figure 2 shows one of the curves for the signal delay on different days. Zero on the abscissa corresponds to the moment of superior conjunction.

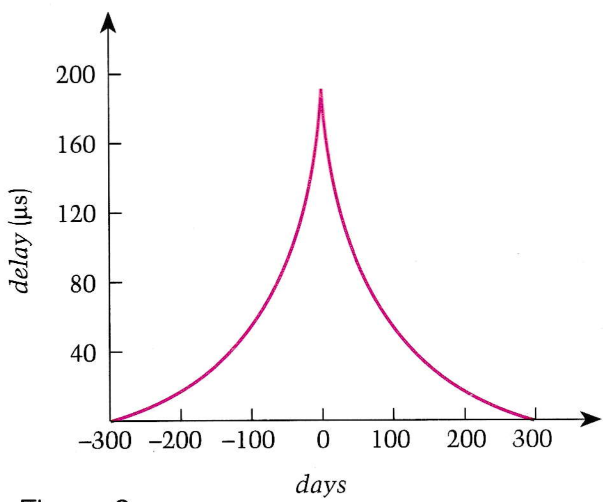


Figure 2

Light deflection in the Sun's field

As was mentioned above, space near the Sun affects a light beam as if it were an optical medium with a refractive index slightly higher than 1. This means that the light of distant stars should curve as it passes the Sun, much like what happens when it passes through a prism. In principle this phenomenon was known long ago. When Sir Isaac Newton presented the theory of light as a flow of tiny particles, it was clear to him that light should be attracted by the Sun. Since in a gravitational field the acceleration of all bodies is the same and doesn't depend on mass, the trajectory of light likewise doesn't depend on the light particle's mass and takes the form of a parabola. Remember, planetary masses aren't present in Kepler's laws of planetary motion.[5] From such considerations Sandemann obtained in 1801 the formula that resulted in a deflection angle of $\theta = R_{gr}/R_{\odot} \cong 0.85''$ for a light beam passing near the edge of the solar disk. However, this result turned out to be wrong. In 1915 Einstein worked out a new formula based on the general theory of relativity, and it gave a value for this effect that was twice as large:

$$\theta = \frac{2R_{gr}}{R_{\odot}} \cong 1.7''.$$

The light deflection was measured for the first time in 1919 by expeditions mounted by the Royal Astronomical Society of London to northern Brazil and the Gulf of Guinea to observe the total solar eclipse. On September 27, 1919, Einstein wrote to his mother: "Good news today! Lorentz just cabled me that the British expedition indeed proved the deflection of light near the Sun."

From then on, the Einstein effect has been measured during virtually every solar eclipse. Nevertheless, it is pretty difficult to obtain a value with a suitable accuracy. One needs to measure very accurately the positions of stars near the Sun and repeat the measurements after a half-year, when the Sun is no longer in that region of the sky. In the meantime, the state of the atmosphere has changed, the refraction in the Sun's atmosphere has changed—in short, a whole system of corrections has arisen, which makes it very difficult to compare data from the two measurements.

Nevertheless, many astronomers have worked long and hard to decrease the error, and have managed to reduce it to the point that it's now possible to talk about agreement between theory and experiment within error limits of no more than 1% of the magnitude of the effect measured.

To date the best results have been obtained in research involving eclipses of quasars (powerful sources of radio waves). The advantage of observing radio sources is obvious: it doesn't have to be dark out to record their radiation, so they can be studied at any time.

In conclusion, we can now consider it an established fact that massive celestial bodies act like huge converging lenses. The refraction is much greater than could be explained by the attraction of a light quantum to the Sun in accordance with Newton's law of universal gravitation. The law was imprecise, as it turned out: light is attracted more strongly than a simple body with a mass calculated according to the formula $mc^2 = h\nu$ (the energy of a quantum). And the phenomenon responsible is the curvature of space near a massive body—in this case, our Sun.

[5]See "The Fruits of Kepler's Struggle" in the January/February 1992 issue of *Quantum*.—*Ed.*

"DERIVATIVES IN ALGEBRAIC PROBLEMS"
CONTINUED FROM PAGE 31

already found a_0 (see example 2)—it's equal to $e^{1/e}$. Similarly we can find the value of a_1: if x_1 is the point where $y = a_1^x$ intersects with $y = x$, then $a_1^{x_1} = x_1$ and $x_1 \ln a_1 = -1$ (this is the slope of a_1^x at $x = x_1$). Eliminating x_1 and solving for a_1, we get $a_1 = e^{-e}$. Here's the final answer: the equation has three roots for $0 < a < e^{-e}$, one root for $e^{-e} \le a < 1$, two roots for $1 < a < e^{1/e}$, one root for $a = e^{1/e}$, and no roots for $a > e^{1/e}$.

(*Editor's note*: Although it's quite correct, this graphic solution seems almost too concise—deceptively simple. It would be a very good exercise for the reader to restore all the missing details. A full-scale solution is offered in the follow-up article on p. 44.)

Exercises

4. Find the number of roots of (a) $3x^4 + 4x^3 - 36x^2 = a$; (b) $e^x = ax$.

5. Solve $(x - 1)e^{x-1} + x^2 - 3x + 2 = 0$.

6. Solve these equations in two variables: (a) $1/x + 2\sqrt{x} = 3y(1 - \ln y)$; (b) $\ln x/x = e^{\cos y}$; (c) $4^x + 1 = 2^{x+1} \sin y$.

7. For every $n = 0, 1, 2, \ldots$, draw the set of points (p, q) in the (p, q)-coordinate plane for which the following equations for x have exactly n roots: (a) $x^3 = 3px + q$; (b) $p^x = x^q$ $(x > 0)$.

8. Without calculating the numbers e^π and π^e, determine which of them is larger.

ANSWERS, HINTS & SOLUTIONS ON INSIDE BACK COVER

The other half of what you see

Don't believe it 'til you prove it!

by Vladimir Dubrovsky

FIGURE 7 IN "DERIVATIVES IN Algebraic Problems" (p. 31), which illustrates the equation $a^x = \log_a x$, looks pretty convincing. But when you think more about it, you understand that the good advice to believe only half of what you see applies here perfectly. As it turns out, an accurate justification that the number of roots of this equation is correctly represented in this figure isn't so easy to provide as it may seem. This short article provides a proof.

First, consider the case $a > 1$. In this case, as the graph suggests, the equation

$$a^x = \log_a x \qquad (1)$$

is equivalent to a simpler equation

$$a^x = x. \qquad (2)$$

Indeed, a little algebra shows that equation (2) implies equation (1). To show that equation (1) implies equation (2), we can proceed indirectly. Suppose $a^x = \log_a x$. Can a^x be greater than x? Well, taking $\log_a$ of both sides preserves this inequality (since $a > 1$), so we would have $x < \log_a x$, or $a^x < x < \log_a x$, which is a contradiction. Similarly, $a^x > x$ implies $a^x > \log_a x$. Therefore, if $a^x = \log_a x$, the only possibility left is that $a^x = x$.

Equation (2) has been already studied in "Derivatives in Algebraic Problems." A question that might need an additional explanation is why there are *not more* than two roots for $1 < a < e^{1/e}$. However, it's easily answered by examining the function $f(x) = a^x - x$, whose zeros are the roots of equation (2). The derivative $f'(x) = a^x \ln a - 1$ is increasing and has only one zero $x_0 = -\ln \ln a$. So the function itself decreases for $x \le x_0$, has a minimum at x_0, and increases for $x \ge x_0$, which means that it can have at most one root in each of the intervals $(-\infty, x_0]$ and $[x_0, \infty)$. The actual number of roots depends on the minimum value $f(x_0)$: there are no roots, one root, or two roots if $f(x_0) > 0$, $f(x_0) = 0$, $f(x_0) < 0$ (or, as the graphs of a^x for different a's clearly show, $a > a_0$, $a = a_0$, $a < a_0$, where $a_0^{x_0} = x_0$), respectively. This is in full agreement with what was said in the article (it was shown there that $a_0 = e^{1/e}$).

The case $0 < a < 1$ needs a subtler inspection. First of all, we note that if $f(x) = a^x - x$, then $f'(x) = a^x \ln a - \ln x$. For $0 < a < 1$ and $x > 0$, we have $a^x > 0$, $\ln a < 0$, *and* $-\ln x < 0$. Hence, $f'(x) < 0$ for $x > 0$, and $f(x)$ decreases from 1 to $-\infty$ as x varies from 0 to ∞. Since it is still true (for $0 < a < 1$) that equation (2) implies equation (1), there is at least one root for equation (1). Since the graphs $y = \log_a x$ and $y = a^x$ are symmetric with respect to the line $y = x$, any other possible root x of equation (1) must have a counterpart, $x' = a^x = \log_a x$, which is also a root (the reader is invited to check this). Points (x, x') and (x', x) in the coordinate plane are symmetric points of intersection of the two graphs.

Now let's rewrite equation (1) as $x = \log_a \log_a x$, and use the formula $\log_a u = \ln u / \ln a$. Taking into account that $\ln a < 0$, we' get the equation

$$\ln(-\ln x) - x \ln a = \ln(-\ln a) \qquad (3)$$

The possible number of roots of equation (3) can be found by the method used above, except that now it's better to take two successive derivatives of the left side of equation (3), which we'll denote by $g(x)$:

$$g'(x) = \frac{1}{x \ln x} - \ln a,$$

$$g''(x) = -\frac{\ln x + 1}{x^2 \ln^2 x}.$$

Since $g''(x)$ has only one zero (at $x = 1/e$) and changes its sign at this point from plus to minus, $g'(x)$ has a local maximum at $x = 1/e$ (see figure 1 on the facing page) equal to $-e - \ln a$, which is the absolute maximum on $0 < x < 1$, the domain of equation (3). So for $\ln a \ge -e$ (that is, for $a \ge e^{-e}$), $g'(x) \le 0$, which means that $g(x)$ is decreasing on $0 < x < 1$, and equation (3) (and equation (1) as well) has exactly one root—that of equation (2). As a gets smaller than e^{-e}, the upper graph in figure 1 is shifted still higher. Clearly, it has two roots for $a < e^{-e}$, so $g(x)$ has three intervals of monotonicity, and, therefore, not

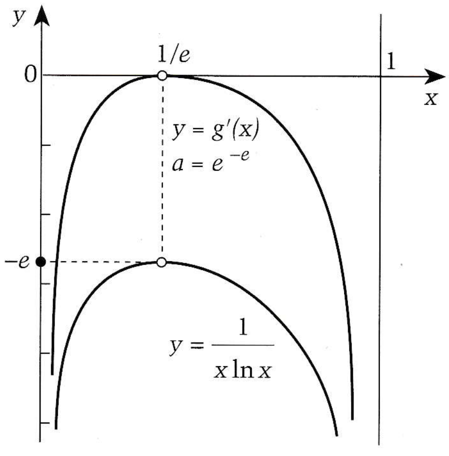


Figure 1

more than three roots. (A different argument can be found in the solution of problem M85 in the May/June issue, where it's proven that the derivative of a differentiable function with n zeros has at least $n - 1$ zeros. This problem is another good example of using calculus in algebraic problems.)

It remains to show that for $0 < a < e^{-e}$ equation (1) really does have three roots. To this end, we'll show that for these values of a the function $\phi(x) = a^x - \log_a x$ is decreasing in some small neighborhood of the root $x_0 = x_0(a)$ of equation (2). It will follow that $\phi(x_1) < \phi(x_0) = 0$ for some $x_1 > x_0$; since $\phi(1) = a > 0$, function ϕ must have a zero between x_1 and 1, and along with it a third, "counterpart" zero smaller than x_0, as we know.

Let's trace the value $s(a)$ of the derivative $a^x \ln a$ of a^x at point $x_0(a)$ as a decreases from e^{-e}. For $a = e^{-e}$ we have $x_0 = 1/e$ and $s(a) = -1$. For $a < e^{-e}$ the graph of $y = a^x$ lies below the corresponding graph for $a = e^{-e}$, so $x_0(a) < 1/e$ (fig. 2). Therefore, $s(a) = a^{x_0} \ln a = x_0 \ln a < \ln e^{-e}/e = -1$. But $\phi'(x_0) = s(a) - s(a)^{-1}$, because $(\log_a x)' = (x \ln a)^{-1}$, and $s(a)^{-1} > -1$. So $\phi'(x_0) < (-1) - (-1) = 0$, which means that $\phi'(x)$ is negative in some neighborhood of x_0, and we're done.

Summing up, we've proved that our equation has three roots for $0 < a < e^{-e}$, one root for $e^{-e} \le a < 1$ and $a = e^{1/e}$, two

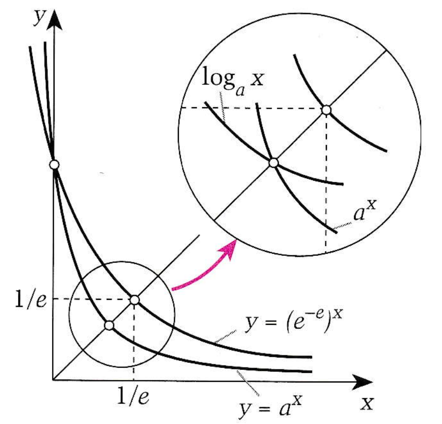


Figure 2

roots for $1 < a < e^{1/e}$, and no roots for $a > e^{1/e}$.

This problem provides an excellent (and nontrivial) opportunity to use graphing calculators and even more sophisticated computer tools. It really requires some effort to make them show you the three roots in the case $0 < a < e^{-e}$ or to compute the roots, even for $a = 1/16$, when two of them are 1/2 and 1/4. ◘

"ONES UP FRONT"
CONTINUED FROM PAGE 20

said to hit M if $\mathbf{u} = \overrightarrow{OU}$, where O is the origin and point U lies in M. Let $\mathbf{a} = (a_1, a_2)$ be a vector such that the numbers a_1, a_2, and 1 are rationally independent. That is, a linear combination of $n_1a_1 + n_2a_2$, with integers n_1 and n_2, is itself an integer only for $n_1 = n_2 = 0$. Consider an infinite sequence of vectors $\mathbf{a} + \mathbf{b}$, $2\mathbf{a} + \mathbf{b}$, ..., $n\mathbf{a} + \mathbf{b}$, ..., where $\mathbf{b}$ is any vector at all. Then the probability that the fractional part of a term in this sequence taken at random hits M is equal to the area of M.[3]

18. A flea is jumping on an infinite chessboard of unit squares. It moves a distance x to the left and a distance y upward with each jump. Prove that if the numbers x and y are rationally independent with one, the flea will necessarily hit a black square. Will this remain true if we require only that x, y, and y/x be irrational?

19. The numbers λ_1, λ_2, and π are rationally independent. Prove that the simultaneous inequalities

$$\begin{cases} \sin n\lambda_1 > 0.999999, \\ \sin n\lambda_2 > 0.999999 \end{cases}$$

have a positive integer solution n. ◘

ANSWERS, HINTS & SOLUTIONS ON PAGE 59

[3]The further generalization of this theorem—to n-dimensional space—is true as well. (The "difference" of the sequence—vector α—must have coordinates none of which is representable as a linear combination of other coordinates and the number 1 with rational coefficients, and the area of the set M must be replaced by its n-dimensional volume.)

Electricity in the air

*"Go, wondrous creature! mount where science guides,
Go, measure earth, weigh air, and state the tides . . .
Go, teach Eternal Wisdom how to rule—
Then drop into thyself and be a fool!"*

—Alexander Pope

by Arthur Eisenkraft and Larry D. Kirkpatrick

THIS MONTH'S CONTEST problem is based on part of one of the theoretical problems given at the XXIV International Physics Olympiad that was held in Williamsburg, Virginia, in July (see the September/October 1993 issue of *Quantum*). The problem was written by Anthony French of MIT, who served as the chair of the examinations committee, and is based on an actual application of physics to a real-world situation. The first part of the solution is based on Gauss's law, one of the most fundamental laws of electricity and magnetism.

Carl Friedrich Gauss was the greatest mathematician of his time and along with Archimedes and Newton may have been one of the three greatest mathematicians ever. He developed the method of least squares for fitting curves to data points and used this method to calculate an orbit for Ceres, the largest of the asteroids, after it couldn't be found. He was honored for this work when the name Gaussia was given to the 1001st asteroid. The gauss—a unit of magnetic field strength equal to 10^{-4} tesla—honors his work in magnetism. While still a university student he devised a method of drawing a seventeen-sided regular polygon using only a compass and straightedge. He then went further to show that certain regular polygons (for example, one with seven sides) could not be constructed this way.

Gauss's law tells us that the electric flux through a closed surface is proportional to the electric charge that is enclosed by that surface. To calculate the electric flux, we imagine dividing the surface into many small regions. For each region the contribution to the electric flux is given by the component of the electric field perpendicular to the surface E_n times the surface area A of that region. By convention, the contribution is positive if the electric field is directed out of the enclosed volume and negative if the electric field is directed inward.

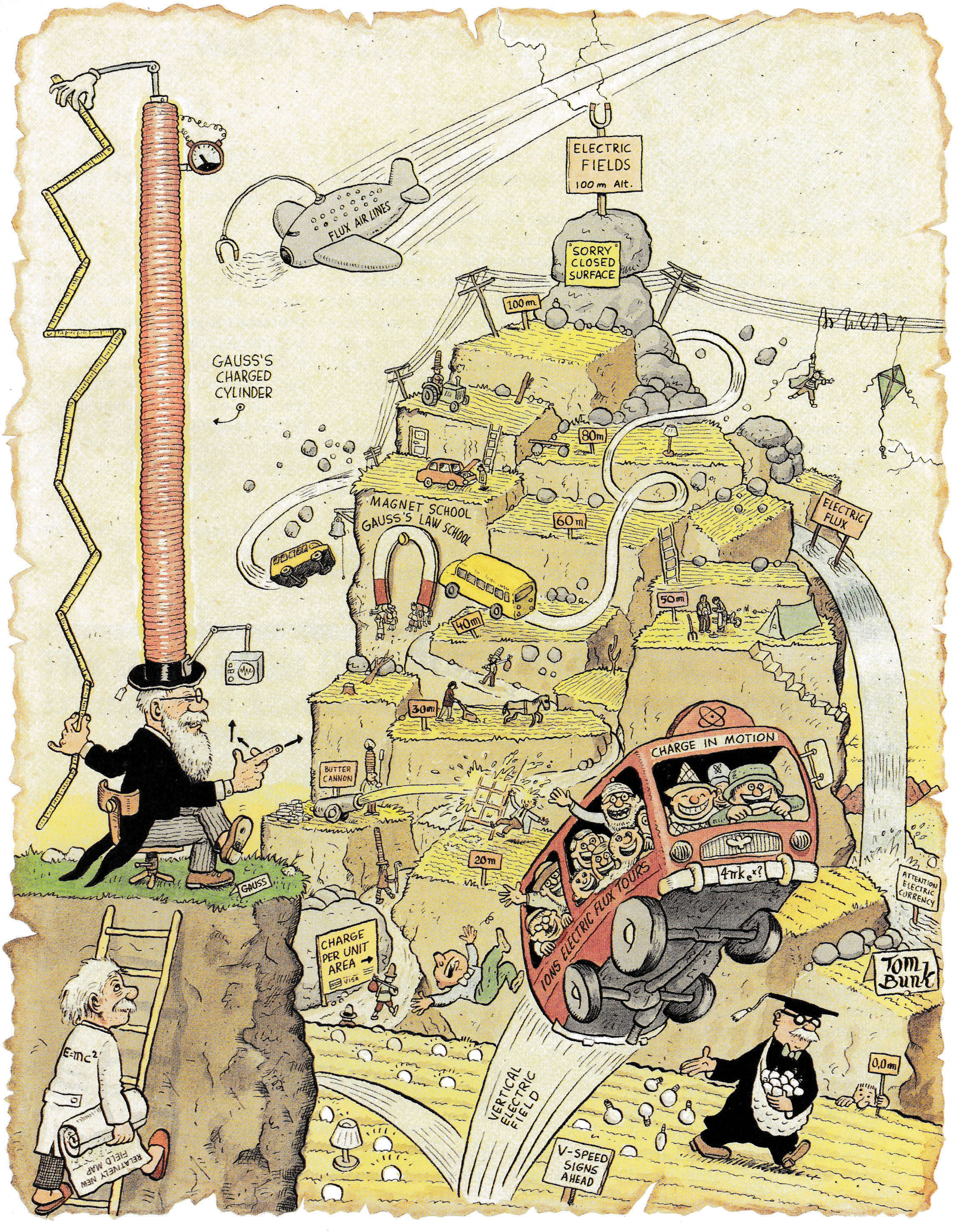
Because total electric flux is just the sum of all of the individual contributions, we can write Gauss's law in the form

$$\sum E_n A = \frac{q_{enc}}{\varepsilon_0},$$

where $\varepsilon_0 = 8.85 \times 10^{-12}\ C^2/(N \cdot m^2)$ is the permittivity of free space. For more information about Gauss's law, see the contest problem in the July/August 1992 issue of *Quantum*.

Gauss's law is very useful for finding electric fields in cases of high symmetry. For example, let's find the electric field outside of an infinitely long, straight wire carrying a positive charge per unit length λ. To exploit the symmetry, we choose the gaussian surface to be a cylinder of radius r and length L that is coaxial with the wire. By symmetry, we expect that the electric field will point radially outward from the wire and have the same magnitude at a given distance from the wire. This means that the electric field will be parallel to the ends of the cylinder and will not contribute to the flux. Therefore, the flux is given by the electric field times the area of the curved surface of the cylinder:

Art by Tomas Bunk



ELECTRIC FIELDS 100 m Alt.
FLUX AIR LINES
SORRY CLOSED SURFACE
100m
GAUSS'S CHARGED CYLINDER
80m
MAGNET SCHOOL
GAUSS'S LAW SCHOOL
60m
ELECTRIC FLUX
50m
40m
30m
BUTTER CANNON
CHARGE IN MOTION
20m
GAUSS
TONS ELECTRIC FLUX TOURS
ATTENTION ELECTRIC CURRENCY
CHARGE PER UNIT AREA
TOM BUNK
E=mc²
RELATIVELY NEW FIELD MAP
VERTICAL ELECTRIC FIELD
0,0m
V-SPEED SIGNS AHEAD

$$\sum E_s A = E2\pi rL.$$

The enclosed charge is equal to the charge per unit length times the length of the cylinder:

$$q_{\text{enc}} = \lambda L.$$

Putting these two expressions into Gauss's law, we can solve for the magnitude of the electric field:

$$E = \frac{\lambda}{2\pi\varepsilon_0 r}.$$

Notice that the length of the gaussian cylinder cancels as we expect.

From the standpoint of electrostatics, the surface of the Earth can be considered a good conductor that carries a total charge Q_0 and an average surface charge density σ_0. We can also consider the Earth a perfect sphere with a radius R = 6,400 km to simplify the geometry. Under fair-weather conditions, this surface charge density produces a downward electric field E_0 at the Earth's surface equal to about 150 V/m.

A. Use Gauss's law to calculate the magnitude of the Earth's surface charge density and the total charge carried on the Earth's surface. Is this charge positive or negative?

The magnitude of the downward electric field is observed to decrease with height and is about 100 V/m at a height of 100 m. This occurs because the air above the Earth's surface contains a net charge.

B. Use Gauss's law to calculate the average net charge per cubic meter of the atmosphere between the Earth's surface and an altitude of 100 m. Is this charge positive or negative?

The net charge density you calculate in part B is actually the result of having almost equal numbers of positive and negative singly charged ions ($q = 1.6 \cdot 10^{-19}$ C) per unit volume (n_+ and n_-). Near the Earth's surface, under fair-weather conditions, $n_+ \cong n_- \cong 6 \cdot 10^8\ \text{m}^{-3}$. These ions move under the action of the vertical electric field and their speed v is proportional to the strength of the electric field:

$$v \cong 1.5 \cdot 10^{-4} \times E,$$

where v is in m/s and E is in V/m.

C. How long would it take for the motion of the atmospheric ions to neutralize half of the Earth's surface charge, if no other processes such as lightning occurred to maintain it?

Please send your solutions to *Quantum*, 1840 Wilson Boulevard, Arlington, VA 22201 within a month after receipt of this issue. The best solutions will receive special certificates from *Quantum*.

Animal magnetism

The best solution to the May/June contest problem was submitted by Eric Joanis of Waterloo, Ontario. This problem appeared on the semifinal exam that was used to select the 1993 US Physics Team that competed in the International Physics Olympiad.

In the problem we asked our readers to show that the mass of a particle in a mass spectrometer is given by

$$m = \frac{qB^2R^2}{2V}. \qquad (1)$$

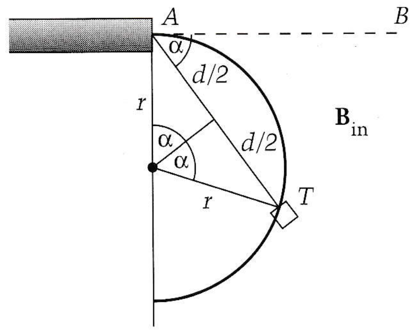
As explained in the problem, a particle with a mass m and charge q gains kinetic energy as it travels through a potential difference V according to

$$\tfrac{1}{2}mv^2 = qV. \qquad (2)$$

Once the particle enters the magnetic field B, the magnetic force provides the centripetal acceleration

$$qvB = \frac{mv^2}{R}.$$

Combining these two equations yields equation (1) for the mass of the particle, which can be determined when the radius of its path



can be measured.

Part B of the problem involved some target practice with an electron in a magnetic field. If the field is perpendicular to the page, the particle travels in a circular path in the page. Since we know the mass of the electron, we can solve equation (1) for the magnetic field:

$$B = \frac{1}{R}\sqrt{\frac{2mV}{q}}.$$

From the geometry in the figure above, we see that

$$\frac{d}{2} = R\sin\alpha.$$

Therefore,

$$B = \frac{2\sin\alpha}{d}\sqrt{\frac{2mV}{q}}.$$

If the magnetic field is parallel to AT, the problem grows in complexity. There are now two components of the velocity. The component parallel to the field, $v\cos\alpha$, is unaffected by the field. The component perpendicular to the field, $v\sin\alpha$, causes the electron to move in a circle. The combined motion is that of a helix.

If the electron traveling along the helical path is going to hit the target T, the time it takes to travel a distance d (due to the parallel component) must equal the time it takes to complete one circle of the helix (due to the perpendicular component). The parallel time is given by

$$t_{\parallel} = \frac{d}{v\cos\alpha}. \qquad (3)$$

The perpendicular time is

$$t_{\perp} = \frac{2\pi R}{v \sin\alpha}.$$

Because the radius R of the helix is determined by the component of the velocity perpendicular to the field, we have

$$R = \frac{mv \sin\alpha}{qB},$$

and the perpendicular time becomes

$$t_{\perp} = \frac{2\pi m}{qB}. \qquad (4)$$

Since the two times must be equal, we can equate equations (3) and (4) and solve for B to obtain

$$B = \frac{2\pi mv \cos\alpha}{qd}.$$

Solving equation (2) for v and substituting, we find that

$$B = \frac{2\pi \cos\alpha}{d}\sqrt{\frac{2mV}{q}}.$$

Note that the direction of the field does not matter.

The electron will also hit the target if it completes two circles or three circles or k circles before it travels the parallel distance to T. In that case we must modify the final equation to take this into account:

$$B = k\frac{2\pi \cos\alpha}{d}\sqrt{\frac{2mV}{q}}.$$

Part C of the contest problem asked readers to find the numerical values for the magnetic field given V = 1,000 V, d = 5 cm, and α = 60°. For the field perpendicular to the page, B = 3.7 mT; for the field parallel to the page, B = k(6.7 mT).

MATH
INVESTIGATIONS

Periodic binary sequences

Complex formulas for simple things

by George Berzsenyi

IN THIS COLUMN WE'LL EXplore periodic sequences of 0's and 1's defined by

$$a_n = \frac{1-(-1)^{f(n)}}{2}, \quad n = 0, 1, 2, \ldots,$$

where $f(n)$ is to be specified. For the simplest choices of $f(n)$—that is, if $f(n) \equiv 0$ or $f(n) \equiv 1$—we simply reproduce them; whereas if $f(n) = n$, we obtain $\langle a_n \rangle = \langle 0, 1, 0, 1, \ldots \rangle$; and if $f(n) = n + 1$, we get $\langle a_n \rangle = \langle 1, 0, 1, 0, \ldots \rangle$. The latter are a bit more promising, but clearly we must employ better machinery to obtain more interesting sequences of 0's and 1's.

If $f(n) = \lfloor n/2 \rfloor$, where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x, we get $\langle a_n \rangle = \langle 0, 0, 1, 1, 0, 0, 1, 1, \ldots \rangle$; while for $f(n) = \lfloor n/2 \rfloor + 1$, we get $\langle a_n \rangle = \langle 1, 1, 0, 0, 1, 1, 0, 0, \ldots \rangle$. My first challenge to my readers is to **prove that the choices $f(n) = \lfloor n^3/2 \rfloor$ and $f(n) = \lfloor n^3/2 \rfloor + 1$ produce the sequences $\langle a_n \rangle = \langle 0, 0, 0, 1, 0, 0, 0, 1, \ldots \rangle$ and $\langle a_n \rangle = \langle 1, 1, 1, 0, 1, 1, 1, 0, \ldots \rangle$**. My next challenge is to **construct the functions that will similarly generate the two other nontrivial sequences of period length 4: $\langle 0, 1, 1, 0, 0, 1, 1, 0, \ldots \rangle$ and $\langle 1, 0, 0, 1, 1, 0, 0, 1, \ldots \rangle$**.

Clearly, one should also be able to generate all sequences of 0's and 1's of period length 3, 5, 6, and so on. Still other challenges await you. For some of them, you might need to investigate functions of the form $f(n) = \lfloor p(n) \rfloor$, where $p(n)$ is a higher order polynomial. Hence the following questions arise: **Do all polynomials $p(n)$ lead to periodic sequences $\langle a_n \rangle$? Can one determine from the degree and/or coefficients of $p(n)$ the period length of $\langle a_n \rangle$? Can one construct in such manner all periodic sequences of 0's and 1's? If not, what other simple machinery is needed to accomplish the task?** Some of these questions may be quite difficult, so you should be happy with partial results.

The purpose of this column is to direct the attention of *Quantum*'s readers to interesting problems in the literature that deserve to be generalized and could lead to independent research and/or science projects in mathematics. Students who succeed in unraveling the phenomena presented are encouraged to communicate their results to the author either directly or through *Quantum*, which will distribute among them valuable book prizes and/or free subscriptions.

Reader response

I'd like to thank my readers for communicating their thoughts to me concerning the problems discussed in earlier columns. In particular, I'm deeply indebted to Brian Platt, whose insightful comments and continued interest are most appreciated. In a future column, I'll share some of Mr. Platt's original investigations about chaotic behavior, which he was kind enough to share with me.

I also wish to thank Mark Rupright for his wonderful solutions of the problems posed in my column "Digitized Multiplication à la Steinhaus" (July/August 1993); and Michael Filaseta and Ben Rahn, who submitted their proofs that three 3's cannot occur in any of the rows of Hilgemeier's "likeness sequence" (presented in the last issue). Mr. Rahn's proof is reproduced below for my readers' scrutiny:

Assume three 3's occur consecutively in certain rows. Let S be the set of all natural n such that row n contains three consecutive 3's. By the Well Ordering Principle, there is a least element of the set S—call it k. If row k contains three consecutive 3's, then either the first two or the last two of the three 3's describe the presence of three consecutive 3's in the $k - 1$ row. Thus, $k - 1$ is also in the set S. Note that k is not 1, so $k - 1$ is still a natural number. But this contradicts the fact that k is the least element in set S. Therefore, three consecutive 3's never occur in any given row.

The American Mathematics Correspondence School

Math by mail for ambitious high school students

Attentive readers of our magazine were probably perplexed when the Math by Mail department, introduced in the March/April 1991 issue of *Quantum*, never resurfaced. Happily, the idea of a mathematics correspondence school in the United States did not fade away. The eminent Russian mathematician I. M. Gelfand, who founded the Mathematics Correspondence School in the Soviet Union almost 30 years ago, has been instrumental in developing a similar project in this country: the American Mathematics Correspondence School (AMCS).

AMCS is sponsored by the Center for Mathematics, Science, and Computer Education at Rutgers University, where Prof. Gelfand now teaches. The program gives ninth-grade students an opportunity to develop their mathematical ability by working with university mathematicians. They hone their skills on highly effective, nonstandard problem-solving models in algebra, geometry, and analytical geometry. The school is independent of the school day, but teachers are encouraged to become mentors to students in their schools.

Interested students who apply to become part of AMCS take an entrance exam to determine their mathematical aptitude. Those who perform satisfactorily are admitted to the school and receive bimonthly assignments. The students write solutions and explanations of their work and mail them to Rutgers, where they are reviewed by faculty members and graduate students in the Department of Mathematics. These mathematicians then send their comments back to the students.

The texts for AMCS are books written by Prof. Gelfand and his colleagues. Two of them—*The Method of Coordinates* and *Functions and Graphs*—have been translated from the Russian and published by Birkhäuser. These paperbacks cover a great deal of mathematics in a compact format. Additional texts in geometry, algebra, and trigonometry are in preparation.

Here are a few examples of the types of problems you might encounter on the AMCS entrance exam:

✎ You have a glass of wine and a glass of water. You take a spoonful of the wine and pour it into the glass of water, stir the mixture, and take a spoonful of that and pour it into the glass of wine. Is there now more wine in the water or more water in the wine?

✎ What is the maximum number of Saturdays there can be in a year?

✎ Into how many parts can four distinct straight lines divide a plane? Draw an example for each case.

The American Mathematics Correspondence School began in 1991–92 with a program for ninth graders primarily in New Jersey. For 1992–93 AMCS continued with these students (who entered Level Two) and began a new program for entering students (Level One). AMCS is currently accepting applications for 1993–94. The registration fee is $50 (due upon return of the entrance exam). However, no student should be deterred from applying because of financial considerations.

The Mathematics Correspondence School in the former Soviet Union graduated 70,000 students, many of whom have gone on to become prominent mathematicians and scientists. Its US cousin hopes to encourage mathematical talent here in much the same way. For further information about the American Mathematics Correspondence School, please contact

Harriet Schweitzer
Assistant Director
Center for Mathematics, Science, and Computer Education
SERC Building—Room 239
Busch Campus—Rutgers University
Piscataway NJ 08855-1179

E-mail: harriets@gandalf.rutgers.edu
Phone: 908 932-0669
Fax: 908 932-3477

Bulletin Board

NYNEX Science and Technology Awards

The NYNEX Foundation invites teams of high school students in seven northeastern states—New York, Massachusetts, Maine, Vermont, New Hampshire, Rhode Island, and Connecticut—to devise their own practical solutions to community problems using science and technology. But in this competition, the winning ideas won't just sit on the drawing board. In addition to $210,000 in scholarship money to be awarded, the NYNEX Science and Technology Awards will provide development grants totaling $250,000 to the top three teams to enable them to bring their winning ideas closer to fruition. How? By working as interns with scientists or urban planners to carry out a pilot project in a real-life setting, or to build the prototype of a new invention, or to test a theory in a sophisticated laboratory.

Administered by the National Science Teachers Association (NSTA), the competition calls for teams of two to four high school students to focus on a specific problem and come up with a scientifically sound solution. Students may choose any issue affecting the public quality of life in a specific geographic area—providing vital services, serving people in need, preventing crime, or protecting the environment, to name just a few possible areas of investigation.

A panel of judges will choose the 12 finalist teams, who will come to Washington, DC, in April for the final judging and awards, including $60,000 for the first-place team and up to $40,000 for the second-place team. All team awards must be used by the students to cover future educational expenses.

In its inaugural year the competition is restricted geographically to the states listed above. Future competitions may expand to include the remaining states.

Application materials are being sent to teachers in October. Additional applications can be obtained by calling 800 9X-TEAMS. The deadline for entering is February 11, 1994. Preliminary judging will be held in New York City in March.

First Step to a Nobel Prize in Physics

The Institute of Physics of the Polish Academy of Sciences announces the Second International Competition in Research Projects in Physics for Secondary School Students. Last year 134 papers were submitted by students from 23 countries. Three students won a diploma and a research stay at the Institute of Physics: Melvin Boon Tiong of Singapore, "Estimating the Attractor Dimensioin of the Equatorial Weather System"; Ian Galloway, "Beta Backscattering by Metallic Elements and Simple Components"; Dmitry Ruslanovich Bituck, "The Dynamics of the Earth's Climate Complex Behaviour." David Zeltser of the USA won an honorable mention with his paper "A Predicted Uncertainty Principle and Mechanical Unit of Charge Based on Analogy and Dimensional Analysis." The organizers have decided to publish a supplement to *Acta Physica Polonica* (in cooperation with its editors) containing selected papers from the first competition.

The general rules of the competition are as follows:

1. All secondary school students are eligible for the competition. The only conditions are that her or his school cannot be considered a university college, and the participant must not turn 20 years of age before March 31, 1994.

2. There are no restrictions on the subject matter of the papers, their level, or the methods used. The student has full discretion in these areas. However, the papers must have a research character and deal with physics or topics directly related to physics.

3. A participant can submit one paper or several, but each paper must have only one author. The paper should not exceed 20 normal typed pages.

4. The papers will be judged by the Organizing Committee. The number of papers receiving awards or other citations is not restricted. All awards in each category are considered equivalent. The authors of award-winning papers will be invited to the Institute of Physics for a one-month research stay. Expenses in Poland will be paid by the institute; winners will be responsible for travel expenses to and from Poland.

5. Two copies of each paper, in English, should be sent by March 31, 1994, to

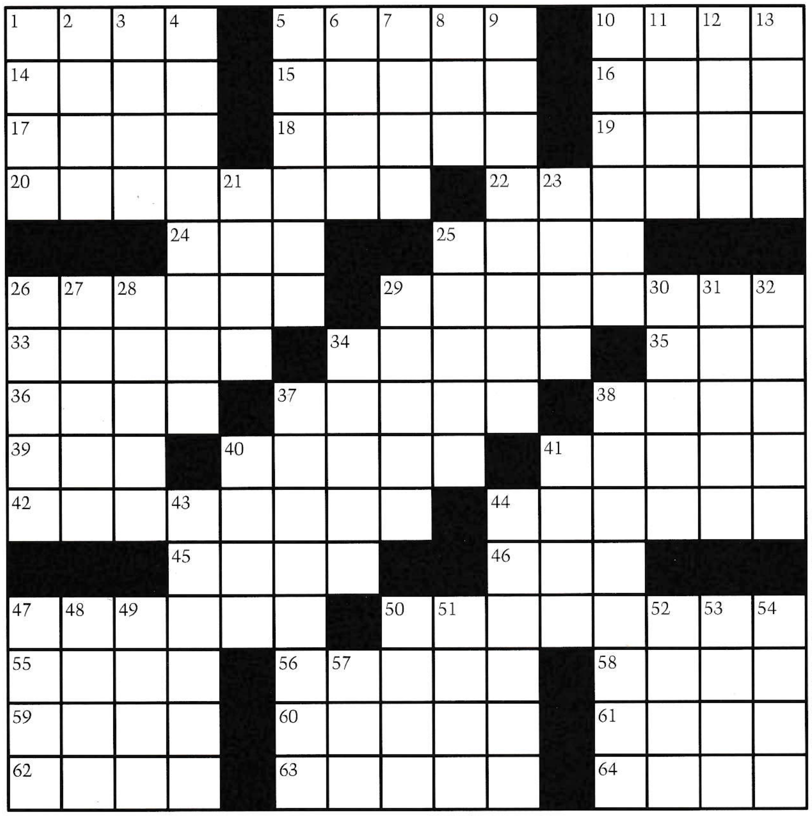
Dr. Waldemar Gorzkowski
Secretary General of "First Step"
Institute of Physics, Polish Academy of Sciences
al. Lotników 32/46, (PL) 02-668 Warszawa
POLAND

6. Each paper should contain the name, birth date, and home address of the author and the name and address of his or her school.

For further information on the competition, contact Dr. Gorzkowski at the address above; by phone at (022)435212; by fax at (022)430926; by e-mail at gorzk@gamma1.ifpan.edu.pl or gorzk@planif61.bitnet.

Criss×cross science

by David R. Martin



Across

1 Repose
5 H_2O
10 Units of radiation dose
14 Turkish title
15 Eagle nest
16 QED word
17 Payoff
18 Lima and pinto
19 Good-bye in London
20 Line of constant temperature
22 Weather map line
24 Atmosphere
25 Russian ruler
26 Delicious and golden
29 Angle from equator
33 Medicinal mass
34 Type of prism
35 Asian country (for short)
36 Type of cheese
37 An arsenic or antimony atom in silicon
38 Teenage problem
39 Department in France
40 Gives up (land)
41 Dermal bony plate
42 Tridiagonal and unitary, e.g.
44 Irish
45 Greek letters
46 Friend: Fr.
47 Foot part
50 Type of gland
55 Hawaiian island
56 African antelope
58 Moslem prayer leader
59 Level
60 Twelve
61 Donuts
62 TV's Alda
63 White poplar tree
64 River in France

Down

1 Jewish teacher
2 Selves
3 Scat!
4 Acid resistant metal
5 Magnetic flux units
6 British philosopher Sir Alfred Jules
7 Ore cart
8 One: German
9 Current impeder
10 Distilling apparatus
11 Bedouin
12 Experiment results
13 Large plasma ball
21 Hurries
23 Go by boat
25 Mexican food
26 Nautical term
27 Small platforms
28 Type of kingdom
29 Shortest paths between points
30 Complete
31 Author of the *Divine Comedy*
32 Moslem ruler
34 Points of minimum disturbance
37 Order of crustaceans
38 Fugacity of a gas
40 Quote
41 Type of truck
43 Keyboard word
44 Salty
47 Very small amount
48 Hit the __ on the head
49 NY stadium
50 Extent
51 Adam's son
52 Hebrew prophet
53 Scarce
54 Earliest being in Scand. myth.
57 Steal

SOLUTION IN THE NEXT ISSUE

SOLUTION TO THE SEPTEMBER/OCTOBER PUZZLE

P	A	S	T	■	L	E	A	S	E	■	M	A	R	T
A	R	E	A	■	E	A	S	E	L	■	A	L	O	E
L	I	E	N	■	S	T	I	P	E	■	G	I	L	L
M	A	N	T	I	S	S	A	■	M	A	N	T	L	E
■	■	■	A	R	E	■	■	R	E	N	E	■	■	■
S	T	A	L	I	N	■	G	E	N	E	T	I	C	S
H	U	M	U	S	■	S	A	L	T	S	■	R	A	H
A	D	A	M	■	B	E	T	A	S	■	D	A	T	E
R	O	T	■	D	E	W	E	Y	■	M	A	T	E	D
P	R	I	M	A	T	E	S	■	M	E	T	E	R	S
■	■	■	A	R	A	R	■	■	E	T	A	■	■	■
O	F	F	S	E	T	■	P	A	R	A	B	O	L	A
B	A	R	E	■	R	O	R	I	C	■	A	P	O	D
I	C	E	R	■	O	V	I	N	E	■	S	A	L	E
T	E	E	S	■	N	A	M	U	R	■	E	L	L	S

ANSWERS, HINTS & SOLUTIONS

Math

M96

Two examples of the required pentagon are shown in figure 1. It can be shown that any pentagon satisfying the condition of the problem is similar to one of these pentagons—with appropriate values of the parameters α and β, of course. The only way to divide a pentagon into two pentagons (one shaded and one white) is to cut it along two adjacent segments AB and BC, where A is a vertex of the pentagon and C a point on the opposite side. Let α and 360° – α be the measures of the angles formed at point B, with α (α < 180°) in the shaded part. Let ϕ and 180° – ϕ be the angles formed at C, with ϕ in the shaded part. Angles α and ϕ of the shaded piece must correspond, in view of the congruence of the pieces, to two angles of the white piece. One of these two angles must be 180° – ϕ, because otherwise either piece would have four angles—α, ϕ, 360° – α, and 180° – ϕ—whose sum, 540°, is equal to the total sum of all five angles of any pentagon (obviously, neither α nor ϕ can correspond to 360° – α). If α = 180° – ϕ, from the shaded piece (fig. 1a) we see that one of the two angles of this piece next to α is ϕ = 180° – α, and the white piece shows that the other angle next to α = 180° – ϕ is 360° – α. Then the remaining angles can be labeled β and α – β (to get the total of 540°). Matching equal angles of the two pieces, we find that the corresponding sides marked in the figure must be equal. The case ϕ = 180° – ϕ = 90° (fig. 1b) is examined similarly.

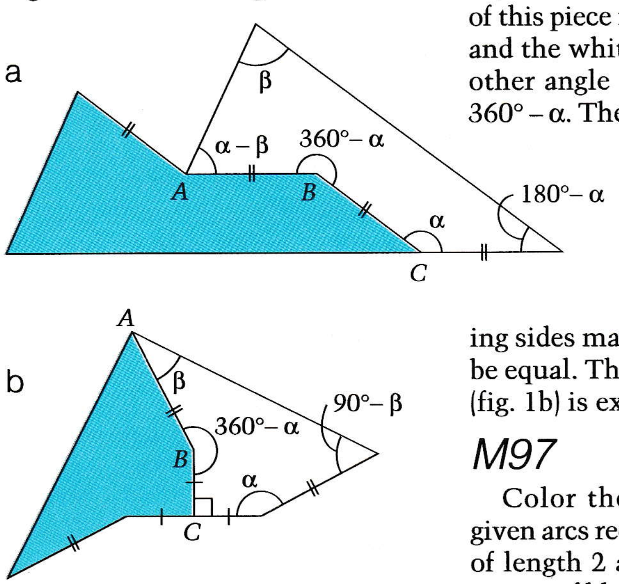


Figure 1

M97

Color the $3k$ endpoints of the given arcs red and subdivide the arcs of length 2 and 3 into unit arcs by means of black points. Thus, we get an additional $k + 2k = 3k$ black points, which, along with the $3k$ red points, make up the vertices of a regular $6k$-gon. The vertices of the figure are diametrically opposite each other.

Suppose the statement of the problem were wrong. Then every red point of the given circle C_0 is diametrically opposite a black point. So opposite every unit arc with red endpoints there lies a unit arc with black endpoints—that is, the middle third of an arc of length 3 with red endpoints (fig. 2a). We remove these two opposite arcs and bend the two remaining (larger) arcs until the endpoints of the arcs we removed overlap. We arrive at a smaller circle C_1 (fig. 2b), with an inscribed $(6k - 2)$-gon. This new circle consists of $6k - 2$ unit arcs with $3k - 1$ red endpoints and as many black ones. Since the points of C_0 that were diametrically opposite remain so on C_1, every red point stays opposite a black one. Our operation decreases by one the numbers of arcs of length 1 and 3 with red endpoints and increases by one the number of arcs of length 2.

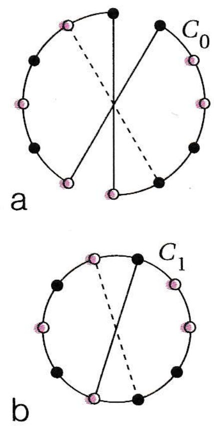


Figure 2

Repeat this operation with the circle C_1 to obtain a circle C_2 of length $6k - 4$, and follow suit until we arrive at a circle C_k in which there are only $2k$ arcs of length 2, with red endpoints and black midpoints, forming a regular polygon with $4k$ vertices. This is a contradiction, because in such a polygon every red point is opposite another red point, whereas our operations preserve diametrically opposite pairs, which originally consisted of a red and black point each.

The statement of the problem is generalized to the case of a circle divided into k arcs of length 1, l arcs of length 2, and k arcs of length 3 with an even sum $k + l$. (V. Dubrovsky)

M98

The only positive integer solution of the given equation is $(x, y) = (2, 5)$. Let's show that there are no other solutions.

For positive integers x and y, the right side of the equation is positive. So let's first find the pairs (x, y) such that the left side of the equation is positive—that is,

$$x^y > y^x, \qquad (1)$$

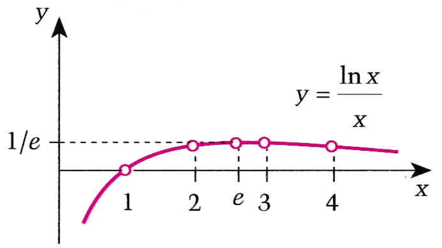


Figure 3

or $(\ln x)/x > (\ln y)/y$. Using the derivative of the function $(\ln x)/x$, which equals $(1 - \ln x)/x^2$, we find that this function increases on the interval $(1, e]$ and decreases for $x \geq e$ (see the graph in figure 3). Since $2 < e < 3$ and $(\ln 4)/4 = (\ln 2)/2 < (\ln 3)/3$ (because $3^2 > 2^3$), we get the following list of the pairs (x, y) satisfying inequality (1): $(x, 1)$ with $x > 1$; $(2, y)$ with $y \geq 5$; $(3, 2)$; and (x, y) with $3 \leq x < y$. We can verify by direct substitution into the equation that the pairs $(x, 1)$ and $(3, 2)$ do not satisfy it and the pair $(2, 5)$ does. If $y \geq 6$, we have $2y > y^2$, and also $2^{y-1} > (y - 1)^2$ (since $2^k > k^2$ for any value of $k > 4$). Hence

$$2^y - y^2 = 2 \cdot 2^{y-1} - y^2 > 2(y-1)^2 - y^2 = y^2 - 4y + 2 > y + 2.$$

So these pairs must be discarded.

It remains to show that our equation has no solutions for $3 \leq x < y$ either. Fix x, putting $x = a \geq 3$, and consider the difference $a^y - y^a$ as a function of y for $y \geq a + 1$ (which is equivalent to $y > a$ for integers y and a). Write it as

$$a^y - y^a = \left(a^y - \frac{y^{a+1}}{a+1}\right) + \left(\frac{y^{a+1}}{a+1} - y^a\right).$$

Then the derivative of the first term equals

$$\ln a \cdot a^y - y^a = (\ln a - 1)a^y + (a^y - y^a) > 0,$$

because $a \geq 3 > e$ and $a^y - y^a > 0$ by inequality (1). So the first term is an increasing function of y and, therefore, is greater than its value at the point $y = a$. This fact, together with some algebraic manipulation, yields

$$a^y - y^a > \left(a^a - \frac{a^{a+1}}{a+1}\right) + y^a \frac{y-a-1}{a+1} = \frac{a^a}{a+1} + \frac{y^a}{a+1}(y-a-1) \quad (2)$$
$$> a^{a-2} + y^{a-1}(y-a-1),$$

(because $a^2 > a + 1$ for $a \geq 3$, and $y \geq a + 1$). For $y \geq a + 2$, we know that $y - a - 1 > 1$, so the right side of equation (2) is not less than

$$a^{a-2} + y^{a-1} > a + y.$$

For $y = a + 1$, $a \geq 4$, it equals

$$a^{a-2} \geq 4a^{a-3} \geq 4a > 2a + 1 = a + y.$$

Finally, for $a = 3$, $y = a + 1 = 4$, we simply calculate

$$a^y - y^a = 3^4 - 4^3 = 17 > 3 + 4 = a + y.$$

This problem provides an additional illustration of how derivatives are used in solving equations (see also the article beginning on page 28). (A. Zaychik, V. Dubrovsky)

M99

The geometric idea of the solution is concentrated in the following lemma:

If five squares with sides parallel to the coordinate axes have a common point, then one of them contains the center of another.

To prove it, assume that the common point of the squares is the origin O. By the pigeonhole principle, two of the five centers must lie in the same quadrant (defined by the axes), and we may assume this is the first quadrant. Denote these squares by S_1 and S_2, and their centers by O_1 (x_1, y_1) and O_2 (x_2, y_2), respectively. Choose the greatest of the four coordinates of O_1 and O_2—let it be x_1. Figure 4 shows that the square with the bottom left vertex at O and the right side passing through O_1 lies in S_1. This new square consists of all the points (x, y) such that $0 \leq x \leq x_1$, $0 \leq y \leq x_1$. In particular, it contains O_2. So S_1 also contains O_2, and we

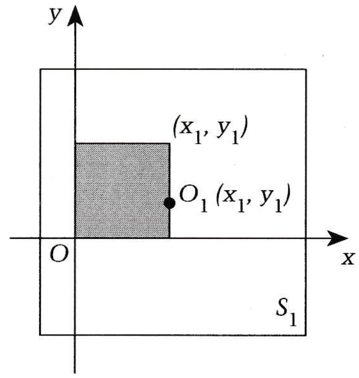


Figure 4

have proved the lemma.

Now we can describe the selection of the squares required in the problem. Let Q_1 be the largest of the given 1,000 squares (or one of the largest, if there are several—the same stipulation is implied in what follows). Further, let Q_2 be the largest of those of the given squares that have their centers outside Q_1 (if there are any); Q_3 the greatest of the squares with centers outside Q_1 and Q_2; and so on. Since there is a finite number of squares, this process ends up with a square Q_n such that all the centers are covered by the set of squares $Q_1, Q_2, \ldots, Q_n$. This set then satisfies the first requirement of the problem. To verify the second requirement, it will suffice to show that no five of the squares Q_i have a common point, for then no five can contain the same center of a square.

Suppose such a point exists. Then by the lemma above, the center of one of these five squares must belong to another square, which is impossible: by construction, for $i < j$, the center of Q_j is outside Q_i. On the other hand, we have constructed the sequence of squares $Q_1, Q_2, \ldots, Q_n$ so that Q_i is smaller than Q_j. Hence the center of Q_i cannot be inside Q_j either. This completes the proof.

M100

The required polygon can be drawn for $n = 3$, 4, and 6 (fig. 5, on the next page). Let's prove that for any other value of n this is impossible. First we make this obvious observation: if points A, B, and C lie on the lines of the given grid, then the point D such that the vector

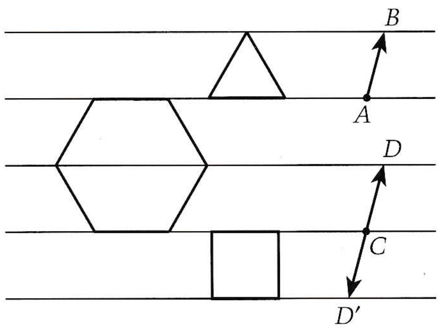


Figure 5

$\overrightarrow{CD}$ is equal to $\overrightarrow{AB}$ or $-\overrightarrow{AB}$ also lies on one of the lines (fig. 5). Now take an arbitrary regular n-gon $A_1A_2\ldots A_n$ with the vertices on the grid and a point O on one of the lines, and draw the vectors equal to $\overrightarrow{A_1B_2}$, $\overrightarrow{A_2B_3}$, ..., $\overrightarrow{A_nB_1}$ from O. The vectors are all equal in length, and the angles between consecutive pairs of vectors are all equal (since each is equal to an exterior angle of the original polygon). Therefore, their endpoints B_1, B_2, ..., B_n form another regular n-gon on the grid. Denote by $k = B_1B_2/A_1A_2$ the factor of similarity of these polygons. For any $n \geq 7$ this factor is less than 1, because $k = B_1B_2/OB_1$ (fig. 6), and in the triangle OB_1B_2 the angle B_1OB_2 is the smallest (since it's smaller than $360°/6 = 60°$). Therefore, repeating this construction an appropriate number of times, we can obtain a polygon $X_1X_2\ldots X_n$ whose side length $k^m \cdot A_1A_2$ is as small as we wish—for instance, smaller than the distance between the lines of the grid. But this is impossible, because any polygon on the grid has a side with the endpoints on different lines, and this side, of course, can't be shorter than the spacing of the grid.

In the case $n = 5$ this proof must be modified. In the first step we draw ten vectors equal to $\pm\overrightarrow{A_1B_2}$, $\pm\overrightarrow{A_2B_3}$, ..., $\overrightarrow{A_5B_1}$ from point O, thus creating a regular *decagon* inscribed in our grid. And this was shown to be impossible. (V. Dubrovsky)

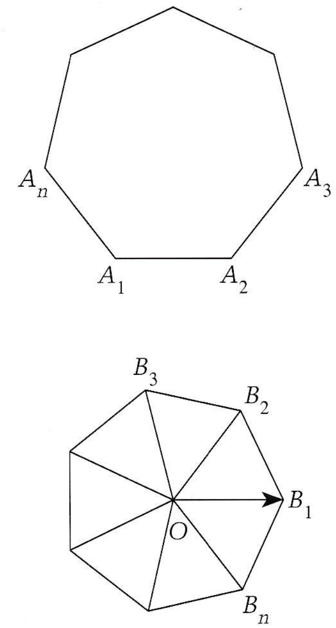


Figure 6

Physics

P96

Let a unit volume of air contain ρ snowflakes. Then

$$\left.\begin{aligned} N_1 &= \rho S(v + v_x) \\ N_2 &= \rho S(v - v_x) \end{aligned}\right\},$$

where v_x is the velocity of the wind in the skier's direction. This yields

$$\rho = \frac{N_1 + N_2}{2vS}.$$

Visibility can be estimated by calculating the average length L of a cylinder with a cross-sectional area A that contains one snowflake, where A is equal to the area of a snowflake. The volume of this cylinder times the density of snowflakes must equal 1:

$$LA\rho = 1.$$

Solving for L and approximating A by d^2 gives us

$$\begin{aligned} L &= \frac{1}{\rho d^2} \\ &= \frac{2vS}{(N_1 + N_2)d^2} \\ &\cong 200 \text{ m}. \end{aligned}$$

P97

We can imagine deforming the array of resistors in figure 7a (on the facing page) to the topological equivalent network shown in figure 7b. We can now see that there is no current flow in the plane of the hexagon *CDEFGH*, since these corners are all at the same potential. Therefore, the network is reduced to seven paths—the direct path *AB* containing one resistor and six parallel paths containing two resistors each, yielding an equivalent resistance of $R/4$.

P98

Even though figure 8 (on the next page) has been drawn with the incident ray at a fair distance from the optic axis for clarity, we need to remember that we only consider rays that are very close to the optic axis. We'll assume that we can ignore the thickness of the lens and use the small-angle approximation that $\sin\theta \cong \tan\theta \cong \theta$.

At the first surface, Snell's law tells us that

$$n_w \sin\alpha = n \sin\beta,$$

or

$$n_w\alpha \cong n\beta.$$

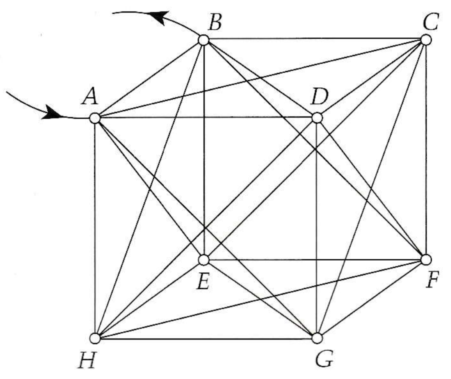
Similarly, at the second surface we have

$$n\gamma \cong \delta,$$

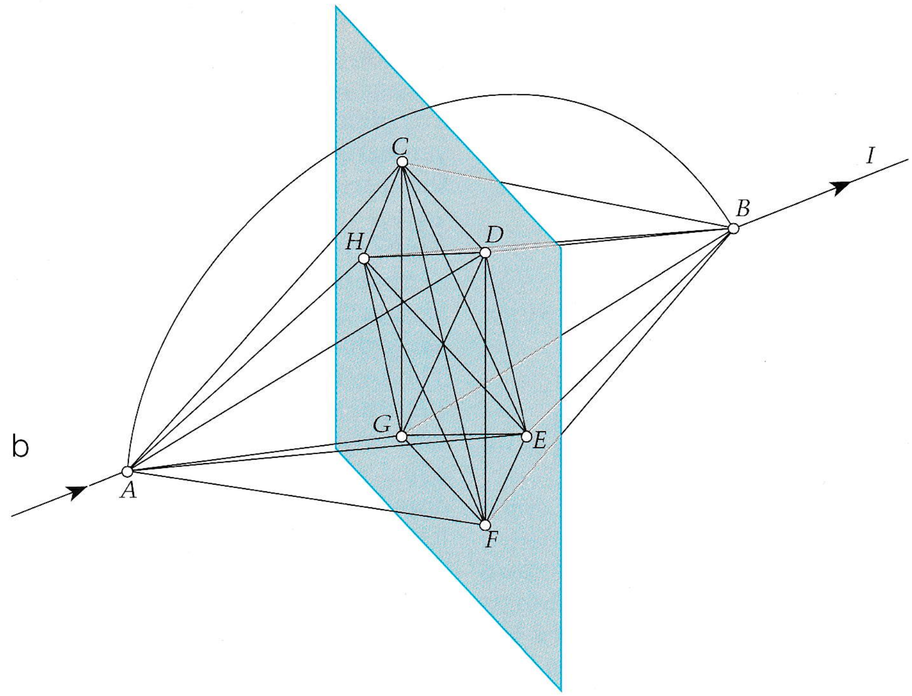
where $\gamma = \alpha - \beta$. From the figure we see that

$$\tan\delta \cong \frac{h_1}{F} \cong \frac{h}{F} \cong \delta.$$

Therefore, $F \cong h/\delta$. But $h \cong R\alpha$ and $\delta \cong n\gamma \cong n\alpha - n\beta \cong n\alpha - n_w\alpha$. Putting this all together, we find that



a



b

Figure 7

$$F \cong \frac{R}{n - n_w} = 15\ \text{cm}.$$

P99

It's clear from the symmetry of the diagram that the electrostatic energy in the initial and final states is the same, so the dissipated energy is equal to the work performed by the battery. In the initial state (fig. 9a), we have $q_1 = q_2 = CV_1$ and $q_3 = CV_3$. But conservation of charge requires that $q_1 + q_2 = q_3$. Therefore, $V_3 = 2V_1$. We also know that $V_1 + V_3 = \mathcal{E}$, which means that $V_1 = \frac{1}{3}\mathcal{E}$ and $V_3 = \frac{2}{3}\mathcal{E}$. Thus, $q_1 = q_2 = \frac{1}{3}\mathcal{E}C$ and $q_3 = \frac{2}{3}\mathcal{E}C$. In the final state (fig. 9b), we can write down that $q_1' = \frac{2}{3}\mathcal{E}C$ and $q_2' = q_3' = \frac{1}{3}\mathcal{E}C$ from the symmetry of the two figures.

To get from the initial state to the final state, we need to take a charge $\Delta q = q_1' - q_1$ from capacitor 3 through the battery to capacitor 1. Therefore, the battery must perform work $W = \mathcal{E}\Delta q = \frac{1}{3}C\mathcal{E}^2$.

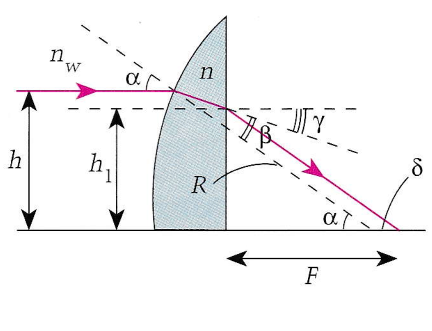


Figure 8

P100

There are two basic techniques for finding the period of small oscillations. One begins with Newton's second law of motion and the other with the statement of the conservation of mechanical energy. Let's use the latter technique.

Rotation of the ring about the vertical axis through an angle ϕ from its equilibrium position results in a horizontal displacement of the lower end of the string

$$\Delta x = r\sin\phi = \frac{L}{2}\sin\phi.$$

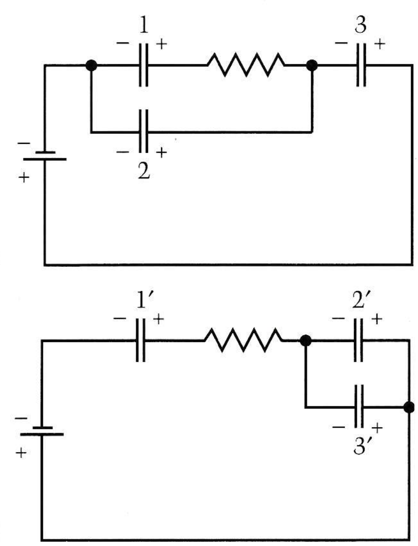


Figure 9

For small angles we get

$$\Delta x \cong \frac{L\phi}{2}.$$

Using the Pythagorean theorem we find that the vertical distance from the support to the string is shortened to

$$L' = \sqrt{L^2 - \frac{L^2\phi^2}{4}} = L\left(1 - \frac{\phi^2}{8}\right),$$

where we have used the binomial expansion to get rid of the square root. This means that the ring rises a distance

$$h = L\frac{\phi^2}{8},$$

and the gravitational potential energy of the ring is given by

$$GPE = mgh = mgL\frac{\phi^2}{8}.$$

The kinetic energy of the ring is given by

$$KE = \tfrac{1}{2}mv^2 = \tfrac{1}{2}mr^2\omega^2 = mL^2\frac{\omega^2}{8},$$

where we have ignored the kinetic energy associated with the vertical motion.

The period of oscillation is given by the square root of the ratio of the coefficient of the velocity squared to that of the coordinate squared. Therefore,

$$T = 2\pi\sqrt{\frac{L}{g}},$$

which is the same expression we get for a simple pendulum.

Brainteasers

B96

If the first digit of the unknown number were 5, the second digit couldn't be 4 (from 543) and the third digit couldn't be 2 (from 562). So the first digit would have to be 1 (from 142). Similarly, the second digit isn't 4. Therefore, from 543, we see that the third digit is 3; then, from 142, that the first digit is 1; and, from 562, that the second is 6. This gives us the answer: 163.

B97

A little experimentation will show that I must be closer to Entropy than to Tesseract. Between the two signs I've driven 150 ents − 10 ents = 140 ents, and also 110 tesses − 26 tesses = 84 tesses. Equating these two, we find that 1 ent = 3/5 tess. It's not hard to find, then, that the distance from Entropy to Tesseract is 20 tesses. Suppose that at the point we seek I am x tesses from Tesseract and also x ents from entropy. Then, measured in tesses,

$$x + \frac{3}{5}x = 20,$$

so x = 12.5 (tesses). This is the required position.

B98

Consider the 32 squares in the odd horizontal rows (the first, third, fifth, and seventh) of the chessboard. Each horizontal domino covers two or none of them, and each vertical domino covers exactly one of these squares. So the horizontal dominoes cover an even number n of these squares, and therefore the number of the remaining squares, $32 - n$, is also even. But it's equal to the number of vertical dominoes, which means that the answer to the question is yes.

B99

Suppose first that diagonal AC bisects BD (see figure 10). We will show that AC bisects midline MN. The key is to note that if AC bisects BD, then the (perpendicular) distances from D and B to line AC are equal. (This can be proven, for example, using congruent triangles.) This means that area(ADC) = area(ABC). Conversely, if area(ADC) = area(ABC), a similar argument shows that AC bisects BD. So AC bisects BD if and only if it divides the area of the quadrilateral in half.

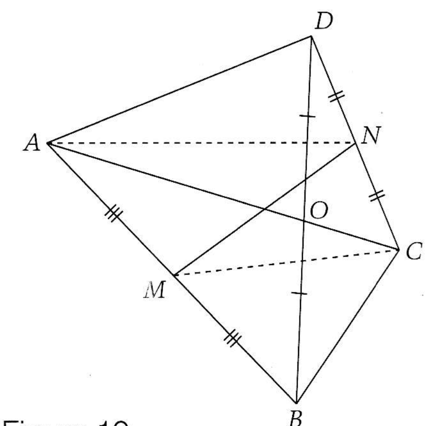


Figure 10

We now show that this same condition is both necessary and sufficient for AC to bisect a midline. For suppose area(ADC) = area(ABC). Then, since area(AMC) = ½area(ABC) and area(ANC) = ½area(ADC), we have area(AMC) = area(ANC), so AC bisects diagonal MN of quadrilateral $ANCM$. Another such argument shows that if AC bisects MN, then area(ADC) = area(ABC). Hence the conditions that AC bisects each of the three segments in the problem statement are all equivalent to the statement that AC bisects the area of the quadrilateral.

B100

The path is $3 \cdot 29 + 13 = 100$.

Bushels of pairs

Here is our answer to problem 7.

If a number x, $0 \le x \le 1$, is written as $0.x_1x_2x_3\ldots$ in ternary notation, we can locate it on the number axis as follows. Divide the segment [0, 1] into three equal parts and choose the left, middle, or right third, if x_1 = 0, 1, or 2, respectively (in each of the three cases x is represented as $0 + x'$, $1/3 + x'$, or $2/3 + x'$, where $x' = 0.0x_2x_3\ldots$ lies between 0 and 1/3, respectively). Then divide the chosen third of [0, 1] into three equal parts again and choose one of these parts according to the value of x_2 by the same rule. The chosen segment (which is a ninth of [0, 1]) is again trisected, and so on. Thus we get an infinite sequence of nested segments whose lengths make up a geometric sequence $1/3^n$, and the point x is their unique common point.

Suppose $x_n = 1$, and $x_i \ne 1$ for all $i < n$. Then, by the definition of $y = C(x)$, in binary notation, $y = 0.y_1\ldots y_{n-1}100\ldots$, where $y_i = 0$ if $x_i = 0$, and $y_i = 1$ if $x_i = 2$ for $1 \le i \le n-1$, no matter what the digits x_j for $j > n$ are. This means that $C(x)$ is constant on all the "middle thirds" that arise in our trisecting process. As for the value of $C(x)$, it's clear that as long as we don't come across a 1 moving along the sequence x_1, x_2, ..., the value of y increases by $1/2^i$ every time we choose the right third in the ith step (that is, when $x_i = 2$). This allows us to sketch the graph (a part of it, of course) on the "middle thirds" (fig. 11).

It's hard to believe that this function, the graph of which seems to consist only of horizontal segments, nevertheless takes any value in [0, 1]. (Notice that the sum of their lengths $1/3 + 2/9 + 4/27 + \ldots + (1/3)(2/3)^n + \ldots$ is equal to $\frac{1}{3}[1/(1 - \frac{2}{3})] = 1$—that is, to the length of the entire segment [0, 1]!) Indeed, the value $y = 0.y_1y_2y_3\ldots$ (in binary notation) is taken at the point $x = 0.x_1x_2x_3\ldots$ (in ternary notation), where $x_i = 2y_i$.

We now show that the function $C(x)$ is nondecreasing. This can be proven by a close examination of the

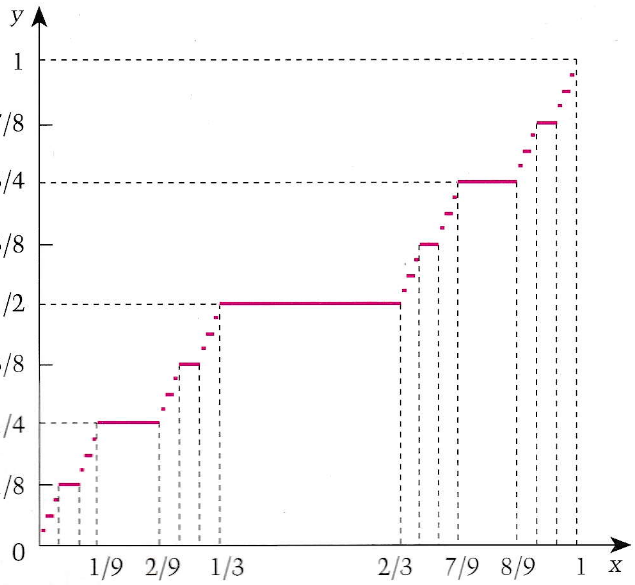


Figure 11

construction of the function. Suppose $x < x'$, and the ternary expansions of x and x' first differ in their nth digit. If $x_j = x_j' = 1$ for some $j < n$, then y and y' are identical, since both have digits 0 after the jth (binary) place. If there is no digit 1 to the left of x_n, things are more complicated. If $x_n = 0$ and $x_n' = 1$ or 2, then $y_n = 0$ and $y_n' = 1$. Even if $y_k = 1$ for all $k > n$, it is still true that $y \le y'$. If $x_n = 1$, then $x_n' = 2$. Here $y_n = 1$ and $y_k = 0$ for $k > n$. But $y_n' = 1$, and so $y' \ge y$. So $C(x)$ is nondecreasing.

Readers familiar with a rigorous definition of a continuous function will prove without difficulty that a nondecreasing function that maps an interval onto an interval—in particular, $C(x)$—is continuous. This function serves as a counterexample to a number of statements about functions that look quite plausible but are, in fact, wrong. Its graph is called *Cantor's staircase* in honor of the creator of modern set theory, Georg Cantor.

Ones up front

(Solutions supplied by the editor)

1. The probability in question is 1/3. If a_n is the number of multiples of three that don't exceed n, then $n/3 - 1 < a_n \le n/3$. So $1/3 - 1/n \le a_n/n \le 1/3$, and $a_n/n \to 1/3$ as $n \to \infty$.

2. Writing out the decimal expansion of 161/222 = 0.7252252252..., we see that every third digit is five, and every $(2 + 3k)$th and every $(4 + 3k)$th digits are twos. So the probability to choose a five is 1/3, and that of choosing a two is 2/3. Note that, in the limit, the presence of the initial digit 7 does not affect the answer.

3. If $1, 2^2, 3^2, \ldots, k^2$ are all the perfect squares not exceeding n, then their number $a_n = k$ is not greater than $\sqrt{n}$, so $a_n/n \le 1/\sqrt{n} \to 0$ as $n \to \infty$. A random integer is a square with zero probability.

4. In the sequence $1, 2, 2^2, 2^3, \ldots$, every number beginning with 1 is followed by a number beginning with 2 or 3, and every number beginning with 2 or 3 is preceded by a number beginning with 1. So if $a_n(q)$ is the number of powers of two not exceeding 2^n and beginning with q, then $a_n(1) = a_n(2) + a_n(3)$ or $a_n(1) = a_n(2) + a_n(3) - 1$, depending on the initial digit of 2^n. In either case, dividing by n and letting n tend to infinity, we get $p_1 = p_2 + p_3$. The other relations are proven similarly.

5. In the notation of the previous solution, $a_n(1) + a_n(2) + \ldots + a_n(9) = 1$. Therefore, $p_1 + p_2 + \ldots + p_9 = 1$, and, by the relations of problem 4,

$$\begin{aligned} 1 &= p_1 + (p_2 + p_3) + (p_4 + p_5) \\ &\quad + (p_6 + p_7) + (p_8 + p_9) \\ &= 2p_1 + p_2 + p_3 + p_4 \\ &= 3p_1 + p_4. \end{aligned}$$

So $p_4 = 1 - 3p_1 = 1 - 3 \log 2$.

6. Suppose $\log l$ is a rational number m/n. Then $l = 10^{m/n}$, or $l^n = 10^m$. This is possible only when l is a power of ten.

7. It was shown in the article that $p_q = \log (q + 1)/q$. It follows that

$$\begin{aligned} p_{2k} + p_{2k+1} &= \log\frac{2k+1}{2k} + \log\frac{2k+2}{2k+1} \\ &= \log\frac{2k+2}{2k} \\ &= \log\frac{k+1}{k} \\ &= p_k \end{aligned}$$

for $k = 1, 2, 3, 4$.

8. The number 2^n begins with the digits 1000 if $10000\ldots0 \le 2^n < 10010\ldots0$ (where the numbers on both sides have the same number of digits). This can be rewritten as $\{n \log 2\} < \log 1.001$, and, by the Fractional Parts Theorem, the unknown probability is equal to log 1.001.

9. Let Q be the number written as $q_1 q_2 \ldots q_r$ in decimal notation. Then the powers of l in the statement must satisfy the inequalities $Q \cdot 10^m \le l^n < (Q + 1) \cdot 10^m$. Then an argument like that in the article, or in the previous solution, leads to the answer $p = \log [(Q + 1)/Q]$. Since $(Q + 1)/Q > 1$, this probability is positive, so any combination of digits appears at the beginning of a power of l sooner or later.

10. (a) By the previous problem,

$$\begin{aligned} p_0^{(2)} &= \log\frac{11}{10} + \log\frac{21}{20} + \cdots + \log\frac{91}{90} \\ &= \log(11 \cdot 21 \cdot \cdots \cdot 91) - \log(9! \cdot 10^9) \\ &= \log(11 \cdot 21 \cdot \cdots \cdot 91) - \log(9!) - 9. \end{aligned}$$

The probability $p_q^{(k)}$ is equal to the sum of the probabilities that the k initial digits of a power of l form a number $10i + q$, where i runs through all $(k - 1)$-digit numbers—from 10^{k-2} to $10^{k-1} - 1$. By problem 9, these probabilities are equal to $\log[1 + 1/(10i + q)]$, respectively.

11. Let $s^{(k)}$ be the sum of $\log(1 + 1/10i)$ over all i from 10^{k-2} to $10^{k-1} - 1$. The hints in the problem statement are the results of various algebraic manipulations, which are left to the reader. Using these, and the formula in problem 10(b), we obtain the following expression:

$$0 < s^{(k)} - p_q^{(k)} < \frac{q}{100\ln 10}\sum\left(\frac{1}{i-1} - \frac{1}{i}\right)$$

$$= \frac{q}{100\ln 10}\left(\frac{1}{10^{k-2}-1} - \frac{1}{10^{k-1}-1}\right)$$

$$\to 0$$

as $k \to \infty$ (the sum here is taken over the values of i specified above and readily "telescopes"). This is true for all $q = 0, 1, \ldots, 9$, so using $p_0^{(k)} + p_1^{(k)} + \ldots + p_9^{(k)} = 1$, we get $10s^{(k)} - 1 = (s^{(k)} - p_0^{(k)}) + (s^{(k)} - p_1^{(k)}) + \ldots + (s^{(k)} - p_9^{(k)}) \to 0$, which means that $s^{(k)} \to 0.1$. The first inequality above then assures us that $p_q^{(k)} \to 0.1$ as $k \to \infty$ as well.

12. In the number system to the base $b > 1$, $p_q^{(k)}$ is expressed by the formula obtained from the "decimal" formula in problem 10(b) by replacing the number 10 with b and replacing log with $\log_b$. The proof is practically the same.

13. If A_0 is the initial point on the circumference, then the length l_n of the arc A_0A_n (measured in the direction in which the unit arcs are marked off) is equal to $n - 2\pi k$, where k is a integer such that $0 \le n - 2\pi k < 2\pi$—that is, $l_n = 2\pi\{n/2\pi\}$. Point a_n hits an arc Q of length h if the number $x_n = l_n/2\pi$ hits a corresponding interval of length $h/2\pi$. This happens with a positive probability according to the Fractional Parts Theorem (with $\beta = 0$, $\alpha = 1/2\pi$).

14. Taking $x = \pi/2 + 2\pi k$ with an integer k, we get

$$f\left(\frac{\pi}{2} + 2\pi k\right)$$

$$= \sin\left(\frac{\pi}{2} + 2\pi k\right) + \left(\frac{\pi\alpha}{2} + 2\pi k\alpha\right)$$

$$= 1 + \sin\left[2\pi\left(\frac{\alpha}{4} + k\alpha\right)\right]$$

$$= 1 + \sin\left(2\pi\left\{\frac{\alpha}{4} + k\alpha\right\}\right).$$

By the Fractional Parts Theorem, $\{\alpha/4 + k\alpha\}$ is arbitrarily close to 1/4 for a certain value of k, so $\sin(2\pi\{\alpha/4 + k\alpha\})$ can be arbitrarily close to $\sin(\pi/2) = 1$. Thus, the function $f(x)$ can take values arbitrary close to 2. On the other hand, $f(x) \le 2$ for all x. If $f(x)$ were a periodic function with a period $T > 0$, it would take all its possible values on, say, the segment $[0, T]$. Being continuous, it must take its maximum value $M \le 2$ at some point in this segment. Since for any $a < 2$ there is a point x such that $f(x) > a$, M must be exactly equal to 2. But $f(x) = 2$ only if $\sin x = 1$ and $\sin \alpha x = 1$—that is, $x = \pi/2 + 2\pi k$ and $\alpha x = \pi/2 + 2\pi n$ for certain integers n and k. It follows that $\alpha = (4n + 1)/(4k + 1)$, which contradicts the assumption that α is irrational. So f is nonperiodic.

15. The answer is yes. Putting $\alpha = d_1/d_2$ we can write the condition that the absolute value of the difference between the kth term of the first sequence and the nth term of the second sequence is less than 10^{-6} as

$$|(a_2 - a_1) + d_2(n - k\alpha)| < 10^{-6},$$

or

$$-\frac{10^{-6}}{d_2} < n + a - k\alpha < \frac{10^{-6}}{d_2},$$

where $a = (a_2 - a_1)/d_2$. Applying the Fractional Parts Theorem to the sequence $a - k\alpha$, $k = 1, 2, \ldots$ and the interval $I = [0, 10^{-6}/d_2]$, we see that for some k (even for infinitely many values of k) $\{a - k\alpha\}$ hits the interval I.

Choose any such value of k, and let $-[a - k\alpha]$, where $[x]$ denotes the greatest integer not exceeding x. For this choice of n and k, we have $n + a - k\alpha = \{a - k\alpha\}$, and this number is between $-10^{-6}/d_2$ and $10^{-6}/d_2$ (in fact, it is between 0 and $10^{-6}/d_2$). Finally, we must be sure that n is positive. To do this, we must choose k so huge that $a - k\alpha$ is negative. The Fractional Parts Theorem assures us of infinitely many possible values of the positive integer k, so we can choose k as large as we need. This completes the proof.

16. The equation for the given line has the form $y = ax + b$, where $a = \tan\phi$ is irrational. Fix $\varepsilon > 0$. It suffices to show that there exists a pair of integers (m, n) such that $|am + b - n| < \varepsilon$, because this inequality means that the line intersects the vertical diameter of the "tree" centered at (m, n). By the Fractional Parts Theorem, $\{am + b\} < \varepsilon$ for some integer m. Using such a value for m, and letting $n = [am + b]$, it's not hard to see that the inequality we need is satisfied.

It will be useful to note that in fact we've proven that even the half-line defined by $y = ax + b$, $x \ge 1$, intersects the forest. Of course, this remains true for any ray on the line $y = ax + b$. If the ray is defined by restricting the line to values of x such that $x \ge x_0$ or $x \le x_0$, we'd only have to consider the sequence $am + b$ with $m \ge x_0$, or $-am + b$ with $m \ge -x_0$, respectively.

17. The statement of this problem can be proven in much the same way as its one-dimensional version in the article, though this would require a considerably more involved technique. Here we give a proof based on a different idea; some of its details are omitted but they are easy to restore.

We'll use the following notations: if a vector $\mathbf{v}$ is drawn from a point x, then its endpoint will be denoted by $x + \mathbf{v}$; $M + \mathbf{v}$ for any figure M is the figure obtained from M by translation along vector $\mathbf{v}$. And any vector (or point) with integer coordinates will be called simply an *integer vector* (or *point*).

The proof consists of several steps.

(1) Let $\mathbf{f}_n = \{n\mathbf{a}\}$, $F_n = O + \mathbf{f}_n$, where $\mathbf{a}$ is the vector from the statement of the problem, and let O be the origin. Then *the set of all points F_n is everywhere dense in the square S*—that is, *for any point X in S and any $\varepsilon > 0$ there is a point F_q such that $F_qX < \varepsilon$*. In vector notation, this means that $|\mathbf{x} - \mathbf{f}_q| < \varepsilon$, where $\mathbf{x} = \overrightarrow{OX}$.[1]

To prove this, we first divide the square S into small equal square "pigeonholes," each with diagonal $< \varepsilon$. We choose n so large that we can find two of the points, F_i and F_j, such

[1]The one-dimensional version of this statement is known as Kronecker's theorem.

Since ε is an arbitrary positive number, this means that $p(M) \geq \text{area}(M)$. Similarly considering the boxes B_i that cover M, we can prove that $p(M) \leq \text{area}(M)$.

Strictly speaking, this proof is still incomplete: we've made an implicit assumption that the probabilities $p(Q)$, $p(M)$, and so on, actually do exist. But it's not difficult to modify the argument so that it yields both the existence and the value of $p(M)$. All the probabilities p for the entire infinite sequence $\mathbf{a}_n$ in the estimates above can be replaced with the corresponding probabilities p_N for a finite segment of this sequence $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_N$ if we add to the bounds appropriate terms of the form c/N, where c does not depend on N. Thus, we can show that for any $\varepsilon > 0$, $|p_N(M) - \text{area}(M)| < 2\varepsilon + c/N < 3\varepsilon$, if N is large enough. But this just means that $p_N(M)$ has a limit, and it's equal to the area of M.

18. Draw coordinate axes parallel to the sides of the squares of our chessboard with the origin O at a vertex of one of the squares and a scale such that a unit segment on each of the axes is twice the side length of a square. Then any two points whose corresponding coordinates differ by an integer number of such units lie in squares of the same color. In particular, a point A and a point B such that vector $\overrightarrow{OB}$ is the fractional part of vector $\overrightarrow{OA}$ are always the same color. So the probability that our flea F hits a black square equals the probability that the fractional part of the vector $\overrightarrow{OF}$ hits (in the sense of problem 17) a black square. Since two of the four chessboard squares that make up a unit square (with respect to our coordinates) are black, this probability equals $1/2 > 0$.

The answer to the second question is no. Consider a flea that starts at the bottom right corner of one of the white squares and jumps such that $x = y = \sqrt{2}$. It will always stay on the extension of the diagonal of the initial white square drawn from the starting point, so it will always hit white squares. But this doesn't satisfy the requirement that y/x be irrational. However, we can imagine another flea that starts at the same point, but jumps with such that $x = \sqrt{2}$, $y = 2 + \sqrt{2}$. For this insect all three numbers x, y, and $y/x = \sqrt{2} + 2$ are irrational. It always lands in the same column, but an even number of squares apart from the first flea—therefore, on the same color (white) as the first one.

19. The inequalities in the statement will be true if, for a sufficiently small $\varepsilon > 0$ and some integer k,

$$\left| n\lambda_i - \left(\frac{\pi}{2} + 2\pi k\right)\right| < 2\pi\varepsilon, \quad i = 1, 2.$$

This inequality can be rewritten as $|(na_i - k) - 1/4| < \varepsilon$ $(i = 1, 2)$, where $a_i = \lambda_i/2\pi$, or as $|\{na_i\} - 1/4| < \varepsilon$ $(i = 1, 2)$. The two-dimensional generalization of the Fractional Parts Theorem (problem 17) applied to the vectors $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = 0$, and the polygon (square) M defined as the set of points (x, y) such that $|x - 1/4| < \varepsilon$, $|y - 1/4| < \varepsilon$, shows that our inequalities in n have infinitely many solutions.

CONTINUED ON INSIDE BACK COVER

Index of Advertisers

Tricky rearrangements

Variations on a rolling permutational theme

by Vladimir Dubrovsky

IN THIS INSTALLMENT OF THE Toy Store I'll continue introducing to you the family of "rolling block" puzzles and games. The last issue contained a rather extensive treatment of the rolling pyramids. This time, though, I'll offer neither solutions nor any substantial hints—only what the puzzles are like and what you're supposed to do with them. In fact, one of the two puzzles below has no rolling parts—it should rather be pigeonholed as "rearrangements on the triangular grid," which links it with the simplified, flat versions of the second puzzle and with the rolling pyramids as well. In an upcoming issue of *Quantum*—after (I hope) you've meditated on these puzzles and come up with your own solutions—we'll return to them and discuss their underlying mathematics. Then you'll see that they have much more in common with each other and with the pyramids than is apparent at first glance.

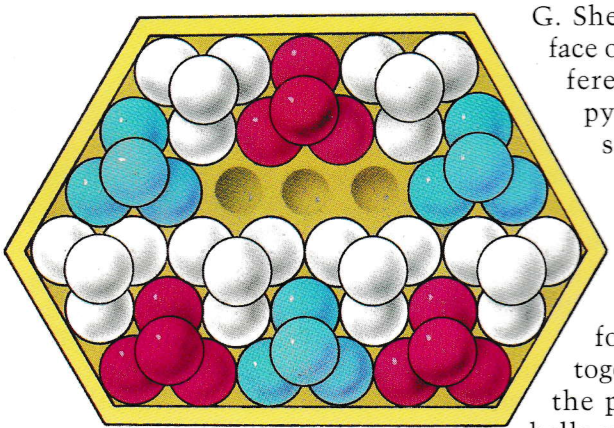


Figure 1

I'd like to thank Anatoly Kalinin, a Moscow engineer who has gathered a wonderful collection of intellectual toys and games. A great many of the items in his collection were sent by their creators from all over the former USSR and are little known in other countries. Among them are the first two puzzles below, as well as the rolling pyramids from the previous issue.

Cannonball pyramids

The puzzle shown in figure 1 was designed by the authors of the rolling pyramids, A. Dryomov and G. Shevtsova. And, on the face of it, it's not much different from the rolling pyramids. It, too, consists of "pyramids" that can be rolled in a hexagonal box. But these are special pyramids—each of them is made from four small balls glued together—they look like the pyramids of cannonballs you see near old cannons in museums. (By the way, if you're going to make this puzzle yourself, notice the round holes in the bottom of the box that prevent the pyramids from slipping when they're rolled. A good material for the bottom is styrofoam.) The number of pyramids (12), their coloring, and the shape of the box are also different, but it's the shape of the pyramid that gives this puzzle its essentially new quality. In the rolling pyramids puzzle we could divide the box into triangles such that each pyramid occupied exactly one triangle in any possible arrangement, and one triangle was left free. With the cannonball pyramids such a division is impossible, because after rolling a cannonball pyramid it still occupies two old spaces (holes) and only one new space. We also see in figures 1 and 2 that the empty holes can wander away from one another and all over the box.

The task you have to accomplish is the same as in all puzzles of this kind: *one given arrangement of pyramids must be transformed into another by a suitable sequence of moves (rolls)*. In particular, the authors offer the arrangement in figure 1 as the standard initial position and those in figure 2 as the target positions.

One problem with this and other original mechanical puzzles is that

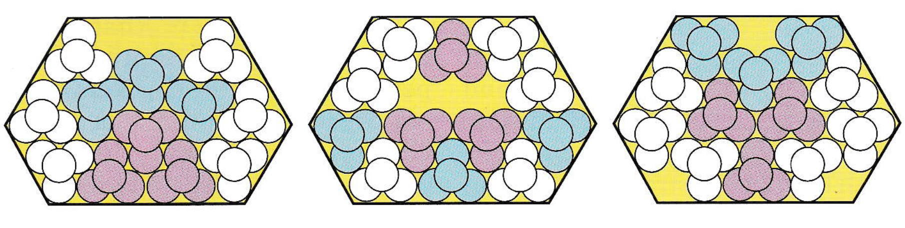


Figure 2

to play with them you have to make them, and this may require more work and time than you can spare. So here's a way out. Draw a triangular grid whose vertices make the same pattern as the holes in the bottom of the box, and replace each pyramid with three round chips of the same color placed at the vertices of the corresponding triangle of the grid. (We'll call such a triangle of chips a "triad.") Then any move (roll of a pyramid) will be represented by a jump of one of the chips of the corresponding triad over the other two chips onto the node of the grid immediately beyond them—if this node is free, of course (fig. 3). In order not to mix the chips from different triads of the same color, you

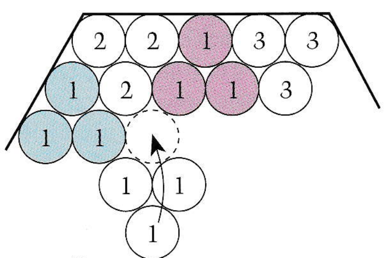


Figure 3

can assign different numbers to such triads and write these numbers on the chips. This is useful in solving the puzzle, too, because the numbers will allow you to follow the displacements of each particular triad.

The term "triad" here was borrowed from the name of the next puzzle, where checkers are also rearranged by moving "triangles of checkers."

Triads

The triad puzzle is shown in figure 4. As you see, it's quite simple—I mean, you don't need any special "equipment" to play with it: just six checkers, three of one color and three of the other. That doesn't mean it's easy to solve. It was created by Sergey Grabarchuk from Uzhgorod, a town in western Ukraine. He has invented a great number of ingenious puzzles and even wrote a book of recreational problems, *The Jar of Diamonds*, with his own illustrations. (We plan to acquaint you with some of them in our Toy Store in the future.)

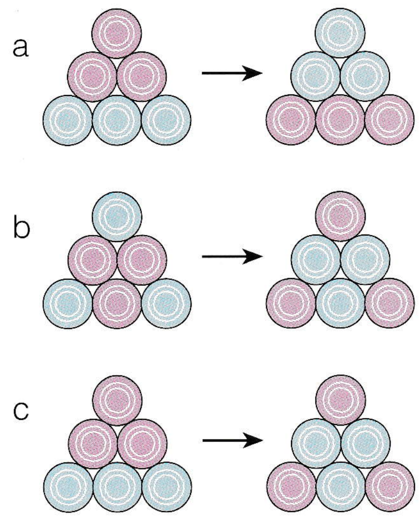


Figure 4

Take the six checkers and place them on a table so as to make the triangle shown in the left part of figure 4a. You're allowed to choose any three checkers touching each other (a triad) and slide this triad as one whole piece along the surface of the table parallel to itself (that is, without turning it) so that it finally joins the remaining checkers again at some new position. Then new triads are formed, and you can repeat a similar move with one of them. As you make your move, be sure the stationary checkers don't budge. Your task is to obtain the pattern shown in the right part of figure 4a.

This is the easiest problem of the three presented in figure 4. Figure 4b presents a more difficult problem, and the trickiest one of all is figure 4c. Try to find as short a solution for each as you can.

In these problems, the final big triangle of checkers might end up shifted with respect to its initial location. It's interesting to *find solutions that bring this triangle (as a whole) back to where it was at the start*, if such solutions exist.

Another problem that arises naturally is to examine all possible permutations of the checkers in the big triangle, assuming that all the checkers are marked differently (say, numbered from 1 to 6)—both with and without returning the triangle to the initial location. And, of course, all of these questions can be posed for the case of another number of checkers and the initial figures they form.

Tumbleweed

The cube is a shape that is as suited for rolling-block puzzles as the pyramid (regular tetrahedron, to be exact). Even more so, perhaps, because the cube has a property that the pyramid lacks: the orientation taken by the cube after rolling to a certain location *depends* on the route it took to get there. This makes the rolling-cube puzzles three-dimensional in essence: we can't just replace the cubes with flat chips, as we did with the pyramids. So this kind of puzzle deserves a separate treatment—here we'll take a look at only one possible application of rolling cubes.

It's a game rather than a puzzle, and it was designed by a 40-year-old Moscow professional artist, Andrey Korovin. Perhaps his bent for abstraction, which is apparent in his landscapes, is responsible for the game's mathematical flavor, and his artistic imagination suggested the name "Tumbleweed" for the game.

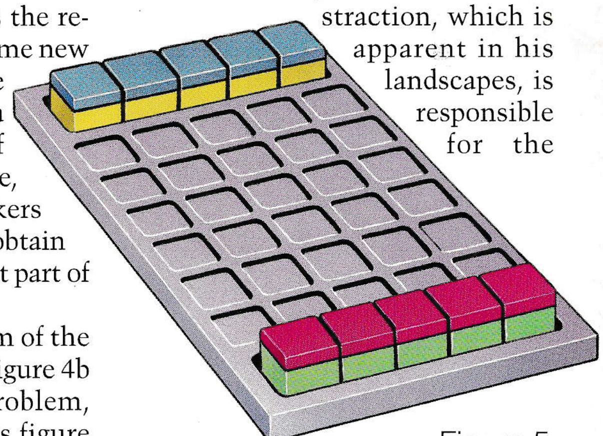


Figure 5

The rules of the game are simple. The playing area consists of a rectangular field measuring 5 × 6, and you need 10 cubes (fig. 5). Each player gets a set of five cubes of different colors (say, green and yellow). One face on each cube the *control face*—is painted a contrasting color (here, red and blue—see figure 6). Initially, the

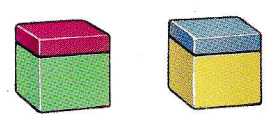


Figure 6

7	8	7	8	9
4	5	6	7	8
5	6	5	6	7
4	5	4	5	6
3	6	5	6	5
	3	4	5	4

Figure 7

cubes are placed like chess pieces along the short sides of the field, their control faces up. The players take turns rolling their cubes (to any free adjacent square), one cube at a time. Each cube must be brought over to one of the squares at the opposite side, its control face up. The player who does this first wins.

There are a few additional rules that prevent draws. We won't dwell on them here—you'll think up rules of this sort after you try to play a couple of times.

From the mathematical point of view, another game suggested by the author of Tumbleweed seems more interesting. A cube is placed, its control face up, on the shaded square of the table in figure 7. One player chooses a square with a number, and another player must find a sequence of rolls that consists of this number of moves and brings the cube to the chosen square (control face up, of course). A correct solution gives the first player a point. Then the players exchange roles.

See if you can find the required sequences of moves for all numbered 29 squares. Can you explain why each square is numbered as shown in figure 7? What other values can the numbers take so that the problem remains solvable? What numbers should be written in the table if the cube's faces were all colored differently and its final orientation must be the same as the initial one?

"ANSWERS, HINTS & SOLUTIONS"
CONTINUED FROM PAGE 62

Derivatives

1. (a)

Value of a	Number of roots
$\|a\| > 216$	1
$\|a\| = 216$	2
$216 > \|a\| > 88$	3
$\|a\| = 88$	4
$\|a\| < 88$	5

(b)

Value of a	Number of roots
$a < 0$	0
$a = 0$	1
$0 < a < 4e^{-2}$	3
$a = 4e^{-2}$	2
$a > 4e^{-2}$	1

2.

Value of a	Number of roots
$a < 0$	1
$0 \le a < e$	0
$a = e$	1
$e < a$	2

3. (a) $x = 1$ ($\ln x < x - 1$ for all $x \neq 1$, $x > 0$); (b) $x = 0$ ($\sin x > x - x^3/6$ for all $x \neq 0$).

4. (a)

Value of a	Number of roots
$a < -189$	0
$a = -189$	1
$-189 < a < -64$	2
$a = -64$	3
$-64 < a < 0$	4
$a = 0$	3
$0 < a$	2

(b)

Value of a	Number of roots
$a < 0$	1
$0 \le a < e$	0
$a = e$	1
$e < a$	2

(Note that this problem is reduced to exercise 2 by substituting $\ln x$ for x.)

5. $x = 1$ (point $x = 1$ is the minimum point of the left side of the equation—examine its derivative).

6. (a) $(x, y) = (1, 1)$; (b) $(x, y) = (e, \pi + 2\pi k)$, $k = 0, \pm1, \pm2, \ldots$; (c) $(x, y) = (0, \pi/2 + 2\pi k)$, $k = 0, \pm1, \pm2, \ldots$. (In equations (a) and (b) the minimum value of the function on one side of the equation is equal to the maximum value of the other side; the same applies to equation (c) after dividing by 2^x.)

7. (a) See figure 15: for the points (p, q) in the shaded area the equation has three roots; the white area, including the origin (0, 0), corresponds to one root; the curves $q = \pm 2p^{3/2}$, $p > 0$, correspond to two roots.

(b) See figure 16: the white area means there are no roots; the gray area—one root; the red area—two roots (for $p \le 0$ the equation is undefined); on the border lines $q = e \ln p$ and $p = 1$, there is one root; on the line $q = 0$, there are no roots, but at the point $(p, q) = (1, 0)$, there are infinitely many roots.

8. $e^\pi > \pi^e$. Hint: $e^\pi > \pi^e$ is equivalent to $\ln \pi/\pi < 1/e$, but $1/e$ is the maximum value of $(\ln x)/x$.

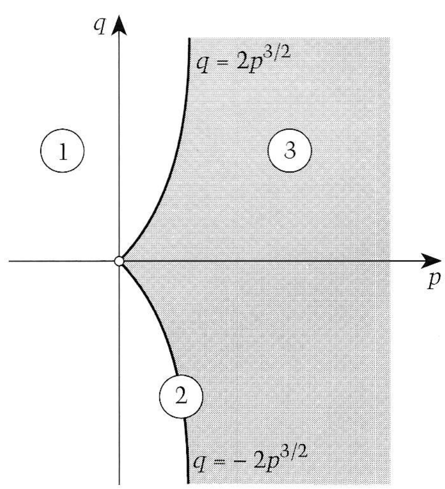


Figure 15

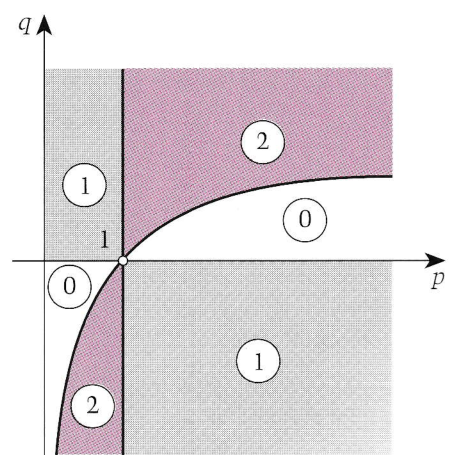


Figure 16

NEW OPTIONS IN MATHEMATICS

College of Arts and Sciences, Northwestern University

MENU

Mathematical
Experience for
Northwestern
Undergraduates

Northwestern University announces the formation of an expanded set of programs for selected students interested in mathematics and its related fields. MENU will be offered for the first time to entering students in the fall of 1993. If you have strong intellectual curiosity in this direction and seek a major university and the opportunity to work closely in a small and personalized setting with other students and professors, read on.

HISTORY

In 1976, Northwestern inaugurated the Integrated Science Program (ISP), which is a successful, one-of-a-kind effort to bring together mathematics and the sciences at the undergraduate level. This was followed by the creation of a parallel program in Mathematical Methods in the Social Sciences (MMSS). Enrollment in these programs is limited to approximately 30 students per year.

FEATURES OF MENU

MENU is a set of special concentrations for students with a strong interest in mathematics and its applications in the sciences and social sciences. MENU is designed to bring together students seeking to develop an active, hands-on approach to mathematics and encourage the exploration of advanced topics in special seminars with program faculty. MENU will provide smaller classes, more individualized advising and seminar programs for selected students wishing to concentrate in mathematics in a multidisciplinary setting. The principal entry route will be Mathematics B90, a three-quarter sequence covering the foundations of analysis. During the first year, students applying to MENU will indicate one of several possible routes: ISP* MMSS* and the new MENU programs in Mathematics, Mathematical Physics, Statistics, Computer Studies, and Decision Sciences.

WHY A SPECIAL PROGRAM?

At Northwestern University, we combine the strengths of a research institution with a small size to offer individualized programs at the undergraduate level, where students receive an in-depth approach that leads to superior preparation and more informed choices for graduate study or professional preparation. As a student interested in obtaining the best possible education, your personal interests are given the highest priority.

APPLICATION PROCEDURE

Applications will be accepted by the director of MENU during the Freshman Year from students who are enrolled in Mathematics B90. The specific MENU programs will begin with Sophomore Year. It is expected that MENU students will take some calculus in high school, including the BC and/or AB examinations administered by the CEEB Advanced Placement Program. Other specific questions can be administered by the Director by writing to: Mark Pinsky, Director of MENU, Kresge 324, Northwestern University, Evanston, IL 60208-2206

* High School students interested in entering ISP or MMSS must complete the relevant application concurrently with the Northwestern application in the last year of high school.

C A S

Northwestern University is an equal opportunity, affirmative action educator and employer.