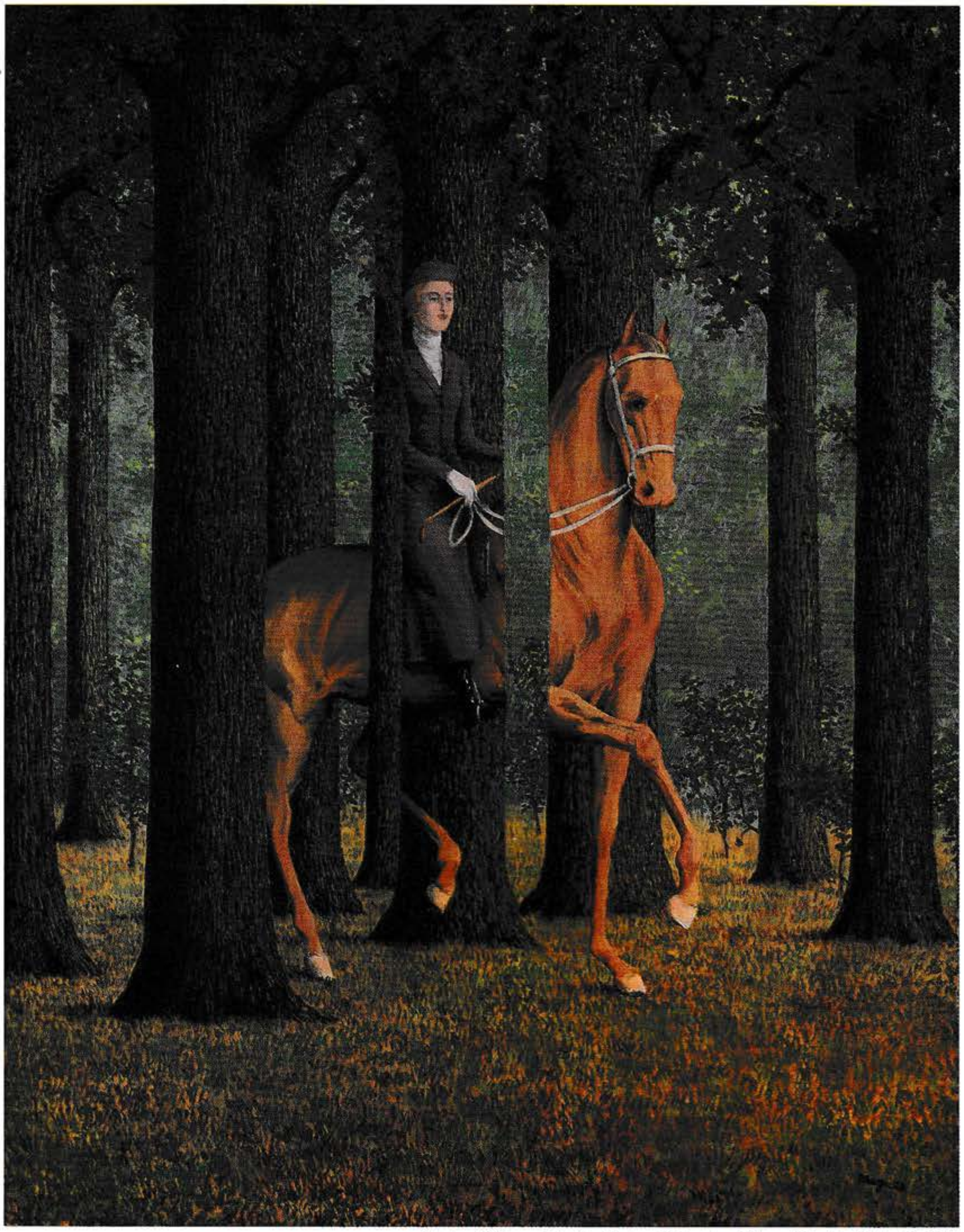
QUANTUM

MAY 1990

The student magazine of math and science



National Gallery of Art, Washington (Collection of Mr. and Mrs. Paul Mellon) © NGA

The Blank Signature (1965) by René Magritte

The Belgian artist René Magritte (1898–1967) was a master of the visual pun and optical paradox. His fellow Surrealist Salvador Dalí called him "one of the most ambiguous painters of our time."

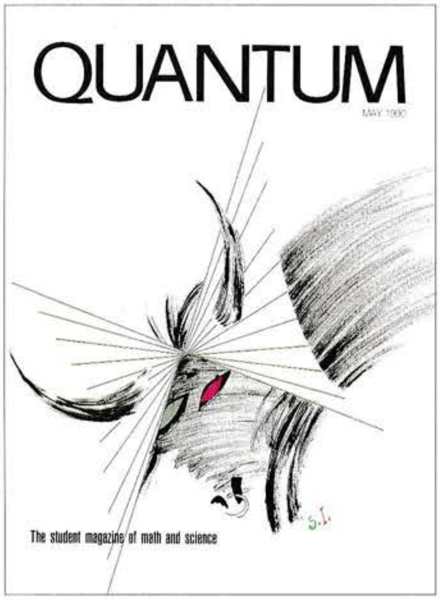
Surrealism was a movement in art and literature that flourished in the 1920s and 1930s. There was no single Surrealist style, as can be seen by looking at the work of such artists as Max Ernst, Joan Miró, and Paul Klee. Magritte's approach was to present minutely detailed, realistic depictions of an impossible or irrational world. By provoking a sympathetic response, the painting induces the viewer to acknowledge the inherent "sense" of the irrational and logically inexplicable.

Perhaps you've had a similar feeling when confronted with the square root of –1 or non-Euclidean geometry. The way Magritte plays with the notions of "subject" and "background" in this painting is not unlike the way scientists and mathematicians play with assumptions and variables, producing valid concepts that fly in the face of common sense.

Even after you've figured out the "trick" in any particular painting by Magritte, it remains unsettling and uncanny and, if you like that sort of thing, weirdly charming.

QUANTUM

MAY 1990



Cover art by Sergey Ivanov

The surface pierced by the furious bull's horns is a hyperbolic paraboloid. It's interesting that this curved shape can be constructed from two interlaced sets of straight lines: any two lines belonging to one set are askew, while any two lines from different sets always intersect. Also, each point of the surface belongs to just one line in each set.

Here's a more precise description. Imagine two points moving at the same speed along two skewed lines. The line joining them sweeps out a hyperbolic paraboloid. In algebraic terms, the equation for this surface with respect to the appropriate space coordinates is $z = xy$.

Our bull in his rage is unable to tell green from red, just as persons suffering from color blindness have trouble with these two colors. Strangely enough, this medical condition has something in common with our remarkable surface—see "The Geometry of Population Genetics," which starts on page 24.

FEATURES

4 Where Biology Meets Physics
How do we breathe?
by K. Y. Bogdanov

8 Where Middling Values Reign
Considerations of continuity
by S. L. Tabachnikov

16 Where Scales Are Weighed
Temperature, heat, and thermometers
by A. Kikoyin

24 Where Biology Meets Mathematics
The geometry of population genetics
by I.M. Yaglom

DEPARTMENTS

3 Publisher's Page

13 Brainteasers

14 Getting to Know...
The chemical elements

22 How Do You Figure?

30 In Your Head
Ballpark estimates

32 Kaleidoscope
What a commotion!

34 Contest
When days are months ... At sixes and sevens

36 Looking Back
The secret of the Venerable Cooper ... From the prehistory of radio

43 Quantum Smiles
Disorder in the court! ... A horse is a horse

47 At the Blackboard
Constructions with compass alone

50 Mathematical Surprises
Friezing our way into summer

52 In the Lab
Walker in a winter wonderland

54 Happenings
Summer study in New York and Tartu ... The American Regions Mathematics League ... Bulletin Board

59 Solutions

64 Checkmate!
Symmetry on the Chessboard

Have you written an article that you think belongs in *Quantum*? Do you have an unusual topic that students would find fun and challenging? Do you know of anyone who would make a great *Quantum* author? Write to us and we'll send you the editorial guidelines for prospective *Quantum* contributors. Scientists and teachers in any country are invited to submit material, but it must be written in colloquial English and at a level appropriate for *Quantum*'s predominantly high school readership.

Send your inquiries to:

Managing Editor
Quantum
1742 Connecticut Avenue NW
Washington, DC 20009-1171

Be a factor in the

QUANTUM

equation!

QUANTUM

THE STUDENT MAGAZINE OF MATH AND SCIENCE

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

Quantum (ISSN 1048-8820) contains authorized English-language translations from *Kvant*, a physics and mathematics magazine published by the Academy of Sciences of the USSR and the Academy of Pedagogical Sciences of the USSR. Copyright © 1990 National Science Teachers Association. Subscription price for 1990-91 (four issues) is $9.95 in the US and $17.95 for all other countries. Bulk subscriptions: 20-49 copies, $9.00 each; 50-99 copies, $8.00 each; 100+ copies, $7.50 each. Correspondence about subscriptions, advertising, and editorial matters should be addressed to *Quantum*, 1742 Connecticut Avenue NW, Washington, DC 20009-1171.

This project was supported, in part, by the
National Science Foundation
Opinions expressed are those of the authors and not necessarily those of the Foundation

Quantum in outer space and the inner space of art

SPACE TRAVEL: PERHAPS nothing else symbolizes quite so well the scientific aspirations and achievements of our age. Countless generations of stargazers would envy us as we explore the outer regions, whether "in person" or by means of sophisticated bundles of instrumentation. I say "we," but how many of us will ever get the chance to see an Earthrise on the Moon, or conduct experiments in orbit, or travel to another planet? Well, I'll have to be satisfied that my *photograph* has flown in space—copies of the premier issue of *Quantum* have been carried aboard Soviet and American spacecraft. It's my belief that more than a few of you will have the opportunity to explore the outer reaches of space, and not just in picture form!

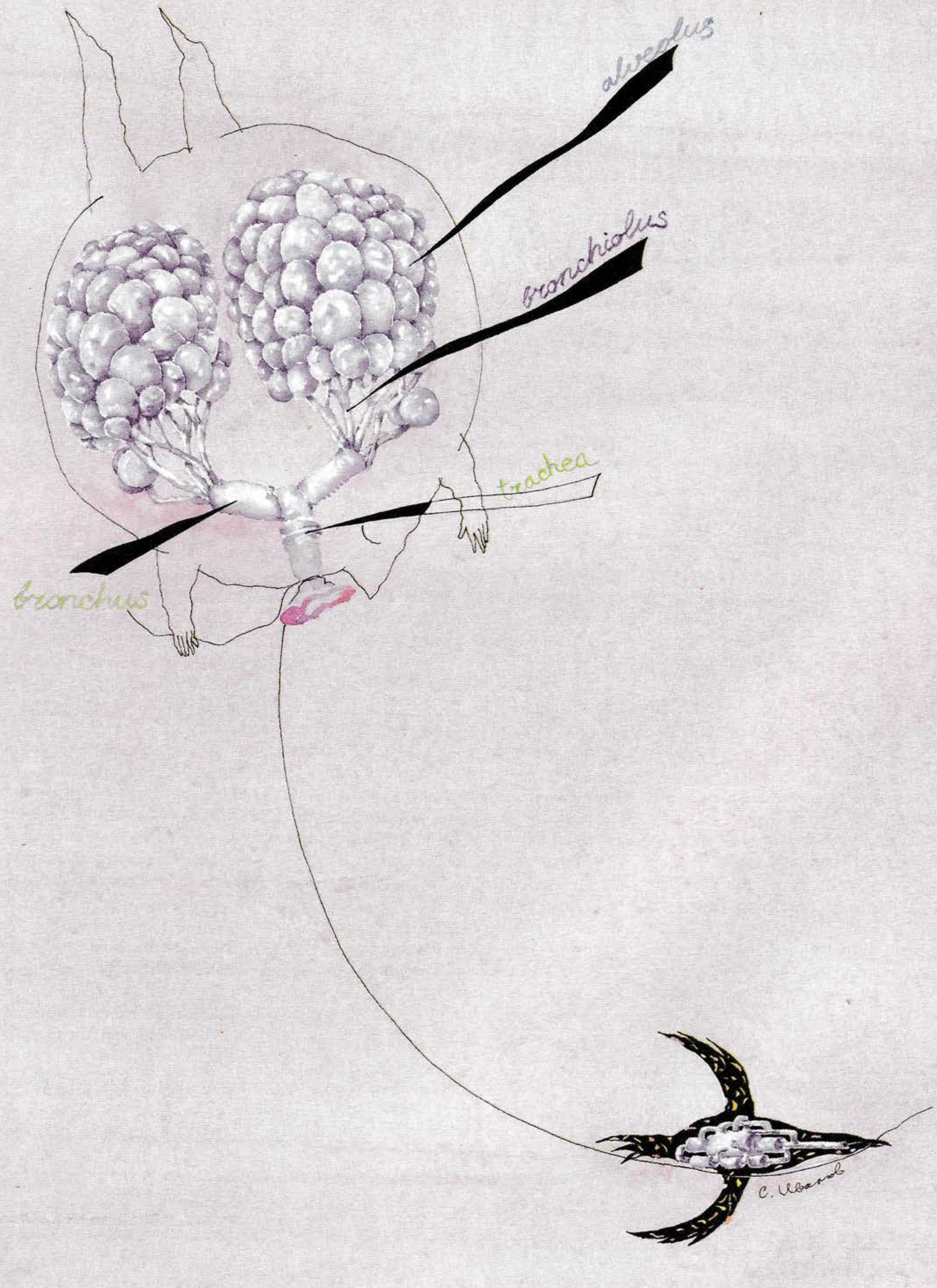
In February a conference was held in Deauville, France, on the International Space Year, which is set for 1992. An educational component was co-chaired by Victor Borovishki, the deputy editor in chief of *Kvant* magazine in the USSR, and myself. Some 40 countries were represented. One of the major recommendations to come out of this conference is to designate the January 1992 issue of *Quantum* an official publication of the International Space Year. Authors would be drawn from as many countries as possible, and the articles would be written by some of the best scientists in those countries. The magazine would be translated into as many languages as possible and distributed throughout the world. It would cover many space science topics of interest to students in all these countries. This special issue of *Quantum* would also contain schedules of activities for the International Space Year, including student exchanges, competitions, and other kinds of events that are likely to occur. I encourage all of you to participate in the International Space Year activities. We'll try to keep you informed through *Quantum* magazine, especially the January 1992 issue.

I'D LIKE TO MAKE a few comments about the illustrations provided to us for *Quantum* by our colleagues in the Soviet Union. Ever since I first encountered *Kvant* magazine, I have been fascinated by the differences between *Kvant* artwork and the kind of technical illustration we usually encounter here in the United States. Our illustrations are technically correct and simple, with nothing superfluous. I recently had the opportunity to explore this topic with *Kvant*'s art director, Sergey Ivanov, who was among the group of *Kvant* editors who visited the *Quantum* offices in Washington to participate in our advisory board meeting.

A talented artist himself, Sergey creates some of the artwork that appears in *Kvant* and *Quantum*. (In fact, Sergey drew the cover of this issue specifically for *Quantum*.) What I discovered in talking with Sergey is that the illustrations in *Kvant* are often prepared by leading artists in the USSR. The technical aspect of the illustration is always coupled with an appeal to the aesthetic side of our nature. Often the art draws upon a fable, legend, or myth; sometimes it's related to a famous painting or sculpture. (The picture of the three cows in the Yaglom article, for instance, was inspired in part by a well-known sculpture by Kanova, "The Three Graces," in the Hermitage Museum in Leningrad.) Sometimes historical figures and events are cleverly woven into the artwork.

Whatever approach the *Kvant* artist takes, the illustrations end up being more meaningful. They convey far more than the technical idea, important as that is—they're also pleasing just to look at. I hope you enjoy the art in *Quantum*. Because of differences in printing technology, the art you see is even more resplendent than it originally appeared in *Kvant*. Perhaps, with the changes that are so rapidly transforming the USSR, this will change as well. But one thing I hope does *not* change is the artistic tradition represented by Sergey Ivanov and his fellow artists. It seems to me this is one of many areas where we can learn something from the Soviet Union.

—Bill G. Aldridge


alveolus
bronchiolus
trachea
bronchus
С. Иванов

How do we breathe?

"In and out" is the silly answer, but the straight answer may surprise you

by K. Y. Bogdanov

"It is now impossible to clarify biological questions if you don't know physics."
—Julius Mayer, 1842

WHAT DOES AN ELECTROcardiogram tell us? How do bats catch flies in the dark? How do carrier pigeons find their way home? The answers to these and many other questions are given by biophysics, the science that studies physical phenomena in living organisms. The biophysicist works on the boundary between physics and biology, using the latest achievements of physics in biological investigations.

Armed with our knowledge of high school physics, let's try to answer a seemingly simple biological question: How do we breathe?

Our respiratory system

When our organism breathes normally, it consumes about 0.5 kg of oxygen a day and exhales almost the same amount of carbon dioxide. Oxygen enters and carbon dioxide exits through the lungs.

The inner space of the lungs communicates with the atmosphere via the respiratory tract. The respiratory tract consists of the nasal cavity (where air is warmed and moistened), pharynx, larynx, trachea, and the two principal bronchi (which supply air to the right and left lungs).

Each bronchus is divided into smaller bronchi (bronchioli) and ends in microscopic bubbles, or "alveoli," encased in a thick network of blood vessels. An adult has 700 million of these interconnected, air-filled alveoli (fig. 1). At any given moment the blood vessels surrounding the alveoli contain approximately 70 ml of blood. Carbon dioxide diffuses from the blood into the alveoli, while oxygen diffuses in the opposite direction. The huge combined surface area of the alveoli makes it possible to saturate the blood with oxygen and cleanse it of excess carbon dioxide.

Breathing and soap bubbles

Is it easy for us to take a breath? How much does the air pressure inside the lungs (in the alveoli) exceed the pressure outside (in the pleural cavity) during the process of taking a breath?

If each alveolus is assumed to be a hollow ball made of an elastic membrane, the pressure needed to keep it inflated for a given external pressure is determined completely by the ball's diameter and the membrane's elasticity. Is this correct?

In 1929 the Swiss scientist Karl von Niirgard discovered that the pressure needed to inflate the lungs can be substantially reduced if the lungs are filled with a physiological solution. In no way did this accord with the idea that alveoli are hollow elastic balls: if in taking a breath we overcome only elastic forces, our efforts don't depend on whether the lungs are filled with a solution because elasticity doesn't depend on that. Niirgard's data could be explained when it was found that the inside of each alveolus is coated with a thin layer of liquid. This circumstance fundamentally alters the mechanical properties of the alveolus. In particular, the pressure needed to inflate an alveolus is greater than that needed to inflate a hollow ball made of pulmonary tissue. Here's why.

It's known that the surface of a liquid behaves like a piece of stretched film—that is, it possesses surface tension. In order to estimate the role of surface tension in the mechanics of an alveolus, let's consider a spherical film made of a liquid. The simplest example of such a film is a soap bubble. Because of the liquid's surface tension, the air pressure inside the bubble is always greater than the outside pressure. The amount of excess pressure inside a sphere of radius R is determined by Laplace's relation $\Delta p = 2\sigma/R$, where σ is the coefficient of surface tension for the film. The greater this excess pressure, the smaller the radius of the sphere. In a soap bubble film, both the inside and outside surfaces are in contact with air, so the coefficient σ for the film is twice the coeffi-

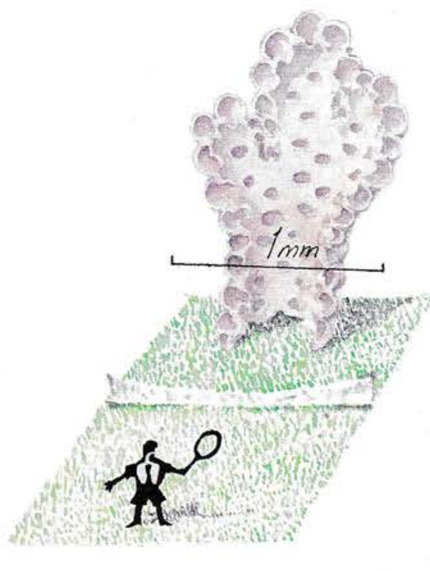


Figure 1
The terminal branches of a lung are the alveoli. The diameter of an alveolus is, on average, 0.1 mm. The walls are 0.4 μm thick. The total surface area of alveoli in an adult human being is about 100 m²—just about half a tennis court!

Art by Sergey Ivanov

cient of surface tension for the liquid. In alveoli, air–liquid contact takes place only on one side, the inside.

Let's estimate Δp. For intracellular fluid, $\sigma = 5 \cdot 10^{-2}$ N/m. We'll also take this value for the fluid that coats the inner surface of the alveolus. Assuming that $R = 50\ \mu\text{m} = 5 \cdot 10^{-5}$ m, we get $\Delta p = 2 \cdot 10^3\ \text{N/m}^2$.

Figure 2 shows how the volume of the lungs depends on the pressure inside them (more exactly, how much this pressure exceeds the external pressure). It's clear from the graphs that, if not all, then at least a considerable part of the pressure that expands the lungs during a breath is spent to overcome the forces of surface tension. But when the lungs are filled with the physiological solution, additional pressure is needed only to overcome the elastic properties of the pulmonary tissue. It's obvious that the difference between the two curves in figure 2 represents the contribution of surface tension to the lung's elasticity. In ordinary breathing, the volume of a human being's lungs is about 50% of their maximum volume. It follows from figure 2 that the contribution of surface tension is more than 30%.

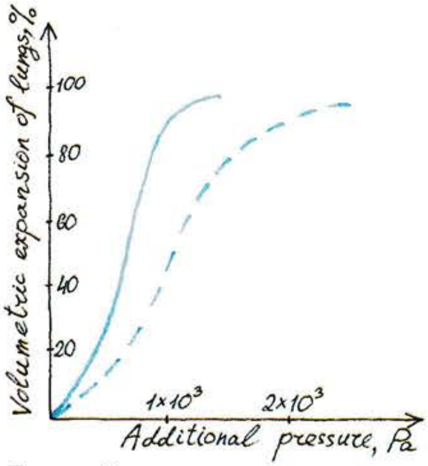


Figure 2
The relation between lung volume and pressure of the air (broken line) or liquid (solid line) inside.

But it's not all so simple

Our use of surface tension to explain the lung's mechanical properties leads to a "paradox" when we study the interaction between neighboring alveoli. Figure 3 shows (in outline) two neighboring alveoli of different sizes. First let's assume that the air cavities of the two alveoli are not connected with each other (fig. 3a). The air pressure p_1 in the left alveolus is greater than the pressure p_2 in the right alveolus (since $R_1 < R_2$). Now, as soon as we open an imaginary valve connecting the air cavities of the two alveoli, the air from the left alveolus begins to flow into the right alveolus until the pressure in both alveoli becomes equal (fig. 3b). So when two alveoli of different sizes are connected, the larger alveolus will inflate and the smaller one will deflate.

It's obvious that such an interaction between the neighboring alveoli would cause all the lung's small alveoli to deflate and the large alveoli to burst. As a result, the lung wouldn't be able to function.

So how *do* we breathe?

Looking at the interaction between neighboring alveoli, we assumed that different alveoli have the same surface tension, which doesn't depend on the state (inflated or deflated) of the alveoli. The surface tension of pure liquids is indeed independent of the size of the surface. But the coefficient of surface tension of liquids containing different impurities depends on the size of the interface between the liquid and gas.

Again nature has foreseen everything. The liquid wetting the inside surface of the alveoli contains, as an impurity, a substance that controls the value of the surface tension in such a way that σ is minimal at the beginning of an inhalation (minimum surface) and maximal at the end (maximum surface). So despite the fact that the radius of an alveolus is very small at the beginning of an inhalation, the contribution of surface tension is small. This makes it possible to blow up a deflated alveolus by means of relatively low pressure. At the same time, the increase in σ with the radius of an alveolus prevents overinflation at the height of air intake. In addition, this dependence of σ on the size of the alveolus regulates relations between neighboring alveoli, preventing the appearance of the mechanism shown in figure 3.

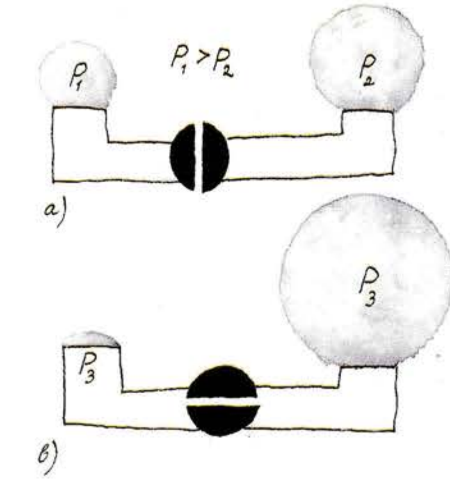


Figure 3
A schematic depiction of two neighboring spherical alveoli with different radii when (a) the air cavities are isolated and (b) the air cavities are connected.

Why do impurities make surface tension dependent on surface area? As a rule, the impurities reducing σ are quite similar to ordinary soap in their chemical structure. When they dissolve, they form a thin film on the water's surface. If the impurity concentration is high, so that the film can cover the whole surface in a continuous layer, then σ for such a liquid is equal to a certain value for the given impurity. But when the concentration of the impurity is insufficient for the film to cover the whole surface, the value of the surface tension will be somewhere between the corresponding values of σ for water and for the impure liquid. An increase in the liquid's surface then leads to a decrease in the surface concentration of the impurity and causes σ to increase, bringing its value nearer to σ for water. If the liquid's surface area decreases, σ will also decrease.

Now look at figure 4. It shows how surface tension depends on the size of the air–liquid interface. Do you notice that for the same surface area of contact, the value of σ at inhalation is always greater than at exhalation? This is because the impurity that reduces surface tension is in the interior of the liquid and not just on the surface. When the surface area increases, the impurity concentration on the surface decreases. Some molecules of the impurity rush to the surface, but equilibrium isn't established instantaneously. So when we begin to inhale, for instance, the rapid increase in the liquid's surface area is accompanied by a sharp increase in σ because the impurities dissolved in the liquid have no time to emerge on the surface. Equilibrium between impurity molecules sets in only at the end of the inhalation. In much the same way, the end of the exhalation corresponds to an equilibrium state. This explains the dependence of σ on surface area shown in figure 4.

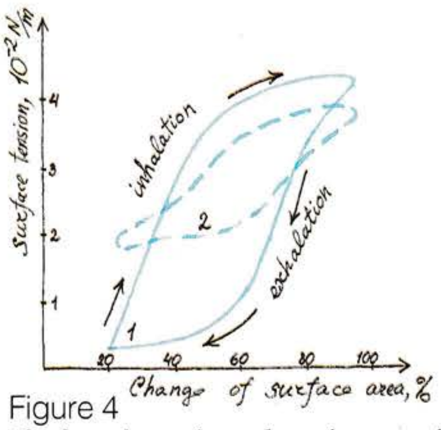


Figure 4
The dependence of σ on the surface area of liquid–air contact (1) for liquid isolated from the alveoli of a healthy person and (2) for liquid taken from the alveoli of newborns unable to breathe independently.

So we have substances in our lungs that reduce surface tension and make our breathing easier. But where do they come from? It turns out they're synthesized by special cells located in the alveoli walls. These "surface-active substances" (SAS) are produced throughout a human being's life, from birth to death. Curve 2 in figure 4 corresponds to pulmonary liquid with a reduced SAS content. You can see that the minimum value of σ is eight times the norm. Some babies are born without the cells that generate SAS. In these rare cases, the newborns cannot breathe on their own. Unfortunately, infants still die all over the world, never taking their first breath, because of a deficiency or absence of SAS in their alveoli.

An exception to the rule

Many animals that breathe by means of lungs do not suffer at all from the absence of SAS in their alveoli. This applies, for instance, to cold-blooded animals—frogs, lizards, snakes, crocodiles, and so on. Since these animals don't need to spend energy to heat their own organisms, their need for oxygen is, on average, ten times less than that of warm-blooded animals. So the surface area across which gas is exchanged between blood and air is less than in warm-blooded animals. The relative reduction of the lung's surface area in cold-blooded animals is due to the fact that the diameter of their alveoli is approximately ten times greater than the diameter of the alveoli of warm-blooded animals. The comparatively larger radius of the alveoli makes it possible to inflate them easily even if there is no SAS on their inner surface. (In fact, $\Delta p \sim 1/R$.)

Another group of animals that have no SAS in their lungs consists of warm-blooded creatures that live a rather fast-paced, active life: birds. Mammals and birds of equal weight have almost the same energy expenditures, and a bird's oxygen needs are great as well. But a bird's lungs have the unique ability to saturate the blood with oxygen when it is flying at great altitudes (about 6,000 m), where the concentration of oxygen is half that at sea level. Mammals (including humans) at such a height start to feel a deficiency of oxygen, sharply limit their activity, and sometimes even become dazed. How can the lungs of a bird, without any SAS, allow it to breathe, saturating its blood with oxygen? And why can't mammals do this?

Let's engage in a little self-criticism. What's wrong with our lungs? For one thing, not all the inhaled air takes part in the exchange of gases with the blood. The air inside the trachea and bronchi at the end of an inhalation can't give oxygen to the blood and take carbon dioxide from it because there are practically no blood vessels in these parts of the lung. So the portion of the lung occupied by the trachea and bronchi is called "dead space." As a rule, the dead space in a human being's lungs is about 150 cm^3. You can artificially increase the volume of dead space by breathing through a long pipe. If you try this, you'll probably notice that you have to increase the depth of your breathing (that is, the volume of inhaled air). Obviously, if the volume of dead space is made equal to the maximum possible inhalation volume (about 4,500 cm^3), you'll start to suffocate—no fresh air enters the alveoli at all! The presence of dead space in the respiratory system of mammals is a "mistake" of nature.

And nature made a second "mistake," I would say, in its design of mammalian lungs. This deficiency has to do with the fact that air moving in the lungs changes direction as inhalation is followed by exhalation. Almost half the time the lungs are practically idle—during exhalation fresh air doesn't enter the alveoli.

But nature again attained perfection in the bird family. Besides ordinary lungs, birds have an additional system consisting of five or more pairs of air bags connected to the lungs (fig. 5). The cavities of these bags are widely distributed throughout the body.

CONTINUED ON PAGE 42

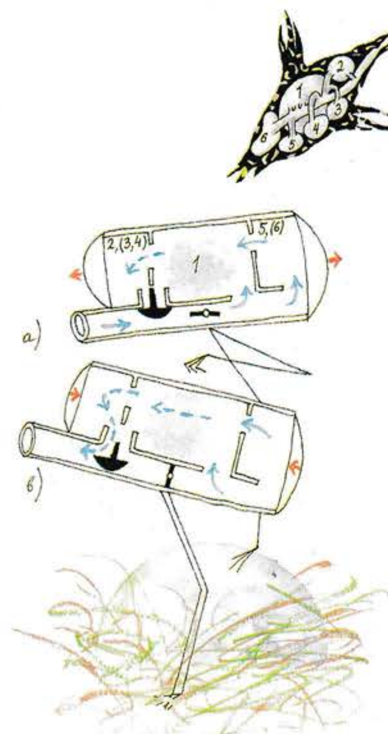
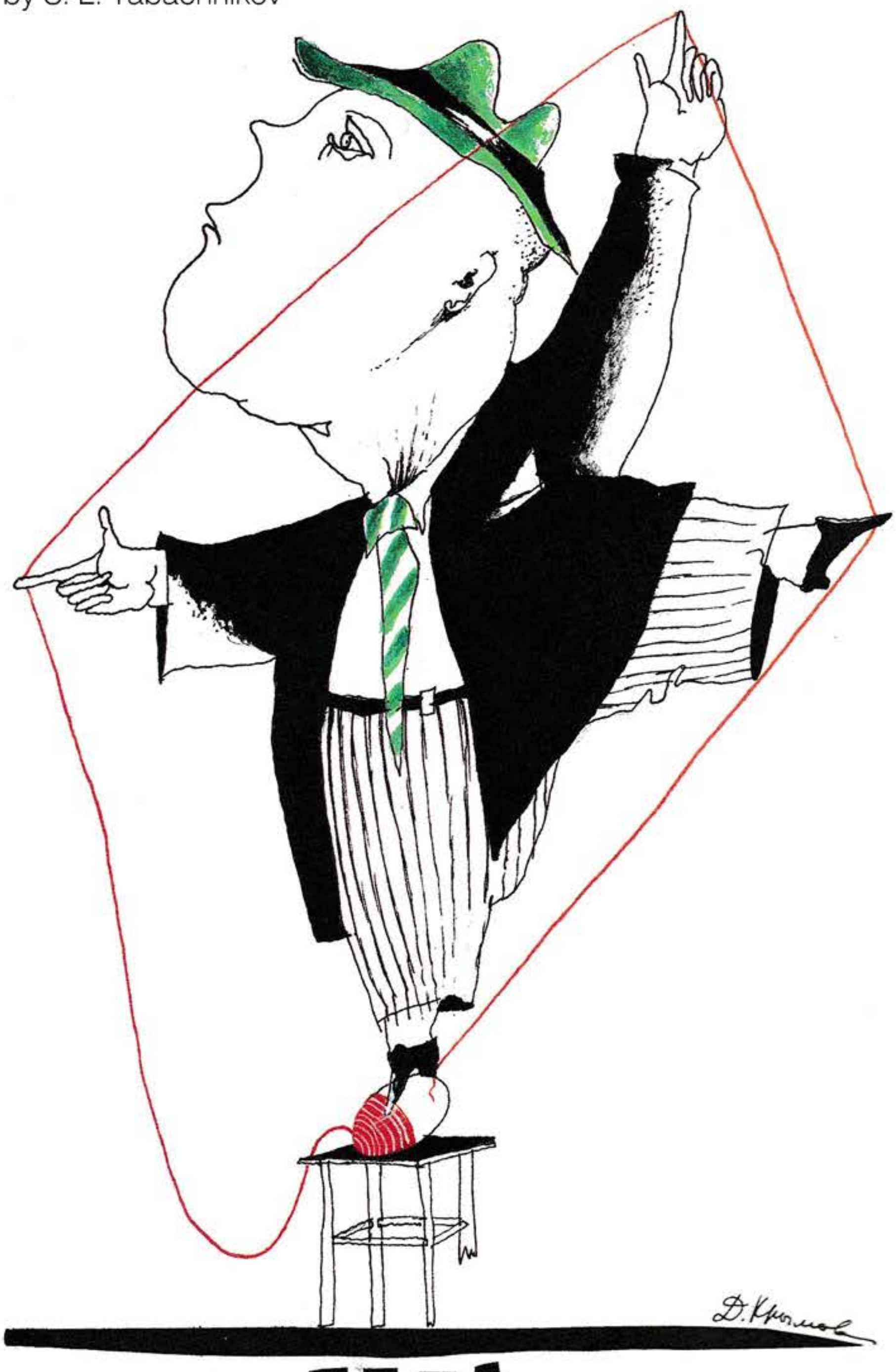


Figure 5
The respiratory system of a bird: (1) lungs; (2–6) air bags. Arrows show the movement of air when the bird (a) breathes in and (b) breathes out.

Considerations of continuity

If you continue reading long enough, you'll learn how to get your wobbly kitchen stool to sit still

by S. L. Tabachnikov



Art by Dmitry Krymov

MOST OF THE PROBLEMS IN your math textbook ask you to solve equations, or compute a certain specific value, or construct a geometric figure with given properties, and so on. This article is devoted to problems of another sort: you're merely asked to prove that the required root, or number, or figure, or whatever, actually exists. The method we'll use to solve them is known as "considerations of continuity."

Plane sets

We'll start with the following problem: Given a certain set bounded by a closed curve in the plane, prove that there exists a vertical line that divides it into two parts of equal area.

Let's take a vertical line to the left of the given set (fig. 1a) and start moving it to the right. The line eventually touches the boundary of the set (fig. 1b), then slides over the set (fig. 1c–1f) until it moves past it to the right (fig. 1g). As the line moved, how did the area of the part of the set to the left of the line (shown in red) vary? Obviously this area changed *continuously* from zero (fig. 1a) to the total area S of the set (fig. 1g). So at some time this area was equal to exactly half of the total area. At that precise moment the line divided the set into two parts of equal area $S/2$.

What do you think of this solution? It doesn't give a method for constructing the required line—it only establishes its *existence*. (Actually, it wouldn't make sense to look for a recipe for constructing this line—the given set is arbitrary.) It shows that there exists a line of *arbitrary direction* dividing the set into two parts of equal area. Also,

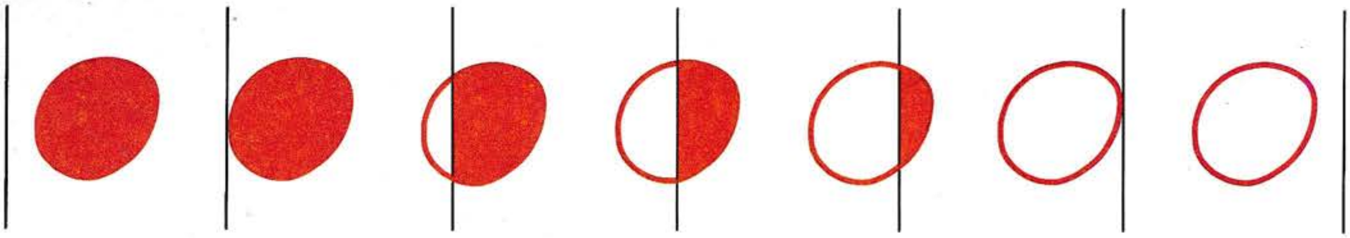


Figure 1 a b c d e f g

the solution implies that there is a *unique* line in any given direction that divides the set into parts of equal area. We'll use this fact later.

Let's look at the solution from a different angle. Choose a horizontal numerical axis in the plane (fig. 2a). The position of an arbitrary vertical line l is determined by the number x (the point where the line intersects the numerical axis). Consider the area of the part of our set to the left of line l as a function $f(x)$. The graph of this continuous function is shown in figure 2b. Finding a vertical line dividing our set into two parts of equal area is then the same as finding a point c on the numerical axis such that $f(c) = S/2$.

Now let's consider the horizontal line $y = S/2$ in figure 2b. The left part of the graph of $f(x)$ lies below this line and the right part lies above it, since $f(a) = 0 < S/2$, while $f(b) = S > S/2$. So there exists a point c where the horizontal line $y = S/2$ and the graph of $f(x)$ intersect. It's precisely at this point c that we have $f(c) = S/2$.

The property of continuous functions used in this proof is known as the "intermediate value theorem"—a continuous function assumes all the intermediate values between any two of its values.

A more formal version of this theorem goes like this: If $f(x)$ is a continuous function on the closed interval $[a,b]$ and c is some number between the numbers $f(a)$ and $f(b)$, then there exists a point x_0 on the interval $[a,b]$ such that $f(x_0) = c$.

For example, in our problem the function $f(x)$ assumes the value $S/2$ contained in the interval between the values $f(a) = 0$ and $f(b) = S$.

The intermediate value theorem is almost self-evident. Like many obvious statements, though, it's not too easy to prove. We'd have to go into an in-depth treatment of the notions of "continuous functions" and "real numbers." This would distract us from the main topic of this article, so let's just leave it at that—the intermediate value theorem is simply obvious to us.

And another thing: we didn't prove that the area to the left of our vertical line continuously depends on its position (that is, that $f(x)$ is a continuous function). This is almost self-evident too. Throughout the rest of this article, we'll dispense with proving the continuity of functions arising in our solutions.

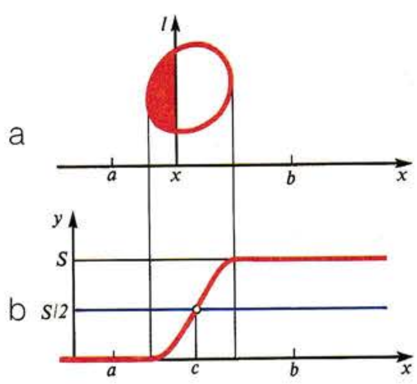


Figure 2

Problems

1. (a) Given a convex set[1] and a point lying outside it, prove that there exists a straight line passing through this point that divides the set into two parts of equal area.

(b) What happens to this problem if the point extends to infinity?

(c) Solve case (a) when the point is inside the set.

(d) Is it true that there is only one solution in cases (a) and (c)?

2. (a) Given two convex sets (which may intersect), prove that there exists a line that simultaneously divides each of the two sets into two parts of equal area. (This problem is often called the plane version of the ham-and-cheese-sandwich problem: Can one always cut such a sandwich with one slash of the knife into two parts so that each contains as much ham and as much cheese as the other?)

(b) What happens to this problem if one of the sets contracts to a point?

(c) Suppose each of the sets in case (a) is a parallelogram. Construct the required line.

3. Given a convex set, prove that there exists a line that simultaneously divides the area of the set and the length of its boundary curve into equal parts.

4. (a) Given two convex sets, one contained within the other, prove that there exist two parallel chords of the bigger set, tangent to the inner one, such that $S_1 = S_2$ (fig. 3).

(b) What happens to this problem if the inner set contracts to a point?

5. Prove that any convex set contains equal and parallel chords dividing its area into three equal parts.

6. Prove that a square can be circumscribed about any convex set.

[1]A "convex set" is a set that contains all line segments joining any two of its points. For example, a disk (the part of the plane bounded by a circle) is convex, while its boundary (the circumference itself) is not.

Chairs and squares

Imagine that the floor in your kitchen is not too even (which is often the case if it's covered by linoleum). A kitchen stool will usually touch the floor with only three of its legs, while the fourth will be slightly up in the air. Is it always possible to move the stool so that all four legs touch the floor or, no

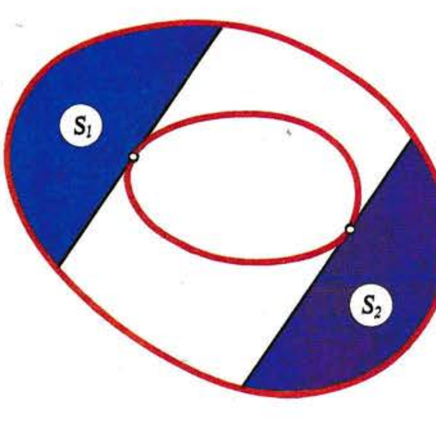


Figure 3

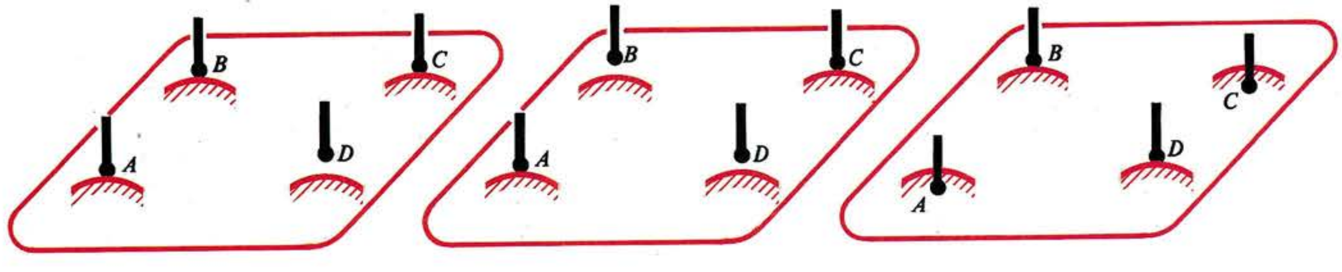


Figure 4 a b c

matter how you move it, will it always wobble? Of course, we assume that when the stool is on all four legs, it doesn't have to be absolutely level.

(Before you begin the mathematical solution of this problem, try to find the experimental solution right in your own kitchen. After you get a positive answer to the question, come back to the mathematics of the problem.)

We'll assume that the kitchen floor, though not a plane, doesn't differ too much from a plane surface. This assumption allows us to ignore floors of "pathological" shape—for example, the floor of a cave covered with stalagmites (upside-down calcium icicles).

So let's suppose that legs *A*, *B*, *C* of our stool are on the floor, while the fourth leg *D* hangs in the air (fig. 4a). Without lifting legs *A* and *C* off the floor, rotate the stool about line *AC* so that legs *B* and *D* are both off the floor at the same distance from it (fig. 4b—the distance from the floor is counted in the direction of the legs). In this position (two legs on the floor, the other two at the same distance from the floor) the stool can be moved.

Imagine for a minute that the floor is made of soft clay into which the legs of the stool can easily penetrate. Let's lower the stool until legs *B* and *D* just touch the floor, while legs *A* and *C* have been pushed into the clay below floor level to the same depth (fig. 4c).

Now, going back to the setup in figure 4b, let's rotate the stool counterclockwise about its center so that legs *A* and *C* continue to barely touch the floor, while the distance from floor level (that is, the depth) of leg *B* remains equal to that of *D*. After we've rotated through an angle of 90°, legs *A* and *C* will occupy the positions of legs *D* and *B* in figure 4c, while legs *B* and *D* will take up the positions of legs *A* and *C* in figure 4c. While the stool is being turned, the distance of legs *B* and *D* from the floor changes continuously. At the outset (fig. 4b) this distance was positive, since both legs *B* and *D* were above floor level; at the final moment, shown in figure 4c, it's negative, since legs *B* and *D* are below floor level. So there must have been a moment when that distance was equal to zero. At that instant the stool was standing on all four legs.

Our solution has not only theoretical but practical value. By rotating a stool about its center by less than 90°, we can always find its stable position. You can check this experimentally.

One more thing. Our solution is based on the fact that the extremities of the four legs of a stool form the vertices of a square. Were they located at the vertices of a rectangle, or some other quadrilateral, our argument would not have worked. I don't know if the statement of the problem is true for any quadrilaterals other than the square. (In any case, it's clear that the quadrilateral must be inscribed in a circle.) If you're able to settle this question, please let us know.

Problems

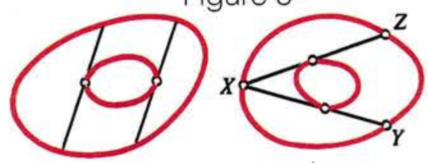
7. Your math teacher asked the class to prove that any convex set can be divided by a pair of perpendicular lines into four parts of equal area. One of your classmates proposed the following solution: "Choosing an arbitrary direction, we can draw a unique line *l* that divides the area of our set in half (fig. 5a). Each half can be divided by lines perpendicular to *l* (fig. 5b). Let's begin to change the direction of *l*, repeating the previous configuration. After the direction changes by 180°, points *A* and *B* will change places. So at some moment these points will coincide. At that time figure 5b will look like figure 5c and we will have obtained the required pair of perpendicular lines."

(a) Find the error in this argument.

(b) Find a correct solution.

8. Prove that any convex set with a center of symmetry possesses an inscribed square. (Actually, a square can be inscribed in any set. But for the proof of this statement, given by the outstanding Soviet mathematician L.G. Shnirelman (1905–1938), simple considerations of continuity are no longer sufficient.)

Figure 6



a b

9. Given two convex sets, one contained inside the other, prove the following statements:

(a) There are two parallel chords of the bigger set that are tangent to the inner set and are of equal length (fig. 6a).

(b) There is a point *X* of the bigger set from which two equal tangents to the inner set can be drawn (fig. 6b).

Hint: Among the chords of the bigger set

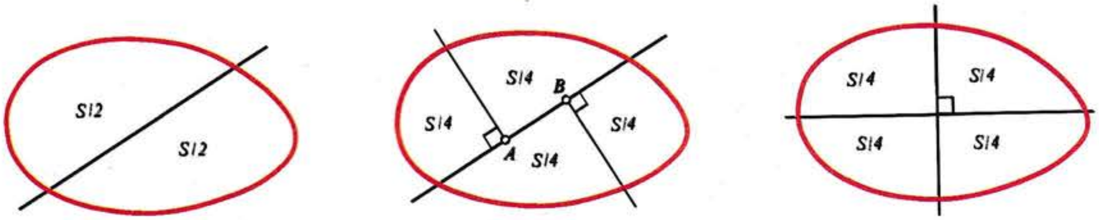


Figure 5 a b c

that are tangent to the inner one there is a longest chord. From its ends, draw tangents to the inner set.

10. Given three nested convex sets (fig. 7), prove that there exists a point X of the biggest set from which tangents to the innermost set can be drawn so that the parts of these tangents contained in the second largest set are of equal length.

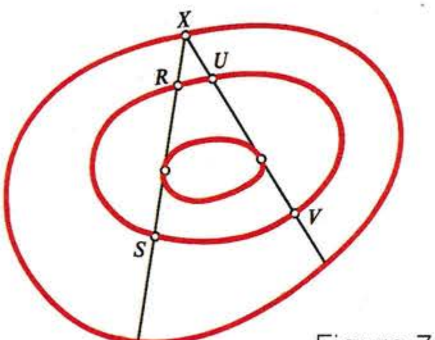


Figure 7

Roots and chords

You know, of course, that a quadratic trinomial might not have a real root. Such is the case, for instance, with the polynomial $x^2 + x + 1$. But what's the situation with cubic polynomials? We'll now prove that *any third-degree polynomial has at least one real root*.

First let's consider a cubic polynomial with leading coefficient 1:

$$f(x) = x^3 + bx^2 + cx + d.$$

We can rewrite it in the following form:

$$f(x) = x^3(1 + b/x + c/x^2 + d/x^3).$$

If the value of $|x|$ is very large, the summands b/x, c/x^2, d/x^3 become very small. In that case the number in parentheses is very close to 1 and is certainly positive. So for large $|x|$ the sign of $f(x)$ is determined by the sign of the number x^3. Likewise $f(x)$ is less than 0 for negative x with large absolute values, and $f(x)$ is greater than 0 for large positive x.

Because a cubic polynomial, like any polynomial, is a continuous function, we can apply the intermediate value theorem. The theorem implies that there exists an x such that $f(x) = 0$, so the cubic polynomial has a root.

The case of the general cubic polynomial

$$f(x) = ax^3 + bx^2 + cx + d$$

reduces to the case considered above by dividing $f(x)$ by a.

Our proof is a good illustration of the strength and weakness of continuity considerations. We obtained the proof of a difficult fact—the existence of a root—almost "for free," but it remains unclear how to find this root for a specific polynomial. (Actually, formulas for the roots of cubic polynomials exist, but that's an entirely different topic.)

Here's another example. Suppose $f(x)$ is a continuous periodic function of period T. Prove that its graph has a horizontal chord of length $T/2$.

The existence of a horizontal chord of length l is equivalent to the relation $f(x) = f(x + l)$ (fig. 8). So we have to prove that there is an x such that $f(x + T/2) = f(x)$. In other words, we must show that the function $g(x) = f(x + T/2) - f(x)$ has a root—that is, $g(x) = 0$ for a certain x.

Let's take an arbitrary number a. If $g(a) = 0$, we're all done. So we'll assume that $g(a)$ is nonzero. To be definite, let's suppose $g(a) < 0$. Let $b = a + T/2$. We can easily compute $g(b)$:

$$\begin{aligned} g(b) &= g(a + T/2) \\ &= f(a + T/2 + T/2) - f(a + T/2) \\ &= f(a + T) - f(a + T/2) \\ &= f(a) - f(a + T/2) \\ &= -g(a). \end{aligned}$$

Therefore, $g(b) > 0$.

The function $g(x)$ is continuous, since it's the difference of continuous functions. So once again we can use the intermediate value theorem, which implies that there is a number x between a and b such that $g(x) = 0$.

The statement about horizontal chords of graphs of continuous periodic functions can be considerably strengthened: *The graph of such a function has a horizontal chord of arbitrary length.*

But for graphs of functions defined on a closed interval, the situation is quite different. Suppose a continuous function is defined on a closed interval of length T and assumes equal values at the endpoints of the interval. If the number l is of the form T/n, where n is a positive integer, then the graph of the function necessarily has a horizontal chord of length l; if l is not of that form, there exists a continuous function with the same values at the endpoints whose graph has no horizontal chords of length l.

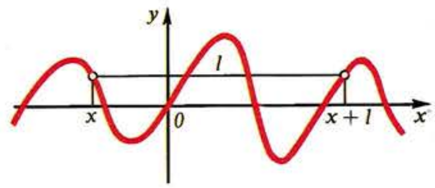


Figure 8

Problems

11. Prove that any odd-degree polynomial has a root.

12. Yesterday at midnight it was colder than it was at midnight the day before yesterday and today. Prove that at some time today the temperature was the same as yesterday at the same time.

13. Suppose $f(x)$ is a continuous function on some closed interval and that all its values lie in that same interval. Prove that there exists an x such that $f(x) = x$.

14. The polynomial $f(x) = ax^2 + bx + c$ is such that the equation $f(x) = x$ has no real roots. Prove that the equation $f(f(x)) = x$ also has no roots.

15. Suppose $f(x)$ is a continuous periodic function (on the numerical line) of period T. Prove that its graph has horizontal chords of length l, where (a) $l = T/3$, (b) $l = T/n$, (c) $l = (p/q)T$, (d) l is any real number.

16. Suppose $f(x)$ is a continuous function on a closed interval of length T whose values at the endpoints are the same. Prove that the graph of $f(x)$ has a horizontal chord of length l, where (a) $l = T/2$, (b) $l = T/3$, (c) $l = T/n$. Also, construct such a function whose graph doesn't have horizontal chords of length $2T/3$.

17. The "mean value of the function $f(x)$" on the closed interval $[a,b]$ is, by definition, the number

$$\frac{1}{b-a}\int_a^b f(x)dx.$$

Prove that if the mean value of a continuous function on a closed interval is 0, then the function assumes the value 0 in the interval.

18. Consider a function of the form $f(x) = a_1\sin x + a_2\sin 2x + \ldots + a_n\sin nx + b_1\cos x + b_2\cos 2x + \ldots + b_k\cos kx$, where $a_1, \ldots, a_n$ and $b_1, \ldots, b_k$ are real numbers. Prove that the equation $f(x) = 0$ has a root.

Conclusion

We've worked out several problems whose solutions are based on considerations of continuity. In each of them the relevant magnitude depended on one parameter. For example, the area of the part of a set to one side of a line passing through a fixed point depended on the line's angle of inclination. Considerations of continuity also work when there's more than one parameter. Now isn't the time to go into a detailed account of this theme—I'll just leave you with a few theorems to mull over.

1. Given three bodies arbitrarily located in space, there exists a plane that divides the volume of each into two equal parts. (This is the space version of the ham-and-cheese-sandwich problem, the three bodies being the slice of bread and the pieces of ham and cheese.[2])

2. There are two diametrically opposed points on the planet right now where both pressure and temperature coincide.

3. A cube may be circumscribed about any convex body. (This theorem generalizes problem 6 to space.)

4. Imagine a sphere covered with hair. Can it be combed smoothly—that is, so that each hair is tangent to the sphere and the directions of nearby hairs do not differ too much? The answer is no: there will always be a hair that will stick out perpendicularly to the sphere. This statement is known as the "porcupine theorem" or the "sphere-combing theorem."

5. This is for those of you who know complex numbers. We've shown that any odd-degree polynomial has a real root. What about complex roots? It turns out any polynomial (except a constant zero-degree polynomial) has at least one complex root. This statement is so important it's often called the *fundamental theorem of algebra*. And one of the approaches to its proof is based on—you guessed it—considerations of continuity. ◘

[2]An open-faced sandwich!—*Ed.*

BRAINTEASERS

Just for the fun of it

Problems offered for your enjoyment by E. Chernyshov, N. Antonovich, A. Savin, B. Proizvolov, I. Slobodetsky, and L. Aslamazov

B6

How can a goat, a head of cabbage, two wolves, and a dog be transported across a river if it's known that the wolf is "culinarily partial to" goat and dog, the dog is "on bad terms with" the goat, and the goat is "not indifferent to" cabbage? There are only three seats in your boat, so you can only take two passengers—animal or vegetable—at a time.

B7

Thirty people took part in a shooting match. The first participant scored 80 points, the second scored 60 points, the third scored the arithmetic mean of the number of points scored by the first two, and each subsequent competitor scored the arithmetic mean of the number of points scored by the previous ones. How many points did the last competitor score?

SOLUTIONS ON PAGE 61

Art by Edward Nazarov

B8

When we multiply multidigit numbers, we have to do some addition as well—the final step is to sum up the subtotals. Here's a puzzle that exposes the steps and tags the various digits as odd ("O") or even ("E"):

```
     EEO
x     OO
    EOEO
+   EOO
   OOOOO
```

Find numbers that satisfy this scheme.

B9

A scuba diver loses his bearings deep in the ocean. How can he tell which way to go to get to the surface?

B10

Will a candle burn in a spaceship, where everything is weightless?

GETTING TO KNOW...

The chemical elements

Just like the old-timer who's "been around," many an element has a story to tell ...

by Sheldon Lee Glashow

Phosphorous

The first known discoverer of a chemical element is the German alchemist Hennig Brand. In 1669 he prepared the "light-bearing" element by distilling a mixture of solid and liquid excrement while trying to find a liquid that would turn silver into gold. Instead, he found a pearly-white waxy stuff that glowed in the dark and could light a pipe. Ironically, Brand's home town of Hamburg was virtually destroyed by phosphorous bombs during World War II.

Mendelevium ($Z = 101$)

Mendelevium is an artificial element first made in California in 1955 and named after the pioneer of the periodic table. (A.N. Znoiko in the USSR suggested the name for element number 97, whose properties he had predicted, but it was named berkelium after the city of its discovery.) Today, there are 21 such artificial elements. The first one, technetium ($Z = 43$), was discovered in Italy in 1936, while the last three ($Z = 107–109$) were made in Germany in the 1980s. Plutonium is needed to make bombs that must never be used.

The rare earths ($Z = 57–71$)

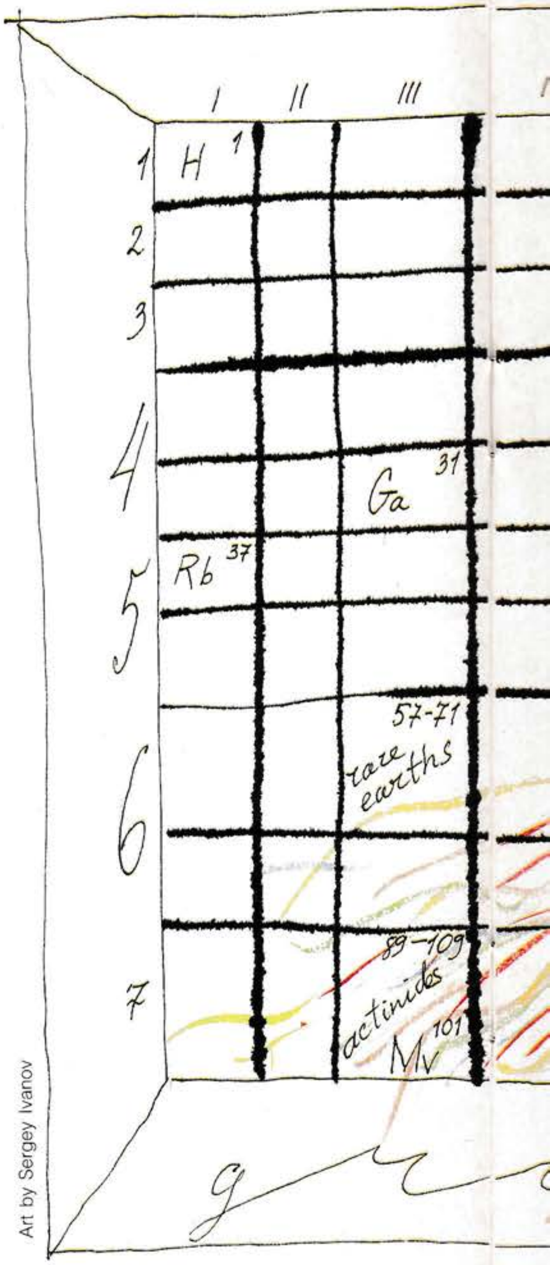
The periodic behavior of the elements reflects the sequential filling up of electron shells. These 15 elements are chemically similar because they have two valence electrons but varying numbers of electrons in an inner so-called 4*f*-shell. Their existence was first hinted at by the work of the Finnish chemist Gadolin in the very same year that Lavoisier died on the guillotine. Rare earths are not so rare and are used in the phosphors of television tubes. A similar family, the actinides ($Z = 90–103$), corresponds to the completion of the inner 5*f*-shell.

Gallium

When Mendeleyev proposed his table, he left three vacant spaces for elements not yet discovered and predicted their chemical and physical properties. Gallium was found in 1875, scandium in 1879, and germanium in 1886. Their names suggest where they were found, and they had just the expected properties. Ga is a rare byproduct of aluminum manufacture and an essential component of up-to-date semiconducting devices. Soviet and American scientists now collaborate in an experiment using 40 tons of Ga (borrowed from the Soviet strategic stockpile!) to study solar neutrinos.

The platinum metals

The platinum metals form a rectangle in the table: ruthenium, rhodium, palladium, and below them osmium, iridium, platinum. They are chemically similar and found together. Platinum metallurgy was developed by American Indians in pre-Columbian times. Along with coffee, tobacco, maize, and potatoes, Pt is a



Art by Sergey Ivanov

New World gift. Four related metals were found in South American Pt ores by two British chemists and friends in the year 1803. Ruthenium was isolated and studied by the Russian chemist Karl Karlovich Klaus. Curiously, all six metals of the Pt family are laboriously separated from the same rocks only to be recombined with one another to make special alloys for things like dental fillings, pen points, and precision instruments.

Argon

Lord Rayleigh measured the density of nitrogen from air and from ammonia and got different answers. He and William Ramsay solved the puzzle in 1894 by showing that 1% of air is a new element they called argon. They went on to find several other noble gases—chemically inert new elements that did not fit into the original table. Mendeleyev reluctantly added an extra column.

Rhenium

Rhenium is the last long-lived element to be found. It was discovered by a collaboration of three German chemists in 1925, two of whom subsequently (consequently?) married one another.

Hydrogen

In 1745 the Russian poet-scientist M.V. Lomonosov wrote, "On solution of any nonprecious metal in acid, there emerges an inflammable vapor that is nothing else than phlogiston." It was hydrogen, whose atom is the simplest of all. Much later, Niels Bohr computed its spectrum with his quantum rules. The structure of all the atoms and the meaning of the periodic table was to be explained by quantum mechanics.

Selenium

Selenium is one of nine elements named after heavenly bodies (the Sun, the Moon, five planets, and the asteroids Ceres and Pallas). A tiny bit of Se is essential to our diet, but too much is toxic. In South Dakota, cows grazing on plants that concentrate Se develop "alkali disease." Se powder smells like rotting horseradish.

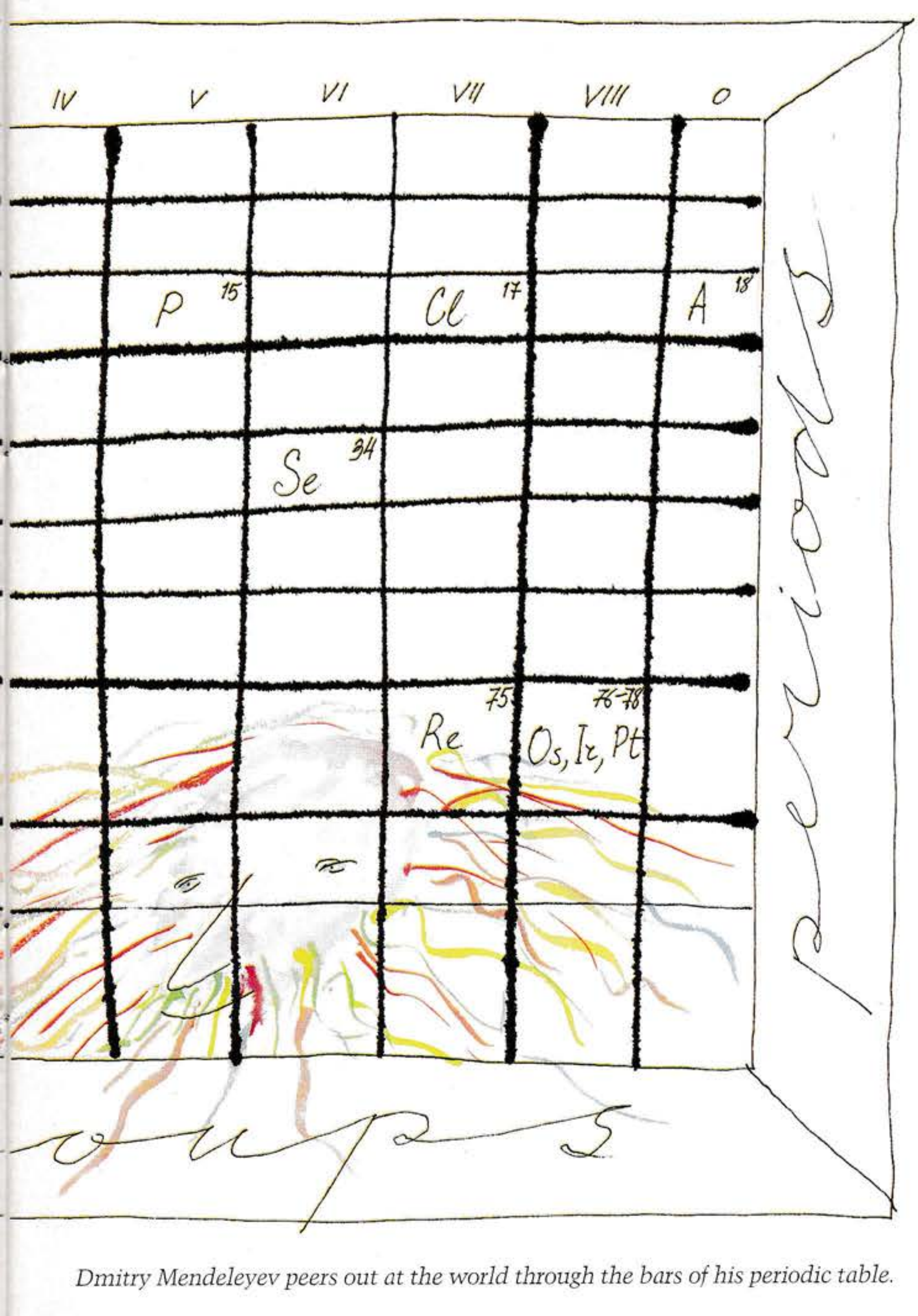
Chlorine

Lavoisier taught us that burning is oxidation. He gave oxygen a name meaning "acid maker" because he thought it was in all acids. Humphry Davy proved him wrong by showing that "muriatic acid" is O-free HCl and that Cl is an element. Fluorine, chlorine, bromine, and iodine are halogens—reactive nonmetals whose atoms have almost complete valence shells.

Rubidium

Each element emits light of definite colors when heated. These appear as bright lines when the light passes through a prism. Many elements were discovered this way and given names from Greek or Latin for the colors produced: rubidium (ruby-red), thallium (bud-green), and cesium (sky-blue). Helium's lines were first seen in sunlight, which is why it was named for the sun.

Sheldon Lee Glashow *is the Higgins Professor of Physics and Mellon Professor of the Sciences at Harvard University.*



Dmitry Mendeleyev peers out at the world through the bars of his periodic table.

Temperature, heat, and thermometers

What does it mean to "measure temperature"? Why are there so many thermometer scales? And how is heat related to temperature?

by A. Kikoyin

THE SHORT LIST OF PHYSICAL notions we become acquainted with before we begin to read, let alone study physics, includes temperature. In infancy we learn that the words hot, warm, and cold that describe our sensations mean different values of temperature; that in summer the temperature is high, in winter it's low. You probably can't recall when you became aware that a healthy person's temperature is "98-point-6" and that a doctor should be called if your temperature is much higher. You've known that "all your life."

Since temperature is so familiar to us, we're not usually aware of the special properties that distinguish it from other familiar notions like length, mass, or volume. But the difference is crucial.

If we take ten rods, each one meter long, and put them end to end so as to form one rod, we get a rod ten meters long. Similarly, if we take ten bodies, each with a mass of 1 kg, and lump them together, we get a body with a mass of 10 kg; and so on. But if we combine ten bodies, each at a temperature of 20°C, we end up with a body whose temperature is 20°, not 200°C, because the temperatures of bodies don't add up when we put them together, unlike their lengths, volumes, masses, and so on. A temperature of 100 degrees is not the sum of temperatures of one degree each, just as a man of 25 is not the same as 25 one-year-old babies!

So the main thing about temperature is that it's not an additive quantity, and this peculiar property will affect how we measure it. To measure the length of a body we have to compare it to another body that's considered the unit length; to determine the mass of some body we have to compare it to the unit mass. This is because a body's length and mass are equal to the sums of the lengths and masses of its parts. But temperature can't be measured this way.

How *is* temperature measured?

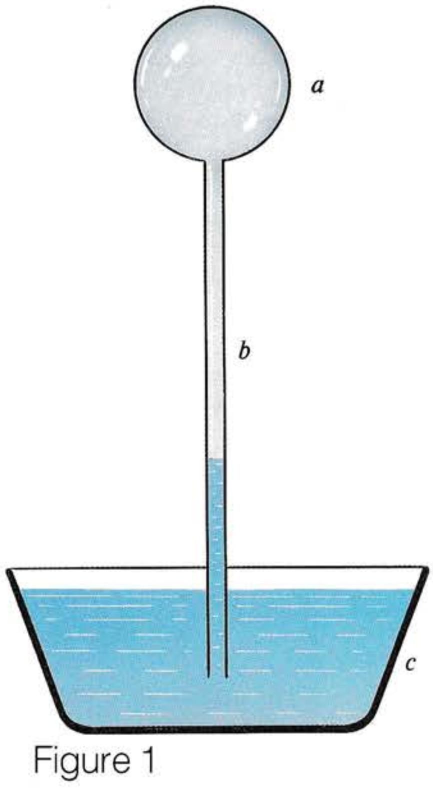
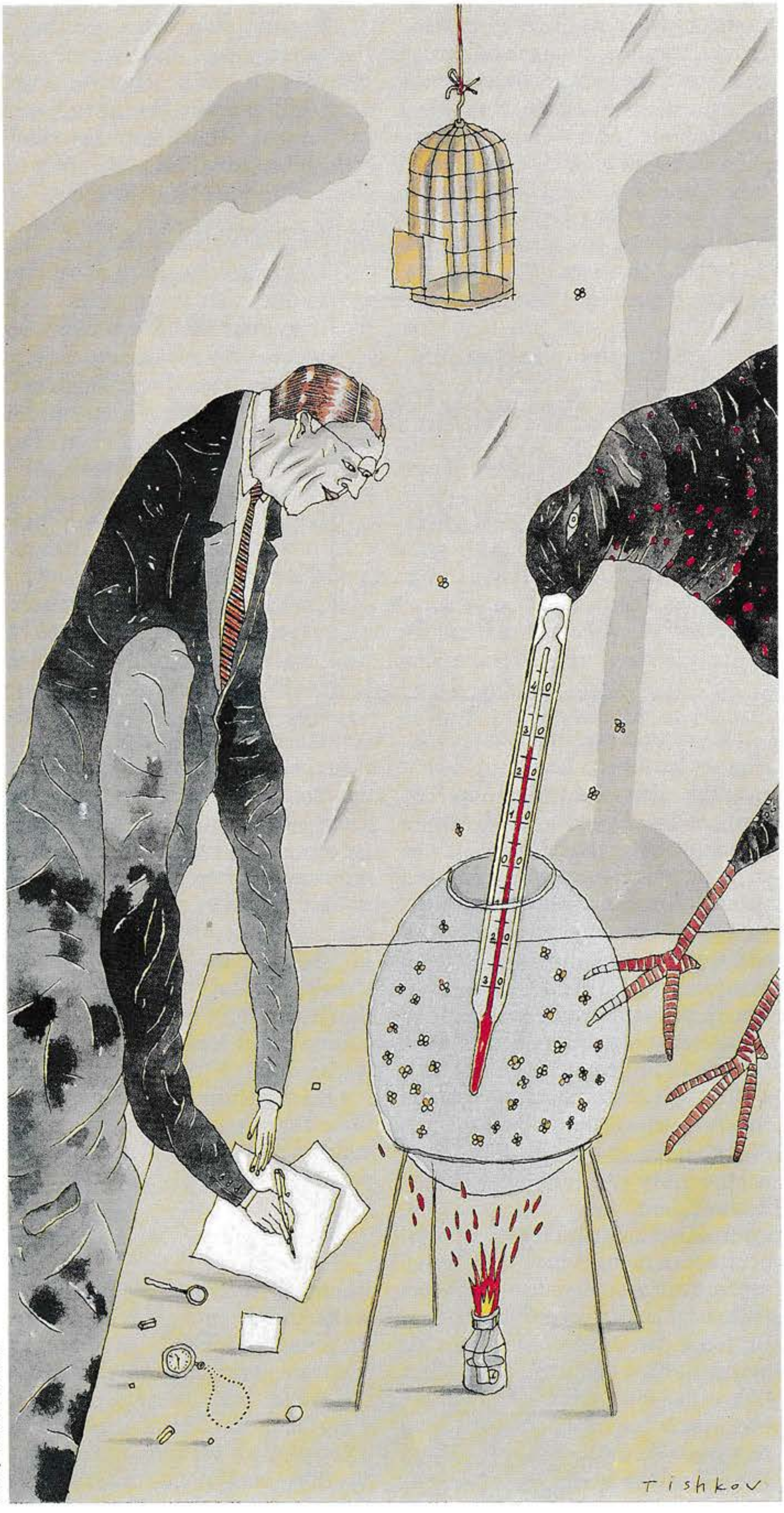


Figure 1

A bit of history

Around 1592 Galileo devised the first instrument for measuring temperature (although the word "thermometer" didn't appear in print until 1624). Galileo's method for measuring temperature, shown schematically in figure 1, is essentially the same as that used nowadays. A small glass bulb (a) is fused to the top of a long, narrow tube (b) that is open at the bottom. The bulb is warmed by somebody's hands and the lower end of the tube is dipped in a container of water (c). As the bulb cools to the temperature of the surrounding air, the level of the water in the tube rises above that of the water in the container.

Galileo's thermometer is obviously based on the fact that the volume of the gas in the bulb depends on its temperature, so that one can judge a change in temperature by the change in the gas's volume. But the instrument has no scale, so we can't assign a numerical value to the temperature. It's not really a thermometer, it's a "thermoscope"—it *shows* temperature



Art by Leonid Tishkov

but it doesn't *measure* it. It took almost 150 years to come up with a scale.

The important point for now is this. Galileo's thermometer is based on an idea we still find useful—that you don't measure temperature directly, you measure a quantity that is dependent on temperature. The quantity used in Galileo's thermometer was the volume of a gas. In a modern mercury thermometer, the quantity that depends on temperature, and whose change indicates a change in temperature, is again volume—but this time the volume of mercury, not a gas. For this purpose we could also use other physical quantities—for instance, the pressure of a gas (at constant volume), the length of a solid rod, or the electrical resistivity of a metal.

A physical law that can't be discovered without a thermometer

The first crude thermometers and even thermoscopes led to the discovery of one of the most important physical laws, the law of thermal equilibrium. The date of its discovery, which no scientist has claimed, is unknown, perhaps because it seemed self-evident. The law asserts that any isolated system of bodies eventually arrives at a state in which all its components have the same temperature. This state is called the state of thermal equilibrium.

Obviously, the law of thermal equilibrium could be discovered only after the invention of the thermometer. On the other hand, temperature measurement itself by means of a thermometer is based on this law. After all, a thermometer is a body at some temperature that indicates its own temperature. So in order to assess a body's temperature with a thermometer, this body must be in thermal equilibrium with the thermometer—their temperatures must be the same if the thermometer is to indicate the body's temperature in addition to its own. That's why you always have to wait for some time before reading a temperature—you're waiting for thermal equilibrium to be established.

Some more history

To recap: the thermoscope appeared at the end of the 16th century and became the thermometer in the middle of the 18th. But what does the thermometer indicate—what *is* temperature? It took scientists another hundred years to find the correct answer to this question.

Temperature is a quantity that characterizes the thermal state of a body. We say that cold and hot bodies have different temperatures. So the nature of temperature comes down to this: how is a cold body different from a hot body?

The first answer to this question was given by Galileo himself. He started from the easily observed fact that when a cold body is placed close to a hot one, the hot body cools while the cold one becomes warmer. Galileo concluded that something travels from the cold body to the hot body (though we might just as well suppose that something travels in the opposite direction). He assumed that it is a specific substance, and in the 17th and 18th centuries most scientists accepted that point of view. The substance was called the "caloric fluid."

According to caloric theory, a hot body differs from a cold one in that it contains more caloric fluid. When thermal equilibrium is established, this caloric fluid has passed from the hot body to the cold one. So in this view, a body is a mixure of two substances: the material of the body itself (for example, water, copper, iron, glass) and the caloric fluid in it. This is where the name for the unit of temperature, a "degree," comes from—the same unit was used to indicate concentrations of aqueous solutions.

This concept of temperature was generally accepted for some 200 years. And that's what they talked about—"degrees of warmth."

At the same time there was another theory, based on the fact that a body can be warmed up by mechanical motion. One of the founders of this theory was the Russian scientist Lomonosov, who wrote: "It is well known that heat is generated by motion; hands are warmed up by their mutual friction; wood can be ignited by rubbing; sparks appear from strokes of flint on steel; iron becomes hot from hammering it with strong, rapid strokes." This led to the conclusion that heat is not a substance but the motion of the "imperceptible particles" (as they called them then) that constitute all bodies.

Galileo's motto:
"To measure that which is measurable,
and to strive to make measurable
that which is not yet so."

Experiments carried out in the 18th century led most scientists to reconsider the caloric fluid theory. In 1760 the Scottish physicist and physician Joseph Black showed that when the same amount of heat is added to equal masses of different substances, different changes in temperature result. If temperature were the concentration of caloric fluid in a body, the acquisition of the same amount of heat by equal masses should produce the same change in temperature. So Black's results were incompatible with the theory of caloric fluid. Black discovered that different substances have what we now call different "specific heats."

In 1764 Black observed that temperature remains constant while ice melts, even though melting requires that a large amount of heat be added to the ice. From the time of Black's experiment this quantity of heat has been called the "latent melting heat." Likewise, a certain amount of heat is removed when water freezes, which also takes place at a fixed temperature. Obviously, heat could not be absorbed or released by a body without a change in temperature if temperature were the concentration of heat in a body.

The true nature of temperature—that quantity whose meaning remained obscure for such a long time—became clear only after the kinetic theory of matter was worked out. To understand how, we have to get a handle on the basic ideas of this theory.

Molecular chaos and its laws

The kinetic theory of matter assumes that any physical body consists of small particles—molecules—that are in constant chaotic motion. Complex forces of attraction and repulsion act between molecules. But for gases under ordinary conditions, these forces are small, and we may even visualize a gas in which any interaction between molecules is absent. Such gas, owing to its purely theoretical nature, is called an "ideal gas."

An ideal gas consists of a huge number of molecules moving chaotically every which way in a container at speeds of hundreds of meters per second, colliding with each other and the walls of the container. The system as a whole, however, is subject to very strict rules. Since the interaction between the molecules of an ideal gas can be neglected, the rules can be stated in a rather simple form. In particular, using the laws of mechanics we can find the pressure—that is, the force with which the gas acts on a unit area of the container's walls.

One can show by certain calculations that in a vessel of volume V containing N molecules of gas, the pressure on the walls equals

$$p = \frac{2}{3} \cdot \frac{N}{V} \cdot \overline{E}, \qquad (1)$$

where $\overline{E}$ (which equals $m\overline{v}^2/2$) is the mean kinetic energy of chaotic motion assigned to one molecule of gas. Equation (1) indicates that gas pressure is equal to 2/3 of the mean kinetic energy of the chaotic motion of molecules per unit volume (since N/V is just the number of molecules per unit volume).

Real gases require more sophisticated calculations, but under appropriate conditions equation (1) can still be used. It gives an approximate relation between p and $\overline{E}$ that is accurate enough for small N/V and $\overline{E}$. The formula can be used for pressures of about 1 atmosphere and lower.

At first glance all this has nothing to do with temperature, which doesn't even appear in equation (1). To see what's really going on, let's turn again to temperature measurement.

Scales: feast after famine

The first thermometers used for practical purposes were liquid thermometers made by a group of scientists in Florence. Later they were manufactured in other countries. Different liquids were used, but the most common ones were alcohol and mercury (sometimes oil).

The liquid thermometer consisted of a thin vertical glass tube that ballooned out at the bottom in the form of a small bulb. The bulb and lower part of the tube were filled with liquid. In addition to being functional, the old Florentine thermometers were quite beautiful. If you saw them, you might even consider them works of art—such a refined artistic approach was taken in creating these scientific instruments.

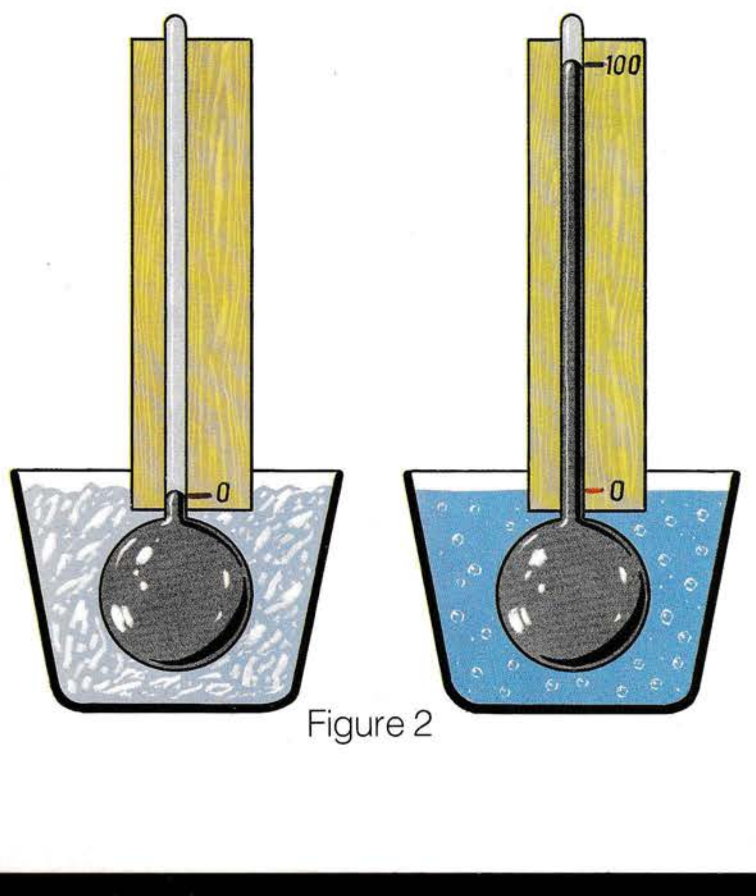


Figure 2

Various methods were used to construct thermometer scales. Every manufacturer or designer of thermometers worked out a different one. By the end of the 18th century there were about twenty different thermometer scales. Three of them (too many, in fact) have survived to the present day.

The three survivors are variations on a single theme. Earlier in the century the German glassblower and amateur physicist Daniel Gabriel Fahrenheit and the Swedish astronomer Anders Celsius came up with a principle for constructing thermometer scales that has since been generally accepted. It's based on the use of two reference points—thermal states that can be characterized as constant. The points they chose for this purpose are the temperature of melting ice and that of boiling water at atmospheric pressure. (The melting point of any solid substance and the boiling point of any liquid at a given pressure are also constant, but water and ice are more readily available.)

In 1742 Celsius proposed the following prescription for dividing up the thermometer scale. A thermometer is put in contact with melting ice. Once heat equilibrium is established, the height of the liquid in the thermometer is marked. Then boiling water is used instead of ice, and the new level is marked with a number that differs from the first by 100. The space between the two marks is divided into one hundred regular intervals, each interval corresponding to the change of one degree in temperature (fig. 2). Curiously enough, Celsius took the temperature of boiling water as 0 and that of melting ice as 100. Eight years later, however, this scale was inverted and has stayed that way for more than 200 years.

Even before Celsius, in 1724 Fahrenheit manufactured thermometers in which two reference points, melting ice and boiling water, were taken as 32 and 212, respectively, and the interval between them was divided into 180 regular intervals (degrees). Like Celsius, the contemporary French scientist René-Antoine Ferchault de Réaumur took the temperature of melting ice as 0, but according to Réaumur's scale water boils at 80.

So we see there was a lot of confusion when it came to selecting temperature scales. The number of intervals between the two fixed points was arbitrary, as were the values of these reference points themselves. In fact, there's no reasonable argument in support of assigning the temperature 0 to melting ice—as if melting ice had no temperature!

Not only that, there's an assumption hidden in our division of the temperature scale into equal parts—80, 100, 180, or whatever. We assume that the volume of the liquid in the thermometer is exactly proportional to its temperature.

If we denote the volume of the liquid at the temperature of melting ice as V_0 and its volume at the temperature of boiling water as V, and these temperatures themselves as t_0 and t, division of the temperature scale into equal parts means that

$$\frac{V-V_0}{t-t_0}=c,$$

where c is a constant. If we assume $t_0 = 0$, then $V - V_0 = ct$, or $V = V_0 + ct$.

Can we verify that volume depends on temperature as a linear function? Obviously we can't, because we'd have to use a thermometer in the verifying experiment, and in constructing the thermometer we've already assumed that volume is linearly dependent on temperature.

There's an old story about a seaport in which a gun was fired every day, exactly at noon, so that ship captains leaving the port could set their chronometers (which were used to determine longitude during the voyage) to the time of the shot. One of the captains wanted to know how confident he could be that the gun was really fired at midday. He found out that the artillery man determined the time with the help of a local watchmaker's "very exact watch." And the watchmaker told the captain he set his "very exact watch" by the shot in the port. Obviously, under these circumstances, you can't judge whether the watch is of good quality or the gun is fired exactly at noon. Similarly, when we use a thermometer whose scale is constructed on the assumption that the volume of its liquid is proportional to the its temperature, we can't tell whether that assumption is valid.

Equally important, in measuring temperature we have to allow for the fact that the actual readings depend on the liquid used in the thermometer or on some of its other physical properties. We need a standard thermometer for grading all the others, so as to make their readings compatible. How do we solve this problem? At present the standard thermometer is the so-called "gas thermometer of constant volume."

Gas thermometers and an absolute temperature scale

From experimentation we know that the pressure of hot gas is greater than that of cold gas. In the gas thermometer, then, pressure at constant volume is the quantity that indicates the temperature. The thermometer shown in figure 3 consists of a container A filled with an "ideal gas" (that is, any gas at low pressure) and a manometer M attached to the container to indicate the pressure. If we put the container in melting ice, then in boiling water, and measure the pressure at these temperatures, we find that the pressure at the boiling point is 1.3661 times greater than the pressure at the melting point. So we have

$$\frac{p}{p_0} = 1.3661, \qquad (2)$$

where p is the pressure at the temperature T of boiling water and p_0 is the pressure at the temperature T_0 of melting ice. Following the tradition of the Celsius scale, let's say

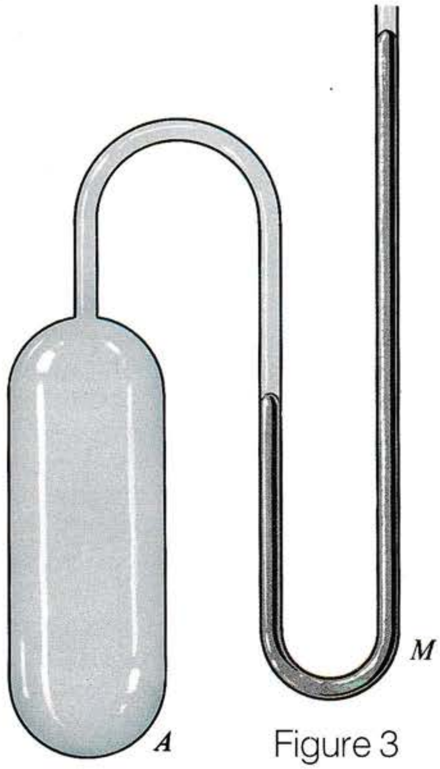


Figure 3

$$T - T_0 = 100. \qquad (3)$$

The difference in pressure at the temperatures of boiling water and melting ice is divided into 100 equal intervals—that is, we still assume the linear dependence of temperature on pressure at constant volume. Of course, we can't justify the assumption, just as the captain couldn't judge if the watch was correct by the gunshot or the gunshot was on time by the watch. The assumption of linear dependence is simply the basis for the method of measurement.

Now, with the gas thermometer we have no need to assign 0 to the temperature of melting ice. We can calculate it. In fact, if temperature is directly proportional to pressure, then

$$\frac{p}{p_0} = \frac{T}{T_0}. \qquad (4)$$

Since the ratio on the left side equals 1.3661, this is the same as $T/T_0 =$ 1.3661, or $T = 1.3661T_0$. Substituting this value for T in equation (3), we get $1.3661\,T_0 - T_0 = 100$, which gives us

$$T_0 = \frac{100}{0.3661} \cong 273.15.$$

We can see the difference between this new scale and the old Celsius scale (fig. 4). According to this scale, the temperature of melting ice is not 0 but 273.15 degrees, and zero temperature is 273.15 degrees lower than the temperature of melting ice. This zero temperature is called "absolute zero." It's the temperature at which an ideal gas must have a pressure equal to zero, if such a temperature could be obtained and the gas would remain in the gaseous state. Since gas pressure can't be negative, the temperature in this scale can't ever be negative.

Perhaps you've already recognized this new scale in my rough sketch. It's the Kelvin (or absolute) scale. The temperature in this scale is called "absolute temperature," whose standard notation is T, and it's measured in degrees Kelvin. The temperature of melting ice is 273.15°K, that of boiling water 373.15°K, and so on.

But for most practical purposes, the Celsius scale reigns throughout the world (although most Anglo-Saxon countries still use Fahrenheit). The temperature in this scale is denoted as t and is expressed in degrees Celsius (abbreviated °C). According to this scale, the temperature of melting ice is 0°C, that of boiling water 100°C, and so on. Obviously, t°C = $(T - 273.15)$°K. Those who use the Kelvin scale are almost invariably physicists.

And now, finally, we're in a position to grasp the real meaning of temperature.

So what *is* temperature?

According to the method of measuring temperature I've just described, the pressure of gas of mass M, consisting of N molecules in volume V, is proportional to its absolute temperature T. We can write equation (4) in the form

$$\frac{p}{T} = \frac{p_0}{T_0}, \qquad (5)$$

which suggests that the ratio of pressure to temperature of a gas at constant volume is constant. On the other hand, we have equation (1) for gas pressure:

$$p = \frac{2}{3} \cdot \frac{N}{V} \cdot \overline{E}\,.$$

Substituting this value of p in equation (5), we get

$$T = \frac{2}{3} \cdot \frac{N}{V} \cdot \frac{T_0}{p_0} \cdot \overline{E}\,. \qquad (6)$$

This equation is true for a gas in a closed container at constant volume. So the number of gas molecules N is constant, and the ratio T_0/p_0 is also constant, as we saw earlier. As a result, the coefficient at $\overline{E}$ in equation (6) is a constant for any gas, and we have

$$T = \frac{2}{3} A \overline{E}\,, \qquad (7)$$

where $A = N/V \cdot T_0/p_0$ (a constant). Equation (7) means that the absolute temperature of a gas is the mean kinetic energy of chaotic motion of one molecule, measured in degrees Kelvin and not in energy units (joules). The coefficient A is a factor that relates the energy and temperature units. The situation is similar to the one we encounter when working with different units of length. For example, to find a length in inches, we need to know that 1 meter is about 40 inches. Equation (7) is usually written in the form

$$\overline{E} = \frac{3}{2} kT\,, \qquad (8)$$

where

$$k = \frac{1}{A} = \frac{V}{N} \cdot \frac{p_0}{T_0}\,. \qquad (9)$$

The coefficient k is called the "Boltzmann constant" after the Austrian physicist Ludwig Boltzmann (1844–1906).

From equation (9) we can infer how to obtain the numerical value of the Boltzmann constant. To do this, we take a container of volume V and fill it with gas of mass M (the mass can be determined by weighing). Then we put the container in melting ice (whose temperature T_0 is 273.15°K) and measure its pressure with a manometer. If we know the mass M, we can easily determine the number of molecules N. In fact, if the molar weight of the gas in grams equals μ, the number of moles of gas in the container equals M/μ. And since each mole contains N_A molecules (Avogadro's number), the number of molecules in the container is $N = (M/\mu)N_A$. So we can determine the Boltzmann constant k if we know the mass M, the molar mass μ, the volume of the container V, and the pressure p_0 of the gas at temperature T_0.

Such measurements have been performed many times, and all of them give the same value for the Boltzmann constant:

$$k = 1.38 \cdot 10^{-23}\ \text{J/°K}.$$

You see how small k is. So we conclude that the mean kinetic energy of chaotic motion of one molecule is a tiny quantity.

It's precisely this energy that determines temperature. At 1°K, the mean kinetic energy of a molecule $\overline{E}$ equals

$$m\overline{v}^2/2 \;=\; \frac{3}{2} \cdot 1.38 \cdot 10^{-23}$$
$$\cong 2 \cdot 10^{-23}\ \text{J/molecule}.$$

This equation can be viewed as the relation between one degree Kelvin and one joule/molecule.

Before we leave the subject of temperature, we should clarify the relationship between temperature and heat, two concepts that have been considered virtually identical through the centuries.

We know now that *heat* is the energy of chaotic motion of particles and that this energy can be transferred from one body to another during heat exchange. So heat isn't a quantity that characterizes the state of a body—we can't say it's contained in the body. But *temperature* characterizes the state of a body because it's determined by the mean kinetic molecular energy. There's no essential relation between temperature and heat. We can only say this: *If two bodies have different temperatures, the body at the higher temperature gives heat to the other.* Temperature is the quantity that determines whether a body gives heat to other bodies or gets it from them.

And that's just how the great Scottish physicist James Clerk Maxwell (1831–1879) defined temperature way back when.

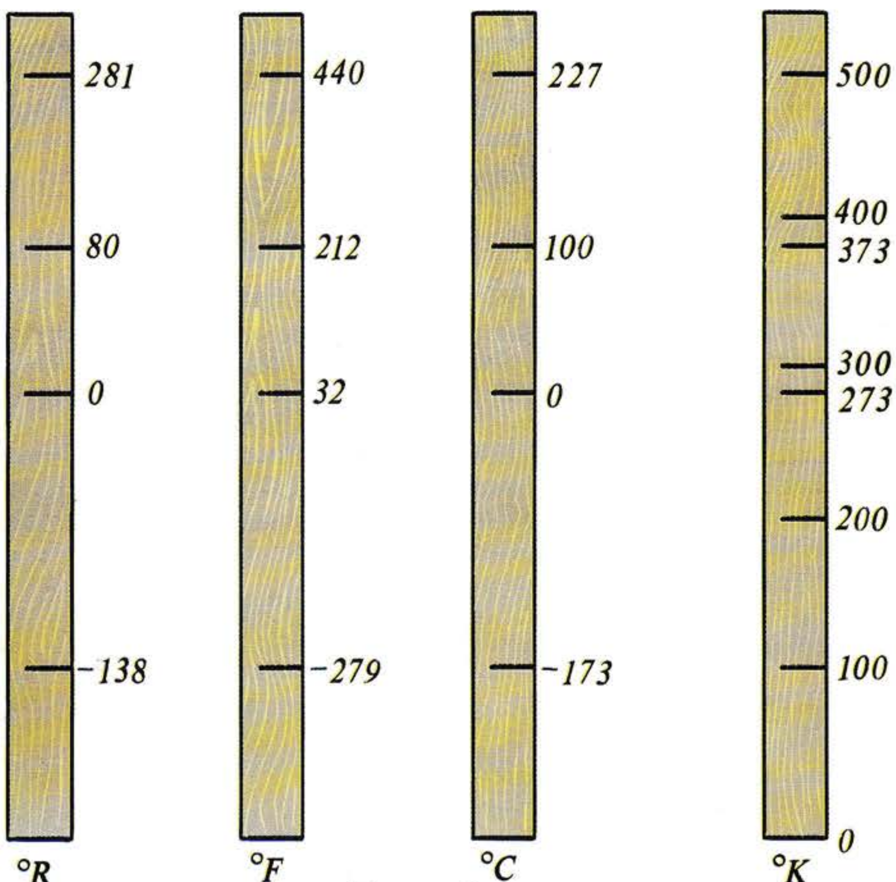


Figure 4

CONTINUED ON PAGE 49

HOW DO YOU FIGURE?

Challenges in physics and math

Anchurian candidate & seven dwarfs, river raft & spaceship, circles & cubes, expanding polygons & merging lines, a pan of water & a string of polarized digits— that is to say, something for everyone

Math

M6

Three equal circles. Three circles with the same radius r all pass through point H (fig. 1). Prove that the circle passing through the points where pairs of circles intersect (that is, points A, B, and C) also has the same radius r.

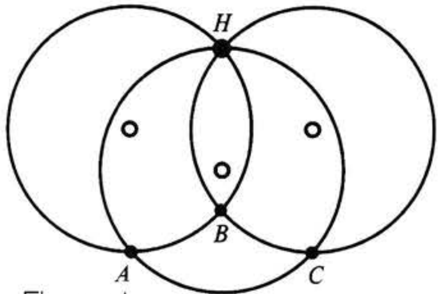


Figure 1

M7

Double parity. Prove that if each of the numbers $x_1, x_2, ..., x_n$ is equal to +1 or −1, and $x_1x_2 + x_2x_3 + ... + x_{n-1}x_n + x_nx_1 = 0$, then n is divisible by 4. (A. Leontovich)

M8

Two similar polygons. When the sides of a convex polygon are moved outward by the same distance 1, they fall on the corresponding parallel sides of a larger similar polygon. Prove that circles can be inscribed in these polygons. (N. Vasilyev)

M9

"Democratic" election. In the country of Anchuria, ruled by President Miraflores, a presidential election is coming around again. There are exactly 20 million voters in the country and only one percent of them (the regular Anchurian army) supports Miraflores. Naturally, Miraflores wants to get reelected, but on the other hand he wants the elections to appear democratic. By "democratic election" Miraflores means the following.

All the voters are divided into equal groups, each group is again divided into a certain number of equal smaller groups, then these smaller groups are divided into still smaller equal groups, and so on. The smallest groups choose representatives—known as "electors." These electors choose representatives of the next larger groups, and so on. Finally, the representatives of the largest groups elect the president. The election proceeds at all stages by majority vote—in a split vote, the opposition wins. Miraflores has the right to divide the electors into such groups as he wishes and can instruct his supporters how to vote. Will he able to organize such a "democratic election" and get reelected? (32nd Moscow Math Olympiad)

M10

Seven dwarfs. Seven dwarfs sit around a circular table waiting for Snow White. Each of them has a big cup, and some of the cups contain milk. The first dwarf pours out all his milk into the other six cups, dividing it into six equal portions. The dwarf to his right then does the same. The next dwarf follows suit. And so they continue until the seventh dwarf pours out all his milk into the other six cups (dividing it into equal portions). Then it turns out that each of the dwarfs has exactly as much milk as when they started. Find the amount of milk originally contained in each cup if the total amount was 3 liters. (V. Gutenmacher)

Physics

P6

Imminent collision. The pilot of a spaceship moving at velocity v = 1 km/s notices an asteroid of diameter d = 7 km straight ahead at a distance of l = 8.5 km. The astronaut immediately switches on the emergency engines, which in a negligibly small interval of

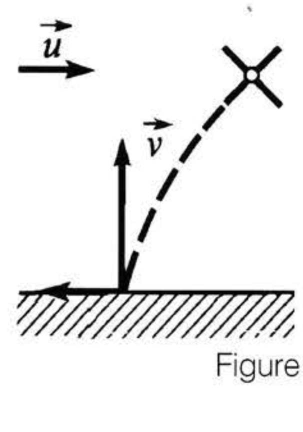


Figure 2

time impart an additional velocity of Δv = 300 m/s to the ship. This additional velocity can be directed in any way the astronaut chooses. Can a collision be avoided? (A. Andrianov)

P7

Raft on the river. A wooden raft is pushed perpendicularly from a riverbank such that its initial velocity is **v**. The raft's trajectory is shown in figure 2. The X on the trajectory shows where the raft will be at time t after the operation began. Assuming that the velocity of the current is constant and equal to **u**, graphically plot where the raft will be at times $2t$, $3t$, $4t$ … . (I. Poterayko)

P8

Cooling water. A pan contains 1 liter of water that can't be brought to a boil by means of a 100-W heating element. How long will it take for the water to cool 1°C when the heating element is turned off? (A. Zilberman)

P9

The cube's magnetic field. The current I flowing in a circuit formed by the four edges of a cube creates a magnetic field of induction B_0 in the center of the cube (fig. 3). Find the value and direction of the magnetic field created in the center of the cube by the current I flowing in the circuit formed by the six edges displayed in figure 4. (M. Tsypin)

P10

Merging rails. Imagine you're down at the railroad tracks, looking far off down the line. Calculate the distance at which the rails seem to come together. (It's assumed you understand the physics of the observed phenomenon and can estimate the other magnitudes involved.) (P. Zubkov)

SOLUTIONS ON PAGE 59

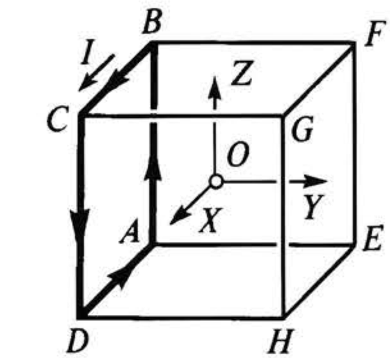


Figure 3

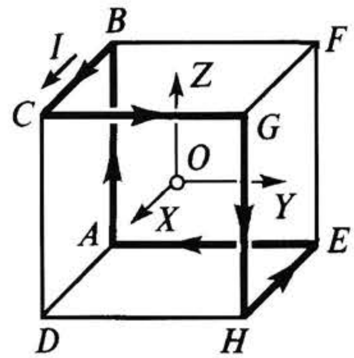


Figure 4


С. Иванов
H
D
R

The geometry of population genetics

The power of mathematical modeling is applied to the phenomenon of hereditary change from generation to generation

by I. M. Yaglom

MATHEMATICS OCCUPIES a privileged position among the sciences. Natural sciences such as physics or chemistry study the real world we live in; social sciences study human society and its features. These sciences have in common the fact that they acquire their knowledge from experience. In contrast, mathematics studies such abstractions as irrational numbers (for example, $2^{1/2}$) and things like lines, which by definition have no width but only length. Mathematical knowledge proceeds from the construction of abstract concepts. A mathematical concept can be defined by giving, as is done in geometry, a full list of axioms that adequately determine its properties.

The outstanding role played by mathematics is largely due to the fact that it can be applied to both the natural and social sciences—that it can, in fact, "model" physical and human phenomena. Mathematical modeling consists of establishing a correspondence between specific features of a phenomenon by using mathematical methods. We'll use the example of heredity—more precisely, population genetics—to illustrate this unique power of mathematics.

Art by Sergey Ivanov

Heredity

Following the classical theory proposed by Gregor Mendel,[1] we'll focus our attention on a certain characteristic, or "trait," transmitted from one generation to the next and therefore called hereditary. The fur color of guinea pigs, for example, is such a trait. The physical unit of inheritance is the gene. A trait is determined by a pair of genes; either gene can be one of two types, G and g. The first one, denoted by the capital letter G, gives the dominant form of the characteristic (in our example, the color black); the other, denoted by g, gives the recessive form (the color brown). The words "dominant" and "recessive" mean that the outward appearance of a hybrid guinea pig, H, with a genetic makeup determined by the combination of genes (or "genotype") Gg, is the same as that of a dominant guinea pig, D, with the genotype GG—that is, hybrid guinea pigs are black. In contrast, recessive guinea pigs, R, with the genotype gg, look different from the dominant and hybrid ones—they're brown.

The law of transmission of traits can be stated as follows: an offspring receives one of its genes from the father and another, independently of the first one, from the mother; either of the father's genes (and either of the mother's) has the same odds of being passed on to the offspring. So a pair of black hybrid guinea pigs can give birth to brown guinea pigs—in fact, one quarter of their offspring will be brown, if there are enough of them.

[1] The Austrian natural scientist Gregor Johann Mendel (1822–1884), the abbot of a monastery in what is now Czechoslovakia, discovered the fundamental laws of heredity that underlie the modern science of genetics.

Problem

1. Find the proportion of dominant, hybrid, and recessive offspring among a sufficiently large number of descendants of each of six possible mating couples: (D,D), (D,R), (R,R), (H,D), (H,R), and (H,H).

Biological populations and geometric points

A biological population is a sufficiently large group of organisms that belong to the same species and form mating couples only within the group. Population genetics studies the changes, from one generation to the next, of the proportions d, h, and r, which represent the numbers of dominant, hybrid, and recessive organisms relative to the total population. So we can describe the population P by three nonnegative numbers (d,h,r) that are subject to the constraint

$$d + h + r = 1.$$

This simple observation suggests that we use the following device for the mathematical modeling of biological populations.

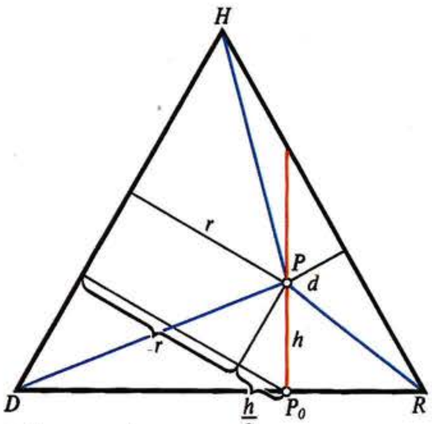


Figure 1

If the side of triangle DHR *is equal to* 2a (a = $1/3^{1/2}$ *since the altitude of the triangle is equal to* 1), *its area is given by the equation* $S_{DHR} = a = S_{DHP} + S_{RDP} + S_{HRP} = ad + ah + ar$. *Therefore,* $d + h + r = 1$.

Let's consider an equilateral triangle DHR whose altitude is equal to 1. It's important that, for any point P of the triangle, the sum of the distances from that point to sides HR, RD, DH is constant and equal to 1 (fig. 1). We'll call the numbers d, h, r coordinates of the point P with respect to triangle DHR and write $P = P(d,h,r)$, as for Cartesian coordinates. Their relation to the usual three-dimensional coordinates in space is illustrated in figure 2.

In this way we can assign to each population P—determined by the ratios d, h, and r of the dominant, hybrid, and recessive organisms—a point in triangle DHR, which we'll denote by the same letter $P = P(d,h,r)$. It's worth noticing that populations consisting only of dominant, hybrid, and recessive organisms correspond to points $D(1,0,0)$, $H(0,1,0)$, and $R(0,0,1)$, respectively.

Problem

2. Find the distance between points $P(d,h,r)$ and $P_1(d_1,h_1,r_1)$ by means of their coordinates with respect to triangle DHR (use figure 2).

The heredity of a population is determined by the composition of its total "gene pool." We can visualize this by looking at our triangle. Let's denote the relative portions of G and g genes by Γ and $\gamma = 1 - \Gamma$ for a certain population corresponding to point $P(d,h,r)$ in the triangle. It's easy to see that $\Gamma = d + h/2$, $\gamma = r + h/2$, and that Γ and γ are the d- and r-coordinates of the projection P_0 of point P onto side DR. Consequently, populations that have a fixed gene pool composition given by the ratio $\Gamma : \gamma$ $(\Gamma + \gamma = 1)$ correspond to points lying on the perpendicular to side DR passing through point $P_0(\Gamma,0,\gamma)$.

Now we've come to the main point of our discussion. Let's consider population P, corresponding to point $P(d,h,r)$ in triangle DHR, and population P', which represents the next generation of organisms bred by mating in population P. Population P' also corresponds to a point in triangle DHR. So we have a transformation of point P

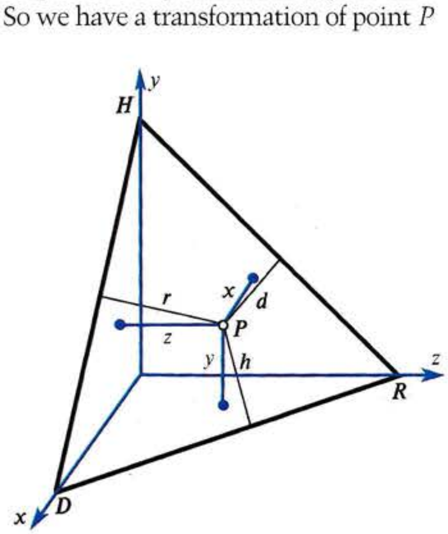


Figure 2

Population P(d,h,r) *corresponds to point* P *in space with coordinates* (x,y,z) = $(2/3)^{1/2}$(d,h,r). *All these points belong to equilateral triangle* DHR, *which is the intersection of plane* x + y + z = $(2/3)^{1/2}$ *with the positive octant* (x ≥ 0, y ≥ 0, z ≥ 0). *The distances from point* P *to the sides of triangle* DHR *are equal to* d, h, r.

into point P', which provides a mathematical framework for studying the transmission of inheritance from one generation to the next.

Examples

Interbreeding with dominant organisms. Let's suppose that organisms in population $P = P(d,h,r)$ interbreed only with the dominant ones, D, so that one of the genes of every descendant will be dominant. The second gene, inherited from parents belonging to population P, is chosen in two steps, in effect: first, through the random choice of a parent; second, through the random choice of one of the genes. It's important that all genes in the gene pool of the entire population have the same opportinuty to participate in the process so that the proportion of descendants with the second G gene is equal to the proportion Γ in the gene pool, while the proportion of g genes equals γ. In the first case dominant organisms are formed, and in the second—hybrid ones. So the ratios d', h', and r' of D, H, and R organisms in the next generation are

$$d' = \Gamma = d + h/2,$$
$$h' = \gamma = h/2 + r,$$
$$r' = 0.$$

Now let's consider the structure of the transformation $f_1: P \longrightarrow P'$ of triangle DHR that is generated by the equations given above. Comparing the coordinates of points $P'(\Gamma, \gamma, 0)$ and $P_0(\Gamma, 0, \gamma)$ (notice that P_0 is the projection of P onto DR—see figure 3), we see that point P' lies on side DH at a distance of $DP' = DP_0$ from vertex D. Let P_1 be the intersection point of lines P_0P and DH—that is, the projection of P onto line DH in the direction perpendicular to DR. (In what follows we'll simply say "in direction h.") It's evident that $DP_1 = 2DP_0$, or $DP' = DP_1/2$. In other words, to get point $P' = f_1(P)$ we first have to find point P_1, which is the projection of P in direction h onto DH, and then find the midpoint P' of line segment P_1D.

And so the inherited transformation f_1 we've been examining in this

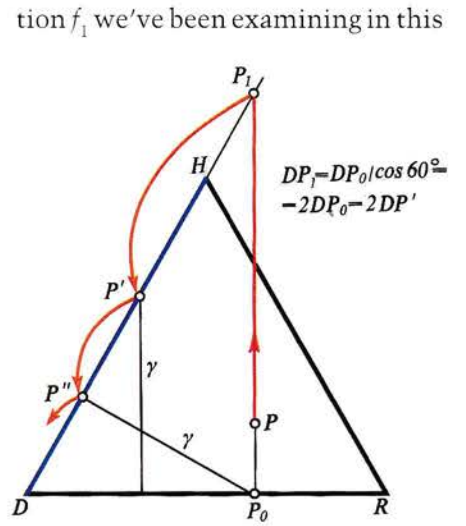


Figure 3

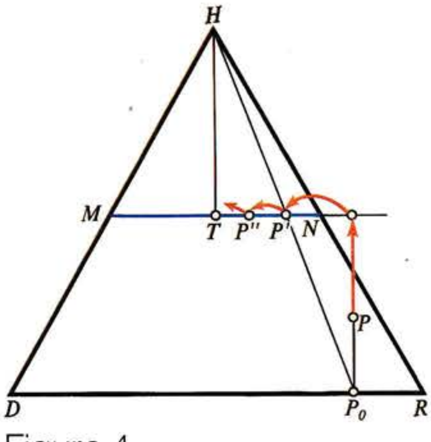


Figure 4

section is the result of two transformations performed sequentially: first, the parallel projection in direction h onto line DH; second, the similitude ("homothety") with center D and scale factor ½.

From this we can infer that the transformation turns the entire triangle DHR (the set of all possible populations) into one line segment—side DH (the set of all populations that don't contain recessive organisms). It's important that the only point that doesn't move (the "fixed point" of the transformation) is point D, which is what we'd expect from our understanding of genetics. Generations that follow P' as descendants of the original population will correspond to images of point P' after sequential similitudes with center D and scale factor ½, and they'll tend to come closer and closer to the pure dominant population D.

Interbreeding with hybrids. Now suppose the organisms in population $P = P(d,h,r)$ interbreed only with the hybrids H. Using arguments similar to those given above, we infer that half the descendants acquire the G gene from their parents and the other half the g gene; then in either half the proportion of descendants having G as the second gene equals Γ, and those having g as the second gene is γ. So for the next generation $P'(d',h',r')$ we have

$$\begin{aligned} d' &= \Gamma/2 = d/2 + h/4, \\ h' &= \Gamma/2 + \gamma/2 = 1/2, \\ r' &= \gamma/2 = r/2 + h/4. \end{aligned}$$

Comparing the coordinates of points $P'(\Gamma/2\tfrac{1}{2},\gamma/2)$ and $P_0(\Gamma, 0, \gamma)$, we see that point P' is obtained from P by the similitude with center H and scale factor ½ (fig. 4). This means the transformation f_2 that turns P into P' can be seen as a combination of (1) the projection in direction h onto the line joining midpoints M and N of sides DH and HR of our triangle and (2) the similitude with its center at midpoint T of MN and a scale factor of ½.

Looking at the construction given above, we can easily see that the hereditary transformation f_2 changes triangle DHR (the set of all populations) into line segment MN (the set of all populations exactly half of which are hybrids) and possesses a unique fixed point $T(1/4, 1/2, 1/4)$ (a stable population that preserves its composition). The sequence of generations P', $P'' = f_2(P')$, ... is depicted as points of line MN; each point is transformed into the next by the similitude with center T and scale factor ½. There's no need by now to make any projections. Consequently, populations P', P'', ... approach the stable population T.

Problem

3. Let f_{abc} be the transformation corresponding to the interbreeding of population $P(d,h,r)$ with a fixed population $P(a,b,c)$. (Here $a,b,c \geq 0$ and $a + b + c = 1$, of course.) For example, f_{100} and f_{010} are the transformations f_1 and f_2 discussed earlier. Find the formulas for the transformation f_{abc} and illustrate its geometrical meaning. Also, find the stable populations. What happens when multiple repetitions of transformation f_{abc} are performed?

Panmixia and the Hardy–Weinberg law

The hereditary transformations considered above are primarily of interest to teachers. Scientists in the field of population genetics mainly study hereditary transformations $P' = f(P)$ such that the population of descendants is generated by a random choice of mating couples. This is called "panmixia" (from the Greek words *pan*, "all," and *mixis*, "mixing" or "mating"). Obviously a random choice of mating couples and then of genes actually amounts to a random choice of gene pairs from the gene pool of population P. Then the frequency of GG pairs—that is, the proportion d' of dominant organisms in the next generation P'—equals Γ^2; the frequency of Gg and gG pairs (the proportion h' of hybrids) is $2\Gamma\gamma$; and the frequency of gg pairs (the proportion r' of recessive organisms) is γ^2. So we have

$$\begin{aligned} d' &= \Gamma^2 = (d + h/2)^2, \\ h' &= 2\Gamma\gamma = 2(d + h/2)(h/2 + r), \\ r' &= \gamma^2 = (h/2 + r)^2. \end{aligned}$$

(Note that $d' + h' + r' = (\Gamma + \gamma)^2 = 1$.)

These formulas express one of the principal laws of population genetics, the so-called Hardy–Weinberg law, which appears in every modern textbook on the subject. They describe how the composition of a population changes when there is panmixia—that is, when there are no additional factors that may influence the formation of mating couples. (It should be pointed out that the male and female portions of a population are assumed to have equal shares of the dominant, hybrid, and recessive types.)

Now let's look at the geometrical meaning of the transformation f described above—that is, the Hardy–Weinberg law. First of all, we see that the composition of the gene pools of populations P and P' is the same. This clearly follows from the equations

$$\Gamma' = d' + h'/2 = \Gamma^2 + \Gamma\gamma = \Gamma(\Gamma + \gamma) = \Gamma.$$

As we know, this means that points P and P' lie on the perpendicular to DR. Let's assume $DP_0 = x$. In that case,

$$\gamma = r_0 = x \sin 60^\circ = x\frac{\sqrt{3}}{2},$$

so that

$$h' = 2\Gamma\gamma = 2(1-\gamma)\gamma = \sqrt{3}x\left(1 - x\frac{\sqrt{3}}{2}\right).$$

In an ordinary (Cartesian) system of coordinates, this equation gives the arc Π of a parabola when $0 \leq x \leq DR = 2/3^{1/2}$ (fig. 5). The equation for a parabola in our "coordinates relative to triangle DHR" reads $h^2 = 4dr$, since $h'^2 = 4\Gamma^2\gamma^2 = 4d'r'$. It's not hard to see that the parabola passes through points D, R, and T (the midpoint of line segment MN) and is tangential to lines DH, RH, and MN at these points.

Point $P' = f(P)$ is where the perpen-

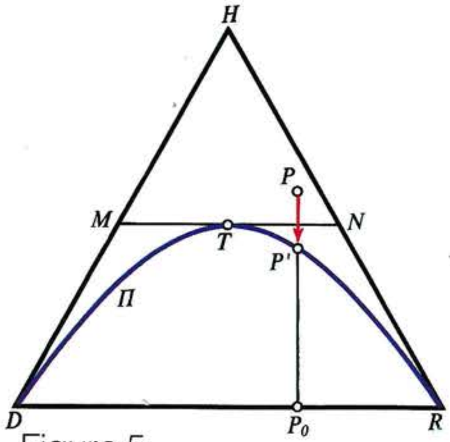


Figure 5

dicular to DR drawn from P intersects the parabola, so that the transformation f is the projection of triangle DHR on arc Π of the parabola in direction h.

From this it follows that all the points of arc Π are fixed—that is, all populations $P(d,h,r)$ for which $h^2 = 4dr$ are stable with respect to transformation f. This result seems paradoxical. It requires that, for a large enough population in the absence of mutations and selection (natural or artificial), evolution always proceeds *in one step*. Any further interbreeding doesn't change the composition of the gene pool: if $P' = f(P)$, $P'' = f(P')$, and so on, then $P' = P'' = \ldots$. But in real life the ideal conditions of interbreeding that we have assumed never hold—we can expect only some crude approximations of them.

So we see that the Hardy–Weinberg law only gives an initial approximation of the real processes that take place in biological populations. Further approximations are given by more sophisticated mathematical models (see, for example, problem 4). Obviously, we need to compare the results obtained by these models with observations of living nature.

Problem

4. Let mating couples in population P be formed randomly (as with panmixia), but assume that the hybrids have less chance of survival than the dominant or recessive organisms. More precisely, only the kth part of the total number of hybrids of the next generation survives. (For $k > 1$ this means, on the other hand, that only the $1/k$th part of D and R organisms survives.) Construct the corresponding transformation of triangle DHR. What's the nature of the transformation for $k = 0$ (the total extinction of hybrids) and k = infinity (the total extinction of D and R organisms)?

The Hardy–Weinberg law was found in 1908, independently, by a young English mathematician, G.H. Hardy, working at Cambridge, and an Austrian physician, Wilhelm Weinberg, in Vienna. Historians of science eventually discovered that the law had been stated five years earlier, in a somewhat different form, by an English biologist, W.E. Castle, but the work had gone unnoticed. Curiously enough, Hardy, an eminent specialist in number theory, was a partisan of pure science. In his largely autobiographical book, *A Mathematician's Apology*, he strongly supports the cause of pure mathematics against applied science, which he considered (perhaps without sufficient reason) boring. Not surprisingly, Hardy's excellent calculus textbook is called *A Course of Pure Mathematics*. Nonetheless, now more than 90% of all references to Hardy in scientific books and articles have to do with his paper on population genetics, which he wrote in his youth, and not his brilliant papers on number theory.

A few remarks on Daltonism

In the simple genetic scheme we've been using, a trait that can exist in two forms is determined by a pair of genes, either of which can be of two types. But the number of genes, as well as the number of their types and laws determining their outward expression, may be quite different. Let's look at a more sophisticated system for the transmission of inherited traits.

Some hereditary diseases are much more common in men than in women. These include color blindness (an inability to distinguish certain colors, usually red and green) and hemophilia (a deficiency in the normal coagulation of blood, which results in prolonged bleeding after even minor injuries). Color blindness was first described by the eminent English chemist John Dalton (1766–1844), and the condition came to be called "Daltonism."

Dalton mistakenly considered color blindness an exclusively male illness. Statistics do, however, tell us that 8% of men suffer from Daltonism while only 0.5% of women do. The American biologist T.G. Morgan explained the difference this way.

The gene for Daltonism, whose dominant variant ensures normal vision while the recessive produces color blindness, is contained in the so-called X-chromosome. This chromosome, together with the Y-chromosome, determines a person's sex (as is the case with other mammals and some insects). The cells in a woman's body contain two X-chromosomes; a man's cells contain one X- and one Y-chromosome. An individual inherits one chromosome from the mother (the X-chromosome, of course) and another, which determines the sex, from the father. So a man receives the unique gene for Daltonism (or any other characteristic related to sex—that is, related to the X-chromosome) from his mother. A woman receives one gene from her father and one gene from her mother.

Now let's consider population P and its series of descendant populations P', P'', ..., generated by random mating (panmixia). We want to study the dynamics of the genetic structure of these populations. In particular, we want to determine the proportion of recessive women and men with the genotypes gg and g—that is, the persons of each sex who suffer from Daltonism.

Three female genotypes—GG, Gg, gg—and two male genotypes—G, g—are possible. We'll denote their proportions in the female and male subsets of population P as d, h, r, δ, and ρ;

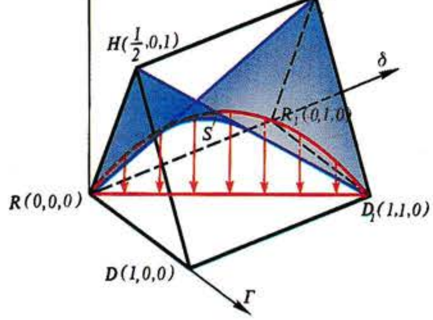


Figure 6

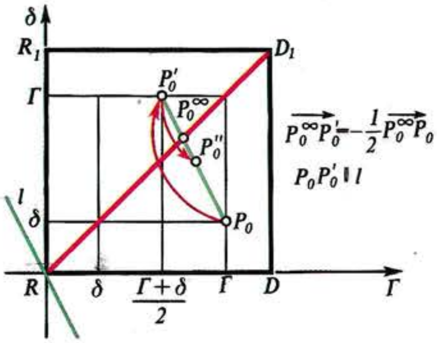


Figure 7

the proportions of genes G and g in the gene pool of the female subset of P are Γ and γ. The corresponding quantities for population P' are given by the equations

$$d' = \Gamma\delta,\ h' = \Gamma\rho + \gamma\delta,\ r' = \gamma\rho;$$
$$\delta' = \Gamma,\ \rho' = \gamma;$$
$$\Gamma' = (\Gamma + \delta)/2,\ \gamma' = (\gamma + \rho)/2.$$

(Can you explain why?)

We'll use a geometric model to study the transformation $P \longrightarrow P'$. Owing to the constraint $d + h + r = \delta + \rho = \Gamma + \gamma = 1$, we can express all the parameters determining the composition of the population by means of three parameters—say, Γ, δ, and h. So to each population we can assign the point given by the Cartesian coordinates (Γ, δ, h). Since the inequalities

$$0 \le \Gamma \le 1,$$
$$0 \le \delta \le 1,$$
$$0 \le h \le 2\Gamma,$$
$$h \le 2\gamma = 2(1 - \Gamma)$$

must be satisfied (the last two because of the constraint $h/2 = \Gamma - d = \gamma - r$), these points lie in the triangular prism $DHRD_1H_1R_1$ bounded by the planes $\delta = 0$, $\delta = 1$, $h = 0$, $h = 2\Gamma$, $h = 2 - 2\Gamma$ (fig. 6). The hereditary transformation changes point $P(\Gamma, \delta, h)$ into point $P'(\Gamma', \delta', h')$ such that point P' and all subsequent points P'', P''', ... belong to surface S, defined by the equation

$$h = 2\Gamma - 4\Gamma\delta + 2\delta.$$

This fact follows from the equation

$$\begin{aligned} h' &= \Gamma(1-\delta) + (1-\Gamma)\delta \\ &= \Gamma + \delta - 2\Gamma\delta \\ &= 2\Gamma' - 4\Gamma'\delta' + 2\delta'^2 \end{aligned}$$

(see figure 6). Consequently, points P', P'', ... are uniquely defined by their projections P_0', P_0'', ... onto plane $h = 0$. We can infer from the equations $\Gamma' = (\Gamma + \delta)/2$, $\delta' = \Gamma$, and figure 7 that segment P_0P_0' is parallel to line l, defined by $\delta = -2\Gamma$, and intersects diagonal RD_1 of face RDD_1R_1 of the prism so that the intersection point divides the segment in the ratio of 2:1. We'll call the transformation that turns P_0 into P_0' a "shear compression" with a scale factor of $-\frac{1}{2}$. It's not hard to see that the points P_0', P_0'', ... obtained from P_0 by repeating the transformation belong to line P_0P_0' and approach point P_0^∞, which is where this line intersects RD_1. Point P_0^∞ and all the points of segment RD_1 are fixed during our transformations. "Lifting" points from plane $h = 0$ onto surface S, we obtain the following statement: the composition of subsequent generations of population P tends to approach (very quickly) that of stable population P^∞; the corresponding point, like all points corresponding to stable populations, belongs to the line of intersection of surface S and plane $\Gamma = \delta$. (Show that this line is a parabola!)

For any stable population $\Gamma = \delta$ we can find all other quantities determining the composition of the gene pool:

$$\rho = \gamma = 1 - \delta,\ d = \delta^2,\ h = 2\delta\rho,\ r = \rho^2.$$

In particular, for $\rho = 0.08$ we get $r \cong 0.006$. These values are roughly the same as the ones given above for the frequency of Daltonism among men and women.

A final word

In this article I've tried to show how geometrical methods illustrate problems in population genetics, enabling us to solve them by elementary methods and obtain important results. I hope you now feel capable of producing similar arguments and solving the problems given below.

Problems

5. Let the scales along the axes h, Γ, δ be equal to $\sqrt{3}/2$, 1, 1, so that the prism $DR \ldots H_1$ is regular. Prove that the distances from point $P(\Gamma, \delta, h)$ to the bases of the prism are equal to δ and ρ, and that the distances to the lateral faces are proportional to the values d, h, r.

6. Prove that the planes $\delta = c$ (a constant) and $\delta - 2\Gamma = c$ intersect surface S along straight lines, whereas the intersections of S and all other planes parallel to axis h are parabolas.

7. Find the coordinates of the limit population P^∞ by means of the coordinates (Γ, δ, h) of the initial population P.

8. Let a population $P(\Gamma, \delta, h)$ be interbred with a fixed population $P_0(\Gamma_0, \delta_0, h_0)$ as follows: female individuals of P are randomly coupled with males from P, while male individuals of P are randomly coupled with females from P_0. We'll denote the descendant population by $P_1 = f_0(P)$.

(a) Find the coordinates Γ_1, δ_1, h_1 and all the rest of the parameters of population P_1 as functions of the coordinates of P and P_0.

(b) Let

$$P_2 = f_0(P_1),\ P_3 = f(P_2),\ \ldots\ P_\infty = \lim_{n \to \infty} P_n.$$

Find the coordinates of the population P_∞ as functions of the coordinates of P and P_0.

(c) Examine specific cases that correspond to various compositions of the gene pool P_0—for example, $P_0 = P_0(1,1,0)$, which is the pure dominant population P_0—as well as the case where population P_0 is one of the many populations that are stable relative to the transformation described in this article.

(d) Describe the maps $f_0: P \longrightarrow P_1$ and $f_\infty: P \longrightarrow P_\infty$ as transformations of the prism. ◘

I.M. Yaglom *is a doctor of physical and mathematical sciences.*

IN YOUR HEAD

Ballpark estimates

How to impress your date and amaze your friends with off-the-cuff answers to questions of magnitude

by David Halliday

SOME PROBLEMS OF PHYSICS involve calculations of the highest possible precision. Many problems, however, call for only an approximate answer. Physicists pride themselves on being able to solve such "order-of-magnitude problems" quickly by breaking them down into their components and making appropriate common-sense estimates.

Here's a typical problem:

> On average, how many atoms of rubber are worn from an automobile tire every time the wheel goes around?

Problems of this kind are often called "Fermi problems" after the great physicist Enrico Fermi, who was a great practitioner of the craft of proposing them and solving them quickly and cleverly.

No doubt you have a few questions.

Yes, I do. Does this problem have any practical significance?

Probably not. Although the problem is an interesting link between the worlds of the very small (the atom) and the very large (the automobile), its real purpose is to help you understand how to make estimates.

But there are no numbers. How can we even start?

We have to estimate the starting numbers—the radius of a tire, the amount of wear ...

But that's just guessing! How can we possibly arrive at an accurate answer?

If by "accurate" you mean an answer good to three significant figures, you're right. But in a problem of this kind, "accurate" means "within a factor of ten either way"—that is, over or under. Actually, it's hard to be that far wrong in estimating the input data.

I get it. Where do we start?

We start with a plan. We'll estimate the volume of rubber worn from the tire and then divide by the volume of an atom. That will give us our answer. Let's deal with the tire first.

Okay. But I don't see any way to guess what volume of rubber is worn from the tire every time the wheel goes around.

We can get an estimate by guessing the volume of rubber worn during the life of the tire and then figuring out how many revolutions the wheel makes during that time. Dividing will give the volume of rubber lost per turn.

Let R be the outer radius of the tire, W the width of the tread, h the depth of wear, and L the distance traveled during the life of the tire. The number of turns N is the total distance traveled divided by the length of the tire's circumference:

$$N = \frac{L}{2\pi R},$$

in which $2\pi R$ is the circumference of the tire. The volume of worn rubber V is the volume of a cylinder of thickness h:

$$V = (2\pi R)Wh.$$

The volume worn per turn is then

$$V_t = \frac{V}{N} = \frac{(2\pi R)Wh}{L/2\pi R} = \frac{(2\pi R)^2 Wh}{L} = \frac{40R^2Wh}{L}.$$

Notice that we've replaced π^2 by 10, which is certainly close enough for our purposes.

But there's no need to replace π^2 by 10. My calculator shows 9.87.

You might feel that you're improving the precision of our answer by doing that, but you're not. Our other estimates will be so approximate that such precision is misplaced. Not only that, 10 is a much simpler number to deal with.

I accept that. What next?

We've already made great progress. We've reduced part of the problem to quantities we can estimate. We'll do that soon. Meanwhile, let's think about atoms.

I've been wondering about that. What is a "rubber atom," anyway? I'm sure you won't find it in the periodic table!

You're right, of course. Rubber is made up of long chain molecules formed from carbon, hydrogen, and oxygen atoms. We're interested here only in a sort of generic atom, whose radius we label r.

I see. Then the volume V_a *of the generic atom would be the volume of a sphere of radius* r, *or* $(4\pi/3)r^3$. *Right?*

You could say that. It's a little better (and simpler) to put the volume at $(2r)^3$—that is, the cube of the diameter. That treats the atoms as little cubes and makes some allowance for the empty space between them.

Now we divide to find our answer. Right?

Right. The number of atoms worn away per turn is

$$n = \frac{V_t}{V_a} = \frac{40R^2Wh}{L(2r)^3} = \frac{5R^2Wh}{Lr^3}.$$

Now we're ready for our estimates. Let's take them one at a time:

R (tire radius) = about 1 ft or 30 cm or 3/10 m,

W (tread width) = about 4 in or 10 cm or 1/10 m,

h (depth of tread wear) = about 1/6 in or 4 mm or 4/1000 m,[1]

L (tire life) = about 50,000 mi or $8 \cdot 10^7$ m,

r (radius of an atom) = about 10^{-10} m.[2]

In putting these numbers into the above expression for n, we must be careful to choose units consistently. Using meters, we find

$$n = \frac{5 \cdot 3 \cdot 3 \cdot 4}{10 \cdot 10 \cdot 10 \cdot 1000 \cdot 8 \cdot 10^7 \cdot 10^{-30}}.$$

Shall I work this out on my calculator for you?

No! It's a point of honor not to use a calculator when solving Fermi problems. Let's rewrite this equation by collecting the integers and the powers of ten:

$$n = \left(\frac{5 \cdot 3 \cdot 3 \cdot 4}{8}\right) \cdot 10^{17}.$$

You can easily see that the number in the parentheses is about 20, so that $n = 2 \cdot 10^{18}$ atoms per turn.

Shouldn't we round that off to 10^{18} atoms per turn?

Yes, indeed. The "2" isn't justified by the precision of our estimates.

So—

When someone asks the "tire question" at a party (and it never fails to come up, believe me!), you can now gaze at the ceiling for a few minutes and say: "About ... 10^{18} atoms per turn, more or less." That's how quickly Fermi himself solved problems like this one!

Try your hand at finding ballpark estimates for these Fermi problems.

1. The population of Boston in 1980 was about 560,000. How many high school teachers were there in that city in that year?

2. How many gallons of gasoline are consumed each year in the United States by private automobiles?

Adapted from the forthcoming book Essentials of Physics *by David Halliday, Robert Resnick, and John Merrill with permission of the publisher, John Wiley & Sons, Inc.* David Halliday *is professor emeritus of physics at the University of Pittsburgh.*

[1] *You* might estimate the depth of tread wear to be 1/2 in (12 mm). If so, your calculations will be slightly different. That's okay—these are estimates.—*Ed.*

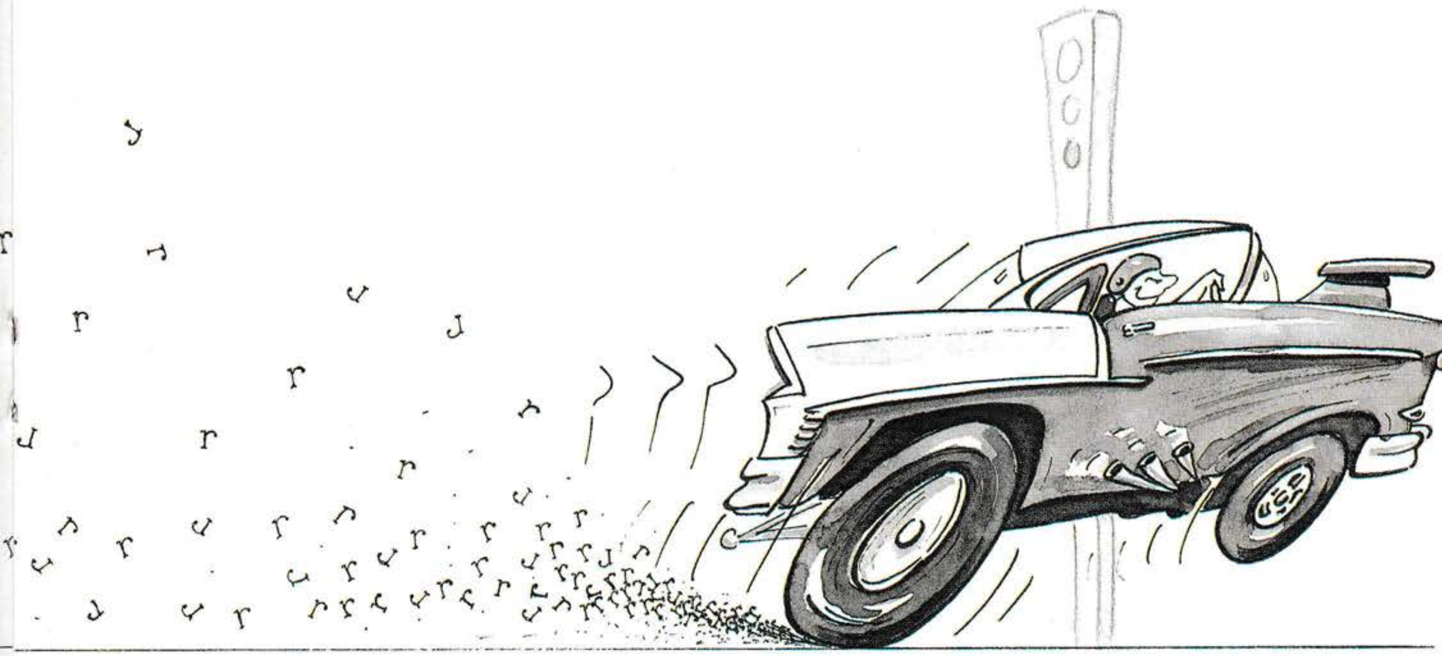
[2] Physicists always use this as an estimate of the radius of an atom. It's a good number to know. (The radius of a nucleus, by the way, is estimated to be 10^{-15} m.)—Ed.

Large or small?

Do you consider the answer to the tire problem (10^{18} atoms/turn) large or small? No answer is possible until you've answered the necessary auxiliary question: Large or small relative to what? As a pure number, 10^{18} seems large. It's 10,000,000 times greater than the number of stars in the Milky Way galaxy, for example.

But the problem deals with 10^{18} *atoms*, not 10^{18} as a pure number. This number of atoms is about 10,000,000 times *greater* than the number of atoms in a typical small bacterium but about 10,000,000 times *smaller* than the number of atoms in a glass of water.

Our conclusion: You can only compare physical quantities of the same kind. There are no absolute standards of "large" or "small."



Art by Nishan Akgulian

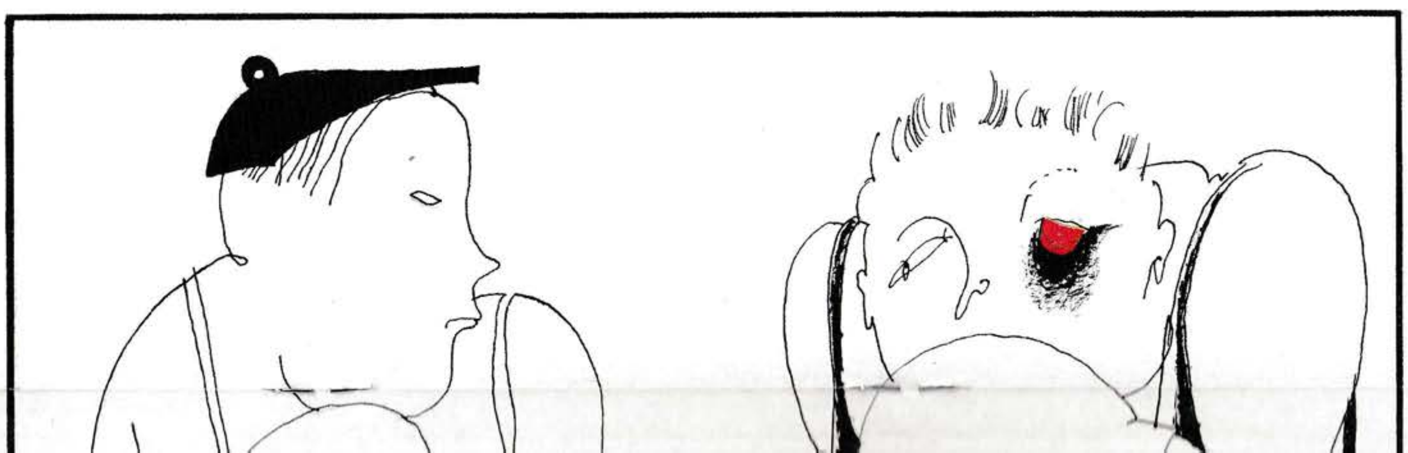
What a commotion!

Are you sure you know how molecules move?

"... heat is an internal motion of matter ..." --M.V. Lomonosov, *A Treatise on the Origin of Heat and Cold*

"The motion of molecules escapes our immediate perception, just as ripples on the surface of the sea do for a distant observer."--Jean Perrin, *Atoms*

THE KINETIC THEORY OF MATTER TAUGHT IN SCHOOL originated in antiquity, when the existence of atoms was conjectured. After a long and arduous pilgrimage, the theory came to rest in its present form at the turn of this century owing to the work of such eminent scientists as Rudolf Clausius, James Clerk Maxwell, Ludwig Boltzmann, and Albert Einstein. The microscopic, or molecular, approach to studying heat phenomena proved to be extremely fruitful, enabling scientists to draw an impressive picture of processes in nature and in machines. This edition of the *Quantum* kaleidoscope contains a few questions that can be answered within the framework of kinetic theory.



Questions and problems

1. How do you explain the phenomenon described by the words "smoke melts in the air"?

2. Why does the layer of cream in a bottle of unhomogenized milk form faster if the bottle is put in the refrigerator?

3. Why isn't Brownian motion observed in the case of tea leaves in a cup of tea?

4. Why does sugar melt faster in hot water than in cold?

5. The velocities of thermal motion of molecules at room temperature are close to the speed of a bullet. Why does the smell of perfume need an appreciable time to fill a room?

6. Why does a ball made from wet wood and coated with varnish keep well, whereas a similar ball not coated with varnish cracks?

7. At an altitude of several hundred kilometers above the Earth, atmospheric molecules possess velocities corresponding to a temperature of a few thousand degrees Celsius. Why don't satellites melt at this altitude?

8. Why does the Moon have no atmosphere?

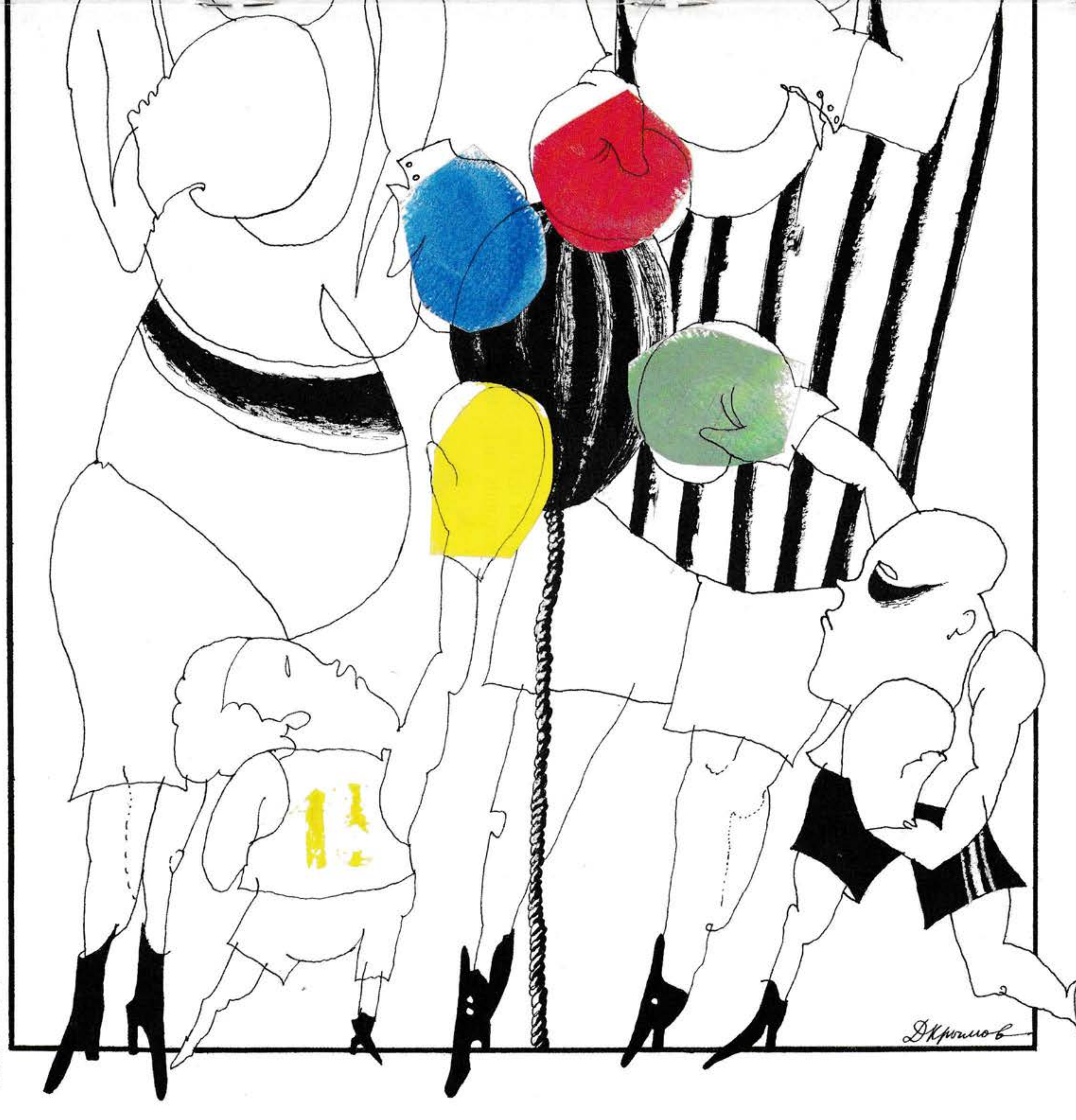
9. Why is it impossible to fix a broken glass by just putting the pieces back where they belong, even though the forces of molecular cohesion are at work?

10. Why is it so difficult to unscrew a nut that has been tightly screwed for quite a while, despite the fact that the nut and the bolt are made of stainless steel?

11. What kind of energy is responsible for capillary phenomena?

12. How would the pressure in a vessel containing a real gas change if attractive forces between its molecules suddenly disappeared?

It's curious that ...



Art by Dmitry Krymov

It's curious that ...

... in 1845 a certain Waterston submitted a paper to the Royal Society (the British academy of sciences). In it he showed that the pressure on the walls of a container can be explained by the collisions of atoms. The idea that elastic properties of gases can be reduced to the classical mechanics of atoms did not get a fair reception, and Waterston's paper was rejected. Many years later, the British physicist Lord Rayleigh found the paper in the archives, and in 1892 he published it in the *Transactions of the Royal Society*. In the meantime, it had taken the work of several persons to rediscover the ignored finding of a lone scientist. In 1859 Maxwell capped their efforts by writing the conclusive formula linking gas pressure and molecular velocity.

... at the end of the last century Boltzmann was so often called upon to refute attacks by opponents of the kinetic theory that he concluded one of his papers by paraphrasing Galileo: "And yet, they move"--"they" being molecules. (Galileo, of course, had bigger things in mind!)

Try this

Pierce three holes in an empty can, making them about 1 mm in diameter and near one another. Pour water into the can and "squeeze" the streams coming out of it by running your fingers over the holes. What happens to the streams? Why?

ANSWERS ON PAGE 62

CONTEST

When days are months

"Then the tides gradually pushed her far away: the tides that the Moon herself causes in the Earth's waters ..."[1]

by Arthur Eisenkraft and Larry D. Kirkpatrick

IT SEEMS PARADOXICAL THAT the Moon's attractive force causes it to move farther from the Earth, but that's exactly what happens. The distance between the center of the Earth and the center of the Moon increases by about 3 to 4 cm per year.

No doubt you're aware that the Moon's gravitational force produces the Earth's tides. As the water moves to form the two bulges that we experience as two high tides each day, there's a lot of friction of the water with the ocean floor and with itself. There are also tides in the "solid" earth. These result in a conversion of some of the Earth's rotational kinetic energy into heat at the rate of about 2 billion horsepower (1.5 trillion joules) each second.

This may seem like a lot of energy, but it's only a tiny, tiny fraction of the total kinetic energy of the Earth. Small as it is, though, this change is measurable. In order to lose kinetic energy the Earth must slow down, but not very much! The length of the day increases by 1.6 milliseconds each century.

But how does this cause the Moon to move farther away? This occurs because the total angular momentum of the Earth–Moon system must remain the same. The loss in the Earth's angular momentum must equal the gain in the Moon's angular momentum. This requires an increase in the size of the Moon's orbit and a corresponding increase in the length of the month—that is, the length of time it takes the Moon to go around the Earth once. So both the length of the day and the length of the month increase. At some time in the distant future, a day on Earth will be a month long!

Although this synchronism of day and month may seem very strange, it already occurs in our solar system. Pluto and its moon Charon have this relationship. They move as if they were two balls on the end of a massless baton—each always has the same face toward the other. Pluto's rotational period is equal to the revolutionary period of Charon—6.4 Earth days.

This leads us to our contest problem: **For the Earth–Moon system, how long will an Earth day be when it's the same as a month?** Let's make a few simplifying assumptions. We'll neglect the revolution of the Earth and Moon around the Sun, the motion of the Earth around the Earth–Moon center of mass, and the rotation of the Moon on its axis. We'll assume that the Moon's orbit is a circle in the plane of the Earth's equator and that the Earth and Moon are uniform spheres.

Please send your solutions to *Quantum*, 1742 Connecticut Avenue NW, Washington, DC 20009. The best solutions will be published in the next issue of *Quantum* and their creators will receive free subscriptions to *Quantum* for one year.

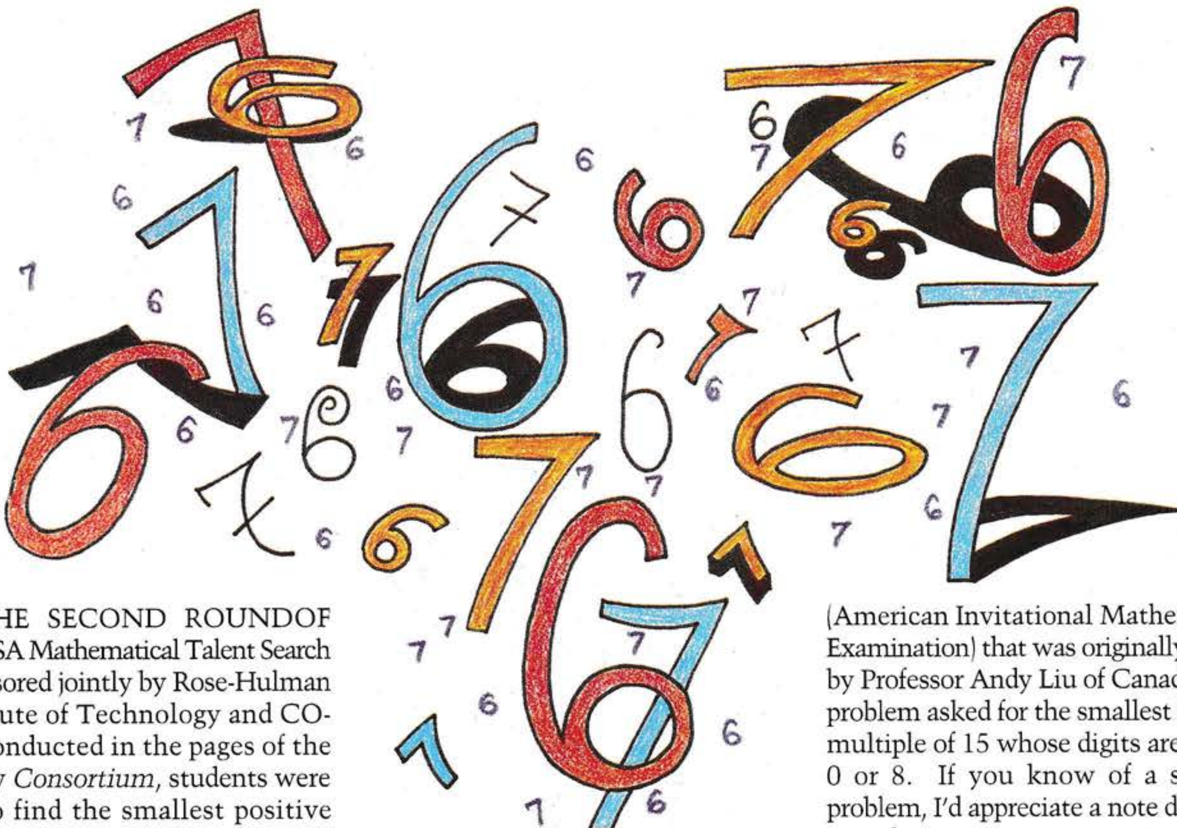
Arthur Eisenkraft *is the chair of the science department and a physics teacher at Fox Lane High School in Bedford, New York.* Larry D. Kirkpatrick *is a professor of physics at Montana State University in Bozeman. Drs. Eisenkraft and Kirkpatrick serve as academic directors for the US Physics Team that competes in the International Physics Olympiad.*

[1]From "The Distance of the Moon" in *Cosmicomics* by the Italian writer Italo Calvino (1923–1985).

At sixes and sevens

—a curious phrase that means "in disorder," but from chaos can come beauty, and order can lie hidden in apparent randomness

by George Berzsenyi



Art by Catherine Lorrain-Hale

IN THE SECOND ROUNDOF the USA Mathematical Talent Search (sponsored jointly by Rose-Hulman Institute of Technology and COMAP) conducted in the pages of the quarterly *Consortium*, students were asked to find the smallest positive integer multiples of 84, 88, and 89 that could be expressed in base 10 using the digits 6 and 7 only. Their response was most encouraging. Nationwide, over 150 students found the answers—76776, 6776, and 667767, respectively—and many provided beautiful analyses leading to their solutions.

Mark Roseberry, a high school senior from Kentucky, also provided a computer printout showing that all positive integers less than or equal to 224 that are not multiples of 5 have an integer multiple consisting of 6's and 7's only.

So here's our challenge to you: **Is the Roseberry Conjecture true for *all* integers that are not multiples of 5?**

If you don't yet feel strong enough mathematically to tackle this problem, maybe you can turn on your computers and—after devising efficient search algorithms—gather further evidence in favor of the claim. Or, alternatively, you may wish to address similar problems using digits other than 6 and 7.

The first problem I mentioned is based on a problem in the 1984 AIME (American Invitational Mathematics Examination) that was originally posed by Professor Andy Liu of Canada. His problem asked for the smallest integer multiple of 15 whose digits are either 0 or 8. If you know of a similar problem, I'd appreciate a note describing where you came across it.

Please send your solutions to *Quantum*, 1742 Connecticut Avenue NW, Washington, DC 20009. The best solutions will be published in the next issue of *Quantum* and their creators will receive free subscriptions to *Quantum* for one year.

George Berzsenyi *is the chairman of the Department of Mathematics at Rose-Hulman Institute of Technology. He was on the committee in charge of the USA Mathematical Olympiad for 11 years, chaired the committee in charge of the AIME for 6 years, and is presently conducting the USA Mathematical Talent Search.*

LOOKING BACK

The secret of the Venerable Cooper

It was an age-old time-saving trick everyone used, but no one understood—that is, until a certain mathematician's curiosity was piqued

by M.B. Balk

JOHANNES KEPLER (1571–1630), the court mathematician of the Austrian emperor Matthew I and a famous astronomer, was observing with curiosity and admiration as a young vintner called out, easily and quickly, the capacities of a succession of wine barrels of different sizes. Kepler recalled the time-consuming measuring procedure used in the vineyards of the Rhine country. The workers there would painstakingly fill each barrel using a container called an "amphora," which served as a standard unit of volume. They'd count the number of amphoras the barrel held and then burn the number into its side with a red-hot iron so that—God forbid!—they wouldn't have to repeat this boring procedure for that particular barrel.

But here in Austria—what a difference! If you've never seen a barrel close up, you might not notice it has a hole, called the "bunghole," on its side right at the middle (that is, its fattest part). The fellow would just shove a brass measuring stick diagonally into the filled barrel's bunghole until it hit the lower edge of the bottom cover (fig. 1). Then he'd simply read off the number on the ruler where it stuck out of the hole. The barrels were all different—large and small, "potbellied" and "skinny"—but this didn't worry the fellow at all. He'd give the answer each time with the same speed and confidence. An unkind thought occurred to Kepler: "He's pulling a fast one!"

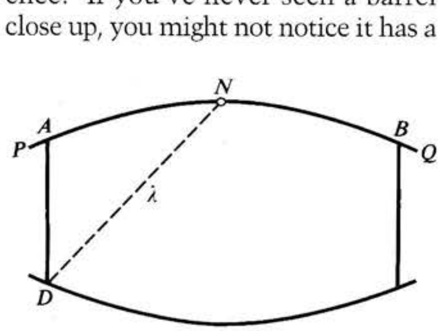


Figure 1

"Don't worry, Honorable Mathematician of His Majesty," the young man said, as if reading Kepler's mind. "This method of measurement is sanctioned here in Linz by the city authorities, and the cooper's guild vouches for the precision of its results".

"For *any* barrel?"

"I don't know about *any* barrel, but for all barrels here in Austria for sure," joked the vintner.

"But how can you be sure this method is correct?"

"What I don't know, I don't know. I won't lie to you. People say that years and years ago the Venerable Cooper lived here, and he proposed this method. But why he proposed it just this way—I'm sure I can't tell you."

Kepler recounts how he deciphered the secret of Austrian cooper in his book *The New Stereometry*[1] *of Wine Barrels, Mostly Austrian, As Having the Most Advantageous Shape, and the Remarkably Convenient Use of the Cubic Ruler With Them, With an Addendum on Archimedean Stereometry*. Although this book was published nearly 375 years ago, it's still quite instructive today, as we shall see.

Kepler first considered the case of a cylindrical barrel (fig.2). Suppose $ND = \lambda$, the length of the generatrix AB is $2x$, and the diameter of the barrel's bottom AD is $2y$. Then

$$x^2 + 4y^2 = \lambda^2, \qquad (1)$$

and the barrel's capacity can be found

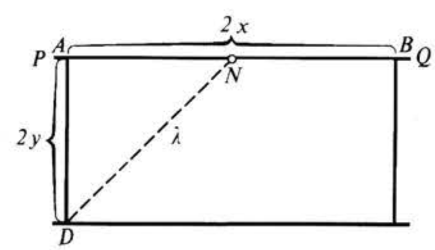


Figure 2

[1]The old name for space geometry—*Ed.*

by applying the formula

$$v = \tfrac{1}{2}\pi(\lambda^2 - x^2)\,x. \qquad (2)$$

It's easy to express this capacity in terms of the distance λ and the ratio t of the generatrix to the diameter of the bottom ($t = x/y$). Equations (1) and (2) imply

$$v = 2\pi\lambda^3 t(4 + t^2)^{-3/2}. \qquad (3)$$

"These formulas show," argues Kepler, "that the capacity of a cylindrical barrel does not depend only on the distance λ. To be able to use the Austrian cooper's method we must deal only with barrels whose ratio t is fixed ($t = t_0$). What is the best choice for this ratio? What is the most profitable choice for the ratio of the generatrix's length to the diameter of the barrel's bottom? Let's argue from the vintner's point of view. His interest lies in the choice, among all values of t, of the value $t = t_0$ such that, of all barrels with distance ND equal to λ, the one chosen has the maximum volume v_0. Then he can calculate this volume according to equation (3). And if the craftsman who made the barrel did not succeed in achieving the ideal ratio too exactly but was somewhat off the mark (not $t \approx t_0$, that is, but $t > t_0$ or $t < t_0$), the vintner will not suffer because of the miscalculation: the actual capacity v of the barrel will necessarily be *less* than v_0. And so the vintner will declare the capacity v_0 and be paid accordingly by the client, but in fact will give the client slightly less wine, namely v.

"But then, no self-respecting vintner wants to swindle his clients. He wants the actual capacity of the barrel (corresponding to the actual choice of parameter t) to differ as little as possible from the volume v_0 corresponding to the best value of the parameter (namely, t_0). The buyer's interest in this is even stronger. It turns out that this state of affairs is also achieved by choosing the parameter t that yields the greatest value for the barrel's volume v."

Art by Sergey Ivanov

Since we are familiar with the notion of derivatives, we can easily justify Kepler's reasoning. (It was much harder for Kepler—his book appeared some 70 years before the differential calculus was born!) Let's make use of the following general principle. Assume $f(t)$ is a function of t (for simplicity's sake, one that has a first derivative for all $t > 0$); and suppose that t_0 is some fixed value of the independent variable t. Then, for a small increment h—that is, $t = t_0 + h$, we have

$$f(t) - f(t_0) = f'(t_0)\,h + \alpha(h)\cdot h,$$

where $\alpha(h) \to 0$ as $h \to 0$ (see any calculus textbook).

So if $f'(t_0) \neq 0$, the deviation of $f(t)$ from $f(t_0)$ will be, for small values of h, practically proportional to h. But if t_0 is a maximum (or minimum) point of the given function, then $f'(t_0) = 0$ and the deviation equals $\alpha(h)\cdot h$, where $\alpha(h) \to 0$ as $h \to 0$. This means that near the extreme value t_0, a small deviation of parameter t from t_0 affects the variation of the function $f(t)$ much less than it does near some value of t where $f'(t) \neq 0$. (Kepler calls the variation "insignificant.") In our case—see equation (3)—it's most advantageous to choose the parameter t in such a way ($t = t_0$) that the cylindrical barrel has the greatest capacity (v_0). In this case, when there are small deviations of t from t_0 (which are inevitable in practice), the deviation of the barrel's volume v from the maximum volume v_0 will be insignificant.

Kepler then computes (for a given λ) the value of the ratio t that yields the maximum capacity for the barrel. Using the notion of the derivative, we can do this much faster and more easily than Kepler could. Let's use equation (2) for v. Then

$$v'(x) = \tfrac{1}{2}\pi(\lambda^2 - 3x^2) = 0,$$

which works out to

$$x = \frac{\lambda}{\sqrt{3}}.$$

Using equation (1) to get y, we get

$$4y^2 = \lambda^2 - x^2 = \frac{2\lambda^2}{3},$$

$$y = \frac{\lambda}{\sqrt{6}}.$$

So

$$t_0 = \frac{x}{y} = \sqrt{2} \cong 1.41\,. \qquad (4)$$

For this value of parameter t the capacity of the cylindrical barrel will be greatest (for a given value of λ) and will equal

$$v_0 = \frac{\pi}{3\sqrt{3}}\lambda^3. \qquad (5)$$

"Now I will consider the more general case," Kepler writes, "in which the staves (that is, the boards that make up the sides of the barrel) are bent, not straight as they are for a cylindrical barrel. Such a barrel can be represented with a sufficient degree of precision as two identical truncated cones joined together along their larger circular bases." (See figure 3.)

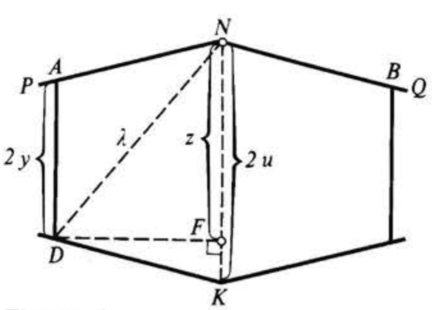


Figure 3

"Of course," Kepler continues, "I am accepting a certain amount of imprecision, but if the barrel is not too "potbellied," this error will be insignificant. Then among all the barrels for which the distance ND equals the given λ, I will choose the one with the greatest capacity."

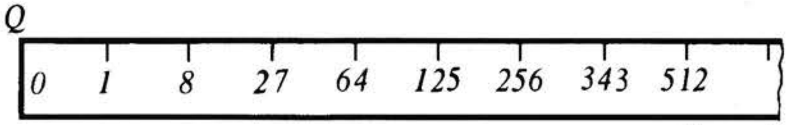


Figure 5

Following the procedure outlined by Kepler, we can express the capacity v of the barrel in terms of the distance λ, the radius y of the smaller base of the truncated cone, and the distance $z = NF$ (where F is the foot of the perpendicular dropped from point D to the larger base of the cone). Let's perform these calculations, using the formula for the volume of a truncated cone and bearing in mind that the altitude of the truncated cones here equals

$$DF = \sqrt{\lambda^2 - z^2}.$$

Simple manipulations then yield

$$v = \frac{2}{3}\pi \cdot [y^2 + (z-y)^2 + y(z-y)] \cdot \sqrt{\lambda^2 - z^2}. \qquad (6)$$

Kepler was able to find the correlations among the various barrel dimensions (that is, the values of y and z for a given λ) that produce maximum barrel capacity. We can obtain Kepler's results by solving the following problem: using derivatives, prove that among all the barrels of the type described (with a given value λ for the distance ND), the one that has the greatest volume is the cylindrical barrel whose generatrix is $2^{1/2}$ times longer than the bottom's diameter.

Now let's use Kepler's idea that, near an extreme value, a small deviation of the independent variable leads to an insignificant change in the function's value. In our case, this means that if the dimensions of some barrel don't deviate much from those of the best barrel (that is, a cylindrical barrel with the ratio $AB : AD = 2^{1/2}$), these deviations will have an insignificant effect on the barrel's volume. So the volume of a barrel differing only slightly from the "best" one can be computed according to the same formula—equation (5).

"It is now clear," writes Kepler, "what an Austrian barrel really is: it is one whose generatrix ANB is about $2^{1/2}$ times longer than the bottom's diameter. Since in choosing the stave length PNQ we have to take into account the thickness of the top and bottom of the barrel and the fact that the staves stick out a little beyond the top and bottom, the staves must be about one and a half times longer than the bottom's diameter."

Not only that, Kepler also understood how to find the capacity of any "Austrian barrel" by means of an ordinary ruler: just measure the distance ND and apply equation (5) to find the volume v_0. "But the vintner called out the capacities immediately, without counting," Kepler recalled. "Apparently his ruler is not an ordinary one, but has marks that are positioned differently. How is this ruler calibrated?"

Kepler turned his attention to a secret amphora the vintner mentioned. He guessed that it was probably some kind of little barrel as well, whose capacity is taken as a unit of measurement. But if Austrian coopers produce barrels with an optimal ratio of generatrix to diameter (that is, $t \cong 1.5$), surely they must use the same ratio in making the standard barrel-amphora. Assume that for the standard barrel-amphora the distance N_1D_1 equals λ_1 (fig. 4). Then, taking the capacity of the barrel-amphora as the unit volume, equation (5) gives us

$$1 = \frac{\pi}{3\sqrt{3}}\lambda_1^3.$$

Using equation (5) again, we get the following adequate approximation of

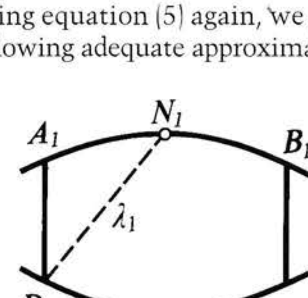


Figure 4

the volume of an "Austrian barrel":

$$v_0 = (\lambda/\lambda_1)^3 \text{ barrel-amphoras.}$$

So, in order to find v_0, it's enough to know the ratio λ/λ_1 (rather than the actual value of λ). In fact, you don't even have to know the ratio itself, just its cube. The values of $(\lambda/\lambda_1)^3$ are in fact the markings on the ruler—that is, we have to calibrate it according to the "law of cubes." At distances of λ_1, $2\lambda_1$, $3\lambda_1$, ... from the end of the brass ruler, the numbers 1, 8 (= 2^3), 27 (= 3^3), ... are written (fig. 5). To state the general case: at a distance $k\lambda_1$ (where k is any positive integer) from the end, the number k^3 is written. If you stick this ruler into a barrel so that its tip touches the lower edge of the barrel's bottom (point D), the reading on the ruler at the bunghole (point N) will indicate how many barrel-amphoras that barrel holds.

IN THE AUTUMN OF 1615, at Kepler's request, the Elder of the Cooper's Guild met with the Emperor's Mathematician at the Linz City Hall.

"Honorable Elder," began Kepler, "I am interested in the method Austrian coopers and vintners use to measure the capacity of barrels."

"Well, you see, Honorable Mathematician of His Majesty," objected the elder, "this is a trade secret of our guild. This secret has been handed down from generation to generation by our craftsmen since the days of the Venerable Cooper."

"I had surmised that at one time there existed an outstanding geometer who taught this method to your guild. I believe I have managed to guess his secret".

"Then tell me, Honorable Mathematician of His Majesty, what the secret of the Venerable Cooper is, in your opinion, and I promise to confirm whatever parts of your explanation are correct."

"My reflections have led me to the conclusion," answered Kepler, "that when the coopers of Linz construct a barrel, they are guided by only one consideration: that the staves be one and a half times longer than the diameter of the bottom."

"Absolutely correct!" affirmed the startled cooper.

"Further, to measure the capacity of the barrels you use a ruler calibrated according to the law of cubes."

"You are right again," said the elder of the cooper's guild with even greater surprise, looking apprehensively at Kepler. "Are you endowed with supernatural powers, Honorable Kepler? I have heard that, as the Emperor's Astrologer, you can read the past and the future from the disposition of the stars. Could it be that the Venerable Cooper's secret was deciphered by means of astrology?"

"No," answered Kepler dryly, "I computed it by means of mathematics." And there was nothing more to be said between the mathematician and the cooper.

... LOOKING BACK

From the prehistory of radio

This article originally appeared in 1984 to commemorate the 125th anniversary of the birth of the great Russian physicist Alexander Popov

by S.M. Rytov

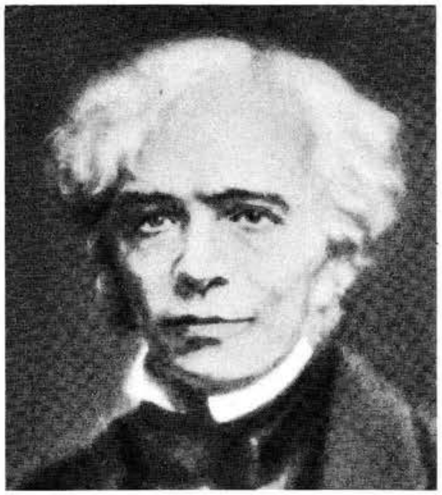
RADIO IS ONE OF THE DISTINCtive features of modern civilization. It's a means of communication and an instrument for scientific research, both on Earth and in outer space. It exerts a tremendous influence on our daily life.

Not that much time has elapsed since the first experiments in radio communication, so we might conclude that the entire history of radio is confined to the 20th century. This conclusion would be wrong. In looking at the history of radio, we have to remember that this technology stemmed from discoveries in the theory of electricity. And the scientists who made the most important contributions in this field are Faraday, Maxwell, and Hertz.

Faraday and his experiments

The most important discovery was made by Michael Faraday in 1831. Faraday found that electricity can be generated by changes in a magnetic field. This phenomenon is called "electromagnetic induction."

Faraday's tremendous experimental work was guided by the concept of interaction through a medium, which he visualized as lines of force traveling through space. This stood in direct contrast to the concept of immediate action at a distance, supported at the time by the authority of Newtonian



Michael Faraday

physics and such eminent physicists as Ampère, Weber, Kirchoff, and Thomson (Lord Kelvin). Faraday's ideas encountered considerable opposition, but he forged ahead and made one important discovery after another.

Faraday used little mathematics—his method of research was almost purely experimental. He derived his ideas under the direct influence of experiments, turned to experimentation as a way of testing the truth of his ideas, and expressed his results in nontechnical language.

Maxwell and his mathematics

James Clerk Maxwell understood the profundity of Faraday's ideas and decided to create a mathematical framework for them. This paved the way for the systematic study of electromagnetism. In his famous book *A Treatise on Electricity and Magnetism*, he wrote: "I have therefore thought that a treatise would be useful, which should have for its principal object to take up the whole subject in a mathematical manner, and which should also indicate how each part of the subject is brought within the reach of methods of verification by actual measurement." Proceeding along these lines, he did quite a bit more. Maxwell introduced the concept of the electromagnetic field, finally rejected the concept of action at a distance, found that a change in an electric field in time generates a magnetic field, and derived the equations that bear his name. To do all this Maxwell had to introduce a new physical quantity (similar to electric current, caused by electrical charges moving in space) that he called "electric displacement." The discovery of electric displacement was the direct result of his efforts to create a mathematical structure for electromagnetic theory.

The beauty of Maxwell's equations, which describe a multitude of physical phenomena, appeals to every physicist or mathematician, and their mathematical structure provides guidelines for research in theoretical physics even today.

An important conclusion that followed from Maxwell's theory was the existence of transverse electromagnetic waves, which have a finite propagation velocity given by the equation

$$c = \frac{1}{\sqrt{E_0 \mu_0}}.$$

The constants of dielectric susceptibility (E_0) and magnetic susceptibility (μ_0) had been measured previously. The value of c turned out to be equal to 193,088 miles per second. This was very close to the speed of light obtained in optical experiments by Hippolyte Fizeau—193,118 miles per second. This led Maxwell to conjecture that light consists of electromagnetic vibrations. Maxwell first mentioned this brilliant discovery in a letter to Faraday dated November 19, 1861. Neither of the values given above are quite accurate. Maxwell no doubt understood that the speed of light had been measured inaccurately, so he played down the numerical disagreement.

James Clerk Maxwell

Maxwell himself felt that the electromagnetic theory of light had been created by Faraday. He wrote something to that effect in 1864, referring to ideas that Faraday had expressed in one of his papers in 1846. But perhaps a distinction should be made between a general idea and a precise statement that can be verified by experiment. Maxwell's theory not only supplied a general framework for previously known electromagnetic phenomena, it also provided a means of predicting new phenomena and describing them quantitatively. The most important phenomenon it predicted was the existence of electromagnetic waves. But neither Faraday nor Maxwell (at least at the early stages of his work) had noticed the direct connection between optics and electromagnetism, which became apparent only after Maxwell had worked out his theories explaining phenomena in electricity and magnetism.

Hertz and his "living equations"

Electromagnetic waves were first detected in the laboratory by the German physicist Heinrich Hertz in 1888. Hertz's experiments have long been considered remarkable examples of experimental ingenuity and theoretical foresight. The Russian physicist O.D. Khvolson called them "eternal classics." Hertz also made important contributions to the mathematical framework of Maxwell's theory. He admired its beauty and wrote that its equations have a life of their own—that they are wiser than we are, even wiser than their creator, and give more information than was put into them. These words proved prophetic. The second half of the 20th century has seen the flourishing of so-called "gauge theories," which from the mathematical point of view are generalizations of Maxwell's equations to suit the needs of subatomic physics.

Immediately after Hertz's experiments, many scientists proposed using electromagnetic waves for communication without wires. It's interesting to see what Hertz himself thought about it.

In 1889 an engineer from Munich named Guber put the question to Hertz

Heinrich Hertz

directly. In 1897, three years after Hertz died, Guber sent Hertz's reply to a German research journal, where it was published without comment. In his letter Hertz wrote that the lines of force of magnetic and electric fields travel like light rays only when their vibrations are very rapid. Then both kinds of line forces are inseparable from each other, so that the rays of waves studied in Hertz's papers can be called "magnetic" or "electric" interchangeably. But vibrations in a telephone wire are much slower. So radio waves of very low frequency are required because (Hertz thought) the frequency of the radio waves should be the same as that of the transmitted sound; consequently, the wavelength would have to be very large. Hertz suggested that radio waves could be focused by concave mirrors to obtain a signal strong enough to detect. He saw no other means to achieve that end. But for very long waves, these mirrors would be the size of a continent!

I'll make no comment about this letter either. Let's just say that even a genius has a hard time trying to guess the solution to a technical problem decades in advance—in this case, wireless telegraphy.

Of course, mere suggestions that electromagnetic waves be used for communication were insufficient. A way had to be found to *do* it. Enter, stage right, Alexander Popov and his apparatus for detecting electric vibrations.

Popov and his "thunder detector"

By the time Popov began to work on the transmission of messages by means of radio waves, the physical phenomena involved and the main parts of his apparatus were known. It was well known, for example, that a spark discharged by a Leyden jar, an electrostatic machine, or lightning consists of damping electric oscillations. Hertz had shown that such oscillations generate electromagnetic waves. So the means of generating electromagnetic waves were already available. The power of the spark had to be increased, but this wasn't considered a serious obstacle—it was a mere technicality, so to speak.

The main problem was how to detect signals at a distance. For this purpose Hertz had used a small spark generated in the tiny gap of a ring. But this method, of course, couldn't serve any practical purpose.

Five years before Popov invented his device, an apparatus called a "coherer" had appeared. It consisted of a glass tube with two electrodes at each end and metal filings between them. When an electromagnetic wave travels through the tube, tiny sparks run between the filings at their points of contact so that the resistance of the filings drops from hundreds of thousands of ohms down to less than a hundred ohms (or even less than ten). The coherer is an excellent device for allowing a high-frequency signal to short a circuit with an electric battery in it. Unfortunately, as soon as the coherer has done its job, it quits. You have to give it a shake to restore its high resistance. So this device was obviously not a receiver constantly ready to detect a signal.

Popov transformed the coherer into a reliable radio receiver by inventing a method of restoring its large resistance automatically. In an article that appeared in January 1896, he wrote: "Having achieved good results in making the sensitivity constant by means of a tube with platinum electrodes and iron filings,[1] I aimed at constructing a device that would automatically destroy contacts between filings caused by electric vibrations." His invention was a great step forward in radio technology. In constructing his apparatus Popov not only achieved electromechanical amplification but also used what we now call feedback.

The new device could sense the distant electrical discharges of lightning, so Popov called it a "thunder detector," even though it also detected discharges from an electrostatic machine and a large Hertz vibrator as well as the rapid strokes of an electrically charged ebonite rod.

As we come to the end of this article, I'd like to quote from Popov's report to the Physical Section of the Russian Physico-Chemical Society on May 7, 1895: "Concluding, I would like to express the hope that my apparatus, after further development, can be used to transmit messages through space by means of rapid electrical vibrations, as soon as a sufficiently powerful source of these vibrations is found."

Actually, though, progress was made during the early years of radio communication by increasing the height of the receiving antennas. But advances on all fronts continued at such a pace that as early as 1897 P.N. Ribkin, Popov's assistant, managed to transmit radio signals to a distance of five kilometers.

The 20th century has brought many new inventions and improvements in

Alexander Stepanovich Popov

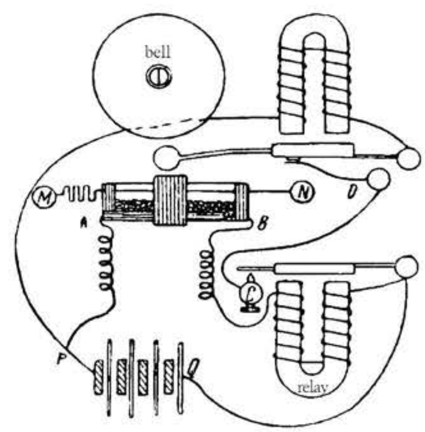
[1] This refers to Popov's earlier work on improving the parameters of the coherer itself.

radio technology—for example, television, which plays an increasingly important role in our lives. In 1931, the centenary of Maxwell's birth, G.G. Thomson, himself an eminent experimenter, said that the discovery of radio waves has influenced the entire development of civilization. Radio waves help draw the Earth's inhabitants together and have led to social, educational, and political changes that we are still in the process of assimilating.

Outline and description of Popov's device

Here's an excerpt from Popov's 1896 paper "An Apparatus for Detecting Electric Vibrations":

"The figure illustrates the positions of the constituent parts of the apparatus. The tube with filings is suspended horizontally between the two clamps M and N by a light watch spring, which is bent in a zigzag so as to be more elastic. There is an electric



bell above the tube; when it rings, it hits the middle of the tube. (A rubber ring keeps the tube from being broken.) It is convenient to attach the tube and bell on a common vertical plank. The relay can be placed elsewhere.

"The apparatus functions as follows. The current generated by a 4- or 5-volt battery flows from the contact P to the platinum plate A, then across the filings contained in the tube to the other plate B, and finally through the coil of the electromagnet back to the battery. The power of the current is not sufficient to draw the anchor of the relay; but if the tube is subjected to the action of electric vibration, the resistance instantly decreases and the current increases enough for the relay anchor to be attracted. At this moment the circuit running from the battery to the bell that is interrupted at point C closes, and the bell begins to function. It shakes the tube, decreasing its conductivity, and the relay breaks the bell's circuit. In my apparatus the resistance of the filings after a strong shaking may reach 100,000 ohms, but the relay (which has a resistance of about 250 ohms) draws the anchor at currents of 5 to 10 milliamperes (the limits of adjustment)—that is, when the resistance of the whole circuit drops below 1,000 ohms." ◘

CONTINUED FROM PAGE 7

They're even located inside some bones, sometimes down to the small bones of the feet. As a result, the respiratory system of a duck, for example, takes up about 20% of its body volume (2% for the lungs and 18% for the air bags), while a human being's respiratory system takes up only 5%. A bird's lung, as opposed to a mammal's, is a set of thin pipes connected in parallel, open from both sides and surrounded by blood vessels. The bags are connected to the lung in such a way that air flows through a bird's lung in the same direction during both inhalation and exhalation. So in the breathing process only the volumes of the bags change while the volume of the lung remains practically constant. Since there's no need to inflate the lungs, surface-active substances aren't needed in a bird's lungs.

Maybe you thought it was a silly question, "How do we breathe?" I wonder now what you think of the answer: "Not as well as birds!" ◘

K.Y. Bogdanov *is a doctoral candidate in physical and mathematical sciences.*

QUANTUM SMILES

Disorder in the court!

When "just the facts, ma'am" just won't do

by V. Fabricant

IN A PROVINCIAL TOWN OF CENtral Russia at the turn of the century, the owners of an electric power plant caught a certain Ivanov in the act of using electricity "free of charge," so to speak, in his home. He had surreptitiously connected a wire to a power line belonging to their enterprise. The company sued Ivanov for theft of its property, but the malefactor won the case in court. According to the existing criminal code, theft was defined as pertaining to an object possessing mass and size. The Senate, however, which considered the case on appeal, determined that electrical energy may be viewed as property, "being mobile real estate, without being a material entity." So the power company triumphed in the end.

A more sophisticated legal issue arose in another case. Beer produced by a certain brewery was stored in a cellar some distance away from the bottling plant. The cellar was cooled by pipes containing a circulating saline solution that came from a central cooling unit. The main pipe connecting the cooling unit and the cellar happened to pass near the cellar of a retailer.

After a while, the owner of the brewery discovered that the retailer was using the saline solution to cool his own cellar and sued the retailer for theft. The judge, however, ruled: "In accordance with Article 242 of the Criminal Code, theft is the unlawful appropriation of commodities belonging to another party. In the present case no theft has been committed, since the saline solution was not misappropriated; rather, it was returned in its entirety to the brewery's main pipe."

The brewery owner appealed this judgment, now arguing his case this way: "The issue is not the theft of saline solution but the theft of energy. If the saline solution is used to cool the defendant's cellar in addition to my own, I have to pay more for electricity to operate the central cooling unit." The ruling of the Court of Appeals was, from the point of view of physics, only half convincing: "The saline solution acquires heat from the retailer's cellar; therefore, energy belonging to the brewery is not being stolen. On the contrary, the brewery is receiving gratuitous energy from the retailer."

We all agree the judge was wrong, but not everyone can correctly explain his error. Can you?

The first story can be found in *Electricity and Heat* by B. Bulyabash and V. Gurevich, Moscow: Nauka Publishers, 1978. The second story is taken from a German scientific monograph, "Questions of Thermodynamical Analysis," by P. Grassman.—*Ed.*

A horse is a horse (of course, of course)

Even cowpokes get the blues when they place a few bets and lose, and lose ...

by A.S. Yarsky

BILL CAME TO MR. COHN'S place right after lunch. At that time of day the Texas sun turns all the cheeses of Texas into gooey cream cheese, all the water of Texas into boiling water, while all the inhabitants of Texas can only dream of a quicker and less painful end.

When his temperature fell below the boiling point of iron and all the ice in the house had evaporated, Bill finally recovered the power of speech.

"You understand, mister Cohn, it's all because of that Stanley! He's a friend of mine, so to speak. Until the other day I only bet against him once't."

"And how did that bet go?" Mr. Cohn was struggling hard to stay awake.

"Well, mister Cohn, Stanley was goin' off on this trip—Natchez, I think it was, in the upper straits of Red River. Me, I been there, so I says to him:

" 'That's seven days on horseback.'

" 'Aha,' says Stanley, 'and to Baton Rouge?'

" 'A day,' I tell him.

" 'And to Lafayette?'

" 'About four days. It's a tough road.'

" 'And how long to ——?' And he names this hole in the wall.

" 'Never been there,' I says. 'I guess it be at least ten or twelve days.'

" 'You know, Bill,' says Stanley, 'these four towns are situated smack dab at the four vertices of a rectangle.'

" 'So what?' I says.

" 'So this. I can tell ya exactly how long it'll take ya to get to ——.' And he names that hole in the wall agin.

" 'You mean you been there?' I ask him.

" 'Nope,' he says, 'ain't never been.'

" 'Then how the heck d'ya know? Somebody tell ya?'

" 'No siree. I just know it's eight days!'

"Well, mister Cohn, the argument goes on and on and ends up in a bet. I get on my palomino and—can you beat this?—I make it to that hole in the wall in *exactly eight days*! Wish I'd never seen it. Eight days later I'm back here, and Stanley tells me how there's this 'theorem' thing.[1] Turns out I dragged my tail to the Red River to prove some word I don't even know what it is!"

"Your friend Stanley's a clever fella." The heat was falling off, and Mr. Cohn was recovering a bit.

"Too clever by a half," Bill snorted. "And so, mister Cohn, the other day this Stanley comes by to see me, we shoot the breeze about the weather, beef prices, and all that, and then he up an' says:

" 'Listen, Bill, what d'ya think, can I find a horse in these here parts that you can't handle but I'll be able to ride?'

"And I answer without battin' an eye:

" 'Well, Stanley, mark my words: I'll use that pony to take you to the nearest graveyard.'

"But he keeps right on:

" 'I figure there's a horse like that, Bill.'

" 'OK,' I says, 'I bet ya 10 to 1 there ain't. Dig out a buck and let's have a look!'

" 'Take your time, Bill, and get your ten smacks ready. The horse is right here in your stable. Look!'

"Then he picks up this notebook I use to do my accounting, opens it to an empty page, and writes:

$$\frac{a}{b}-\frac{c}{d}=\frac{a-c}{b-d}. \qquad (10)$$

"Well, mister Cohn, first thing, I ask him:

" 'Stanley, why'd ya start your count from ten, there on the right, in parenzathes?'

" 'That's how much you owe me. And that's the place to write the bill, Bill—right in your account book!'

" 'Not so fast,' I tell him. 'S'been a long time since I went to school, but I can tell you this: your number ten horse is lame on all four legs. And you ain't getting that ten spot 'til it takes at least one step.'

" 'Then take a gander at this,' says Stanley, and he writes:

$$\frac{8}{2}-\frac{9}{3}=\frac{8-9}{2-3}.$$

" 'Check for yourself,' he says, 'and fork over the dough!'

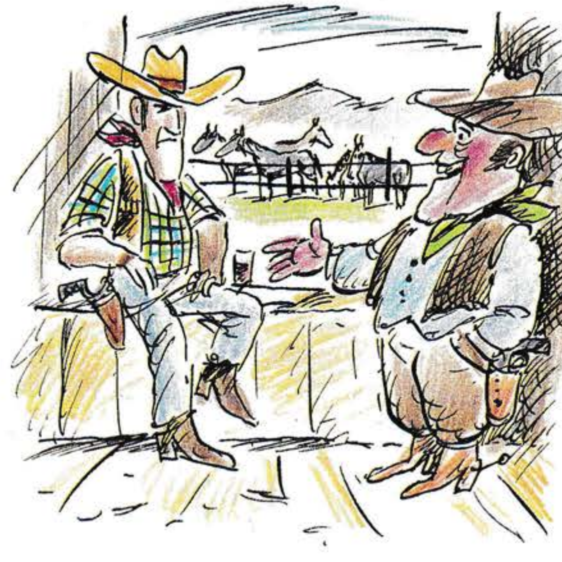
"Now tell me, mister Cohn, have you ever seen a guy make a ass of himself and then admit it? Or a guy who had no use for ten bucks?

" 'Hey, Stanley,' I says, 'I'll bet you another ten that your gelding can't take another step!'

" 'Then fork over another ten, 'cause here it goes.' And he writes:

$$\frac{5}{3}-\frac{9}{9}=\frac{5-9}{3-9}=\frac{-4}{-6}=\frac{2}{3}.$$

"Lose two races in row—that's never happened to me before, mister Cohn.



Art by Edward Nazarov

But then I remember the terms of the bet, and I figure everything ain't lost yet.

" 'Listen here, Stanley,' I tell him, 'that horse rides under you, no argument there. But you said I couldn't handle it. My twenty to your twenty says I'll ride that mustang easy as pie!'

" 'All right, my twenty says you won't.'

"Well, mister Cohn, I lost that race too. Mister Cohn, you don't make as much money in a week teaching school as I lost to that son of a gun Stanley in half an hour ... And, dang it, I still can't figure out how the heck that lame mustang—I mean that incorrect equation—gives correct answers! In fact, that's what I came here to ask you about."

"Science requires sacrifices," declared mister Cohn, "and you, Billy boy, were sacrificed at the altar of science! First of all, whoever told you an incorrect equation can't give correct results? Ba-*lo*-ney! Even the biggest liar of them all can accidentally tell the truth. And besides, while you were telling me about your lost 'races,' one after the other, I discovered the underlying logic of your disgrace. Take a peek at this formula:

$$\frac{a}{c}=\frac{b}{d}\left(2-\frac{b}{d}\right).$$

Never mind where I got it. What's important is that this is just another way of writing that same equation (10). It looks a little different, but it's really the same equation. But now,

---

[1] "If $ABCD$ is a rectangle and S is an arbitrary point, then $SA^2+SC^2=SB^2+SD^2$." Try and prove it.—*Ed.*

Billy, you can take any b and d and use them to compute a and c! That's it. Simple and foolproof! Want an example? Easy enough. Name any two numbers. Two and seven? Fine. Let's take $b = 2$, $d = 7$, and work it out:

$$\frac{a}{c} = \frac{2}{7}\left(2 - \frac{2}{7}\right) = \frac{2}{7} \cdot \frac{12}{7} = \frac{24}{49}.$$

So we can take $a = 24$, $c = 49$, or $a = 48$, $c = 98$ ... Now let's check it with equation (10):

$$\frac{24}{2} - \frac{49}{7} = \frac{24-49}{2-7};$$
$$\frac{48}{2} - \frac{98}{7} = \frac{49-98}{2-7}.$$

Incidentally, you can double, or triple, the denominators, or numerators:

$$\frac{24}{6} - \frac{49}{21} = \frac{24-49}{6-21}.$$

Multiply them by *ten* if you want to!"

Satisfied with his explanation, Mr. Cohn took a deep breath. Bill looked over the computations carefully and summed it all up.

"Twenty at most. Won't win back more than that from Stanley using this trick."

"Now hold on a minute," said Mr. Cohn impatiently, "this Stanley of yours may already know the rule I discovered. But here's something he most surely won't be able to do, simply because it's impossible! Bill, just ask him the following question: *Can his number 10 horse rear its head?!* In other words, ask him to satisfy relation (10) for $a > c$! But don't forget, Bill, that all these letters—a, b, c, and d—must be positive integers. You got that?"

IS RELATION (10) POSSIBLE WHEN $a > c$?

Bill's sullen face lightened up a bit.

"If everything is like you say, mister Cohn, I'll charge him five dollars for every hour he sits around wracking his brains. Five bucks an hour—that's fair enough, ain't it?"

"Don't be so greedy, Bill," Mr. Cohn objected. "This is something that'll set you up for life. Even fifty cents an hour will net you the tidy sum of twelve dollars a day ...!"

Mister Cohn fell silent for a moment.

"You know, Bill, you should assign this *impossible* task to someone who's richer than Stanley. Otherwise the idea won't last for more than a couple days."

"Let me worry about that," replied Bill. "But still, I'd like to know *why* that horse can't rear its head. And one other thing, mister Cohn: what if he knows this one, too?"

"If he does"—mister Cohn was getting hot—"let him find *irreducible* fractions a/b and c/d that satisfy equation (10)!"

IS RELATION (10) POSSIBLE WHEN THE FRACTIONS $\frac{a}{b}$ AND $\frac{c}{d}$ ARE IRREDUCIBLE?

"All right," said Bill. "But what if he knows that one, too?"

"Enough, already!" Mr. Cohn cut him off. "This is getting a little *tedioso*. Tell me, Bill, do you like the fraction 19/95?"

"Looks like a regular old fraction to me. Reducible, ain't it?..."

"Yes, indeedy," Mr. Cohn exclaimed, "it's reducible all right! But how? Look at this. Just cross out the identical digits—the last digit in the numerator and the first digit in the denominator:

$$\frac{1\cancel{9}}{\cancel{9}5} = \frac{1}{5}.$$

Check it out! Here's another one:

$$\frac{1\cancel{6}}{\cancel{6}4} = \frac{1}{4}.$$

And another:

$$\frac{4\cancel{9}}{\cancel{9}8} = \frac{4}{8}."$$

A weird expression came over Bill's face. He grabbed a pencil and wrote with a shaky hand:

$$\frac{5\cancel{3}}{\cancel{3}5} = \frac{7\cancel{7}}{\cancel{7}7} = \frac{3\cancel{33}}{\cancel{33}3}.$$

"Look, mister Cohn," he said in a whisper. "I made that up myself ..."

"These examples are *triv*ial," Mr. Cohn declared pitilessly. "We won't bother with them. Now, you ask your friend Stanley to come up with *one more fraction* that can be simplified like this—and doesn't, of course, consist of identical digits. Remember, Bill—*if there are two-digit numbers in both the numerator and the denominator, there are only four fractions like that!* I've written out three of them for you. Let old Stanley find the fourth one.

WHAT FRACTION DIDN'T MISTER COHN NAME? IS THERE ONLY ONE?

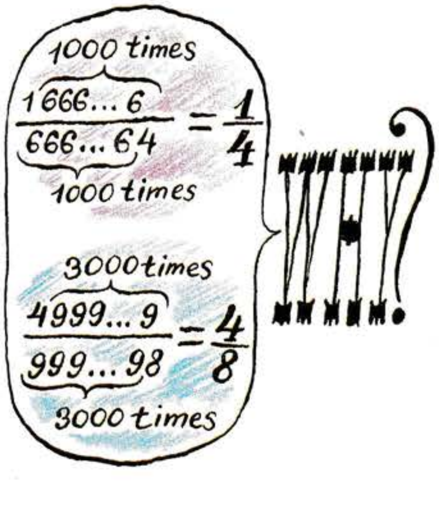
"Oh, yes, I almost forgot," continued mister Cohn. "If you repeat the last digit in the numerator several times, and the first digit in the denominator the same number of times, our incorrect canceling continues to yield correct results. Trust me, Bill, but verify, as someone once said:

$$\frac{1\cancel{9}}{\cancel{9}5} = \frac{1\cancel{99}}{\cancel{99}5} = \frac{1\cancel{999}}{\cancel{999}5} = \frac{1}{5}.$$

You can dare Stanley to calculate this in his head:

$$\frac{16666666666}{66666666664}.$$

"Well, Bill, is that enough?"



"Yes, mister Cohn, thanks. But—" Bill got up, looking sheepish. "I don't have ... I mean ... uh, until I win my money back ..."

"How much?" asked Mr. Cohn, getting right to the point.

Now poor Bill completely and utterly embarrassed. He stood up, wiped the sweat from his forehead, and heaved a big sigh.

"Well, I actually need three seventy-three".

"Three dollars?" Mister Cohn cleared his throat in disbelief. "Take forty for starters."

"No, mister Cohn. For what I have in mind, I need exactly three seventy-three ... That's right. Thanks, mister Cohn ..."

About a week later, as you may have guessed, Stanley came to visit Mr. Cohn. His appearance was eloquent testimony to the fact that he wouldn't be a source of income for Bill or for any other fast-talkin', free-wheelin', wager-lovin' cowpoke.

"Glad to see you," said Mr. Cohn. "Sit down. So tell me, how are things going?"

"What can I say, mister Cohn? Life goes on ... I just stopped by to show you this little doohickey."

Stanley walked up to the table, took a sheet of paper, and wrote:

$$\sqrt{3\frac{3}{8}} = 3\sqrt{\frac{3}{8}}.$$

"Clever little thing, ain't it?" he said. "The number three is incorrectly taken out from under the square root sign, but the answer's correct anyway. But that trick didn't do its magic for me: Bill figured everything out in five minutes and even wrote a generalizing equation! That totally messed up my plans ... But no matter. That isn't why I came. Actually, I came to ask you ... I mean ... well ..."

WHAT FORMULA DID BILL WRITE?

"How much do you need, Stanley?" mister Cohn asked point-blank.

"If it's not asking too much, lend me—" Stanley took a deep breath. "Lend me three seventy-three. I'll pay you back in a couple days."

"You know, Stanley," said Mr. Cohn, handing over the money, "that's the second time this ridiculous sum of \$3.73 has cropped up. What's it all about?"

"Oh, just a lot of nothin', really, mister Cohn. Less than nothin' ..." Stanley turned red, muttered something unintelligible, and ran off.

A few days later Mr. Cohn had to visit a nearby town on business. Over by the general store he noticed an advertisement:

NOVELTY!!!
M.Y. Gardelman's[2]
Living mathematicks for the millions!
Useful and amusing applications!
Only \$3.73!

In the corner of the poster someone had written:

$$\frac{3\cancel{\frac{1}{4}}}{4\cancel{\frac{1}{3}}} = \frac{3}{4}.$$

WHAT FORMULA IS HIDDEN BEHIND THIS SIMPLIFICATION?

[2] An amalgam of the names Martin Gardner (the great American puzzle master and science writer) and Yakov Perelman (the most famous Soviet popularizer of physics and math)—*Ed.*

Constructions with compass alone

—the geometer's version of "one hand tied behind my back"

by Dmitry Fuchs

Among the countless geometric construction problems there are, one often encounters those requiring construction of a figure "with straight edge alone" or "with compass alone." Yet it has been known for centuries that the lack of a straight edge in no way narrows the circle of possible constructions. *Everything that can be constructed with straight edge and compass can be constructed with compass alone.*

The idea of constructing with compass alone was suggested long ago by the Italian scientist Giovanni Batista Benedetti (1530–1590). In 1672 the book *Euclidius Danicus* by the Danish geometer Georg Mohr (1640–1697) appeared. There it was shown that all problems reducible to quadratic equations can be solved geometrically with compass alone. More than a century later the problem was restated and solved by the Italian Lorenzo Mascheroni (1750–1800). Since then the resulting statement has been called the Mohr–Mascheroni theorem. A proof of this theorem will be presented below.

In all the construction problems that follow I'll restrict myself to describing the construction. The proof that it leads to the required result is left to you in each case.

Statement of the result

One can scarcely hope to draw a straight line with a compass, so all our problems will be to construct a certain point (in the plane).

THEOREM. Suppose that a point M can be constructed from the points A_1, ..., A_n with straight edge and compass. Then M can be constructed from A_1, ..., A_n with compass alone.

To prove this theorem we have to know precisely which constructions can be performed with a straight edge—

Art by Dmitry Krymov

Figure 1

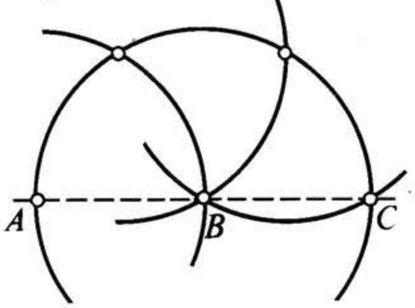


Figure 2

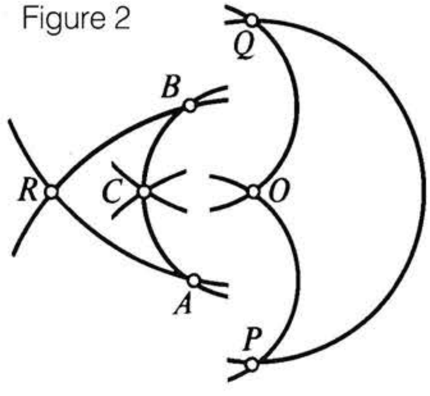
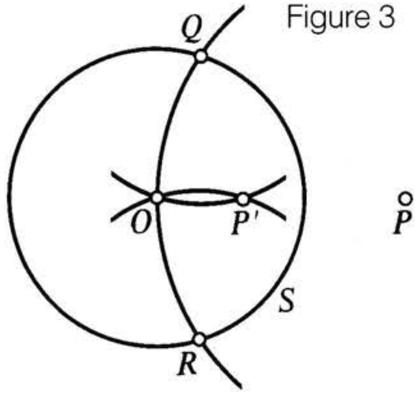


Figure 3



that is, with a ruler. With a ruler we can draw a line through two given points and find its intersection points with straight lines and circles constructed earlier. But since at the very outset we were only given some points, each of our straight lines must have been drawn through two points constructed still earlier; similarly, each of our circles passes through a point constructed earlier and has another point constructed earlier as its center. So in the course of the construction the ruler must be used only to solve one or the other of these elementary problems:

Problem 1. For given points A, B, C, D, construct the point where the straight lines AB and CD intersect.

Problem 2. For a given circle S, its center O, and points A and B, construct the points where the circle S and the straight line AB intersect.

Auxiliary constructions

From now on, by a "construction" we mean a construction with compass alone. We'll begin by solving four auxiliary problems.

Problem 3. Let two different points A, B be given. Construct the point C of the half line AB such that $AC = 2AB$.

Construction (fig. 1). Draw a circle with center B through point A and, starting from point A, mark off three subsequent arcs on this circle spanned by segments of length AB. The end of the third segment is the required point C.

Problem 4. Let a circle with center O and an arc AB on it be given. Construct the point C dividing arc AB into two equal parts.

Construction (fig. 2). Draw circles with centers A and B through point O and then draw a circle with center O and radius AB. Let P and Q be the points where this circle intersects the two circles constructed before; then the arcs OP and OQ are equal to arc AB. Now draw circles with centers P and Q through points B and A, respectively, and take their intersection point R. Finally, draw a circle with radius OR and center P (or Q, it makes no difference). The point where this last circle intersects arc AB is the required point C.

Problem 5. Let a circle S with center O and a point P be given. Construct the point P' on the half line OP such that $OP \cdot OP' = r^2$, where r is the radius of circle S. (Such a point P' is called "symmetric to point P with respect to circle S.")

Construction. Case 1: point P lies outside circle S (fig. 3). Draw a circle with center P through point O. Let Q and R be the points where it intersects circle S. Now draw circles with centers Q and R through point O. The point where these circles intersect (other than O) is our point P'. (This construction is also valid when point P lies inside circle S but the distance from it to point O is greater than $r/2$.)

Case 2: point P lies inside circle S. Using the construction from problem 3, we construct the points P_2, P_3, ... on the half line OP one after another such that $OP_2 = 2OP$, $OP_3 = 3OP$, ... until we reach a point P_n that lies outside circle S. Then, using the previous construction, we find the point P'_1 symmetric to point P_n with respect to S. Finally, we construct the points P'_2, P'_3, ... on the half line OP'_1 such that $OP'_2 = 2OP'_1$, $OP'_3 = 3OP'_1$, ... Then P'_n is the desired point.

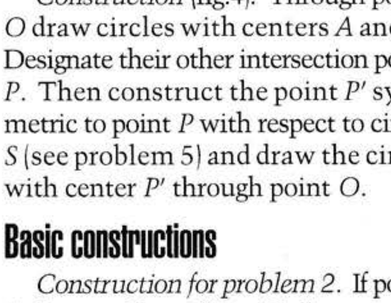
Problem 6. Let a circle S with center O and different points A and B be given. Construct the circle passing through point O and the points where a straight line AB intersects circle S. (We assume that line AB doesn't pass through the center O of circle S and intersects this circle at two points.) Prove that the constructed circle is precisely the set of points symmetric to the points of line AB with respect to circle S.

Construction (fig.4). Through point O draw circles with centers A and B. Designate their other intersection point P. Then construct the point P' symmetric to point P with respect to circle S (see problem 5) and draw the circle with center P' through point O.

Basic constructions

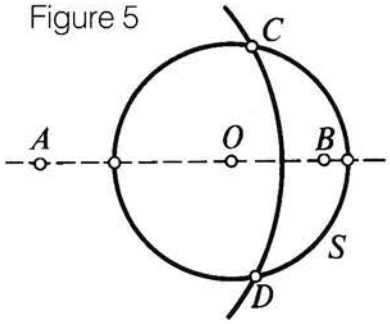
Construction for problem 2. If point O does not lie on line AB, then we can use the construction from problem 6 or even this simplified version: we find the point P exactly as in the problem 6 construction and then draw the circle with center P and radius equal to that of circle S (if we know the center of the circle S, we can measure its radius with a compass); the points where this circle intersects the given

Figure 4



circle S are the desired points. If point O happens to lie on the straight line AB, then this construction won't do: point P will merge with point O. Then we apply another construction (fig. 5): we draw an arbitrary circle with center A (or with center B if $A = O$) that intersects the circle at two points; then we designate these intersection points C and D and divide the arcs CD and DC

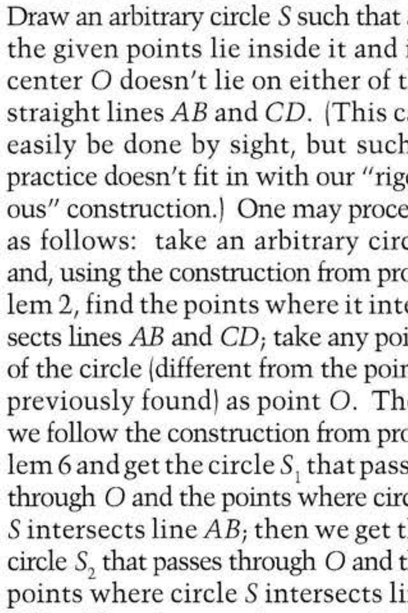
Figure 5



into two equal parts each (see problem 4); the dividing points are the ones we're looking for.

Construction for problem 1 (fig. 6). Draw an arbitrary circle S such that all the given points lie inside it and its center O doesn't lie on either of the straight lines AB and CD. (This can easily be done by sight, but such a practice doesn't fit in with our "rigorous" construction.) One may proceed as follows: take an arbitrary circle and, using the construction from problem 2, find the points where it intersects lines AB and CD; take any point of the circle (different from the points previously found) as point O. Then we follow the construction from problem 6 and get the circle S_1 that passes through O and the points where circle S intersects line AB; then we get the circle S_2 that passes through O and the points where circle S intersects line CD. After that we denote as P the point (other than O) where circles S_1

Figure 6



and S_2 intersect and construct the point P' symmetric to point P with respect to circle S. This is our point.

Concluding remarks

If the construction problem data do not consist of points only, then it may turn out that a ruler is needed to solve it. Consider this problem, for example: given a curve c (drawn in the plane) and two points A, B, find the points where curve c intersects the line AB. In general, it's impossible to do this without a ruler. Nevertheless, a problem like this can sometimes be reduced to problems of the kind we've been looking at, and then they can be solved with compass alone. Here's an important example: a circle is drawn in the plane; find its center. This can be done in the following way: mark three different points A, B, C on our circle; it's well known that one can construct the center of the circumcircle of triangle ABC with straight edge and compass; so this can be done with compass alone. (This problem has a much simpler solution—can you find it?) By the way, we see that in problems 2, 4, 5, and 6 it was necessary to indicate the centers of the given circles in advance.

Not every construction that can be performed with straight edge and compass can be performed with straight edge alone. (A proof can be found in *Numbers and Figures* by Rademacher and Toepliz, for example.) Nevertheless, a theorem by Steiner asserts that any construction that can be done with straight edge and compass can be performed with straight edge alone if a single circle with the center marked is drawn beforehand.

So if you're going to make constructions with straight edge and compass and discover that you've lost your ruler, don't despair: you can do any of your constructions with compass alone. It's worse if you've lost your compass, but everything will be okay if you can borrow one from somebody just for a moment: you draw a circle, mark its center, and return the compass to its owner—now you can do without it.

The worst case, of course, is when you've managed to lose both your ruler *and* your compass. Science can hardly bail you out *then*.

Dmitry Fuchs *is a researcher in I.M. Gelfand's math and biology laboratory. (His article "Bend This Sheet" appeared in the January 1990 issue of* Quantum.*)*

CONTINUED FROM PAGE 21

Do we really need this thing called "temperature"?

Scientists had worked out the concept of temperature and introduced it as a physical quantity long before they understood its real meaning. But now that we know what it means, is it worth keeping this archaic quantity? Wherever we're accustomed to talking about temperature, degrees Kelvin, and so on, maybe we should think of what they really are—the mean kinetic energy of a particle—and measure them in joules.

There are very good reasons why we don't do this.

Do you think a doctor, for instance, would find it convenient to diagnose a patient as sick because of a mean molecular kinetic energy of $6.64 \cdot 10^{-21}$ J? It's easier and more direct to talk about a temperature of 100.4°F.

Besides, using joules instead of degrees might cause confusion. Consider this: 100 J of energy usually implies that 100 J of work can be done. But if the temperature of a body is 100 J/molecule (a fantastic value for temperature, by the way), that does *not* mean we can get that amount of work from it.

So let's stick with degrees—agreed?

Exercises

1. Find the values that correspond to absolute zero temperature in the Fahrenheit and Réaumur scales.

2. Calculate the mean kinetic energy of a molecule at 1,000°K.

3. A nuclear explosion generates a gas ball at a temperature of approximately 20 million degrees Celsius. Find the mean kinetic energy of one particle in the ball.

4. The unit of energy in atomic and nuclear physics is the electron-Volt (eV). Find the mean kinetic energy of a molecule at room temperature in electron-Volts.

5. Calculate the Boltzmann constant in the Fahrenheit scale.

Friezing our way into summer

Introducing a new column on mathematical oddities, quiddities, and just plain surprises

by John Conway

HERE'S A SIMPLE LITTLE ARITHMETICAL GAME. You start with two rows of 1's joined by a zigzag of 1's, as in the example, where the dots mark the places in which further numbers are to be inserted:

```
1   1   1   1   1   1   1   1   1 ...
  1   .   .   .   .   .   .   .
    1   .   .   .   .   .   .
      1   .   .   .   .   .
    1   .   .   .   .   .   .
  1   .   .   .   .   .   .
    1   .   .   .   .   .   .
      1   1   1   1   1   1   1 ...
```

Now you fill in more numbers by the rule that whenever four numbers form a diamond

$$\begin{matrix} & N & \\ W & & E \\ & S & \end{matrix}$$

the product EW must be 1 more than NS, so that

$$E = (NS + 1)/W.$$

In our example, here's what happens:

```
1   1   1   1   1   1   1   1   1
  1   2   2   4   2   1
    1   3   7   7   1
      1  10  12   3   1
    1   3  17   5   2   1
  1   2   5   7   3   1
    1   3   2   4   1
      1   1   1   1   1   1   1   1.
```

Try it with other zigzag patterns. There are some odd things you'll notice when you play with these patterns.

1. All divisions "come out exactly" to give whole numbers.

2. Each line "closes" by getting back to 1.

3. There's no need to continue the lines after this point, because these 1's form a zigzag just like the one you started with but upside down, so that the complete pattern would be a repeating one:

```
1   1   1   1   1   1   1   1   1   1
  1   2   2   4   2   1   3   2   4   1
    1   3   7   7   1   2   5   7   3   1
      1  10  12   3   1   3  17   5   2   1
    1   3  17   5   2   1  10  12   3   1
  1   2   5   7   3   1   3   7   7   1
    1   3   2   4   1   2   2   4   2   1
      1   1   1   1   1   1   1   1   1   1.
```

Can you explain why?

Some similar things happen when you replace multiplication by addition. This time you start with two rows of 0's joined by a zigzag of 0's, and you complete little diamonds by using the formula

$$E + W = (N + S) + 1,$$

so that

$$E = (N + S + 1) - W.$$

Of course, now 0's close the rows:

```
0   0   0   0   0   0   0   0   0 ...
  0   1   2   5   6   4   3   0
    0   2   6  10   9   6   2   0
      0   5  10  12  10   4   1   0
    0   2   8  11  12   7   2   0
  0   1   4   8  10   8   4   0
    0   2   3   6   5   4   1   0
      0   0   0   0   0   0   0   0 ...
```

Can you explain why the numbers never go negative, why they close by getting back to 0's, and why these 0's form an exact copy of the original zigzag, now the right way up?

Some interesting patterns turn up if we drop the requirement that there be a zigzag of 1's joining the top and bottom rows. Multiplicative frieze patterns arise in which all the numbers in any given row are the same and positive. For instance,

$$\begin{array}{ccccccccccc}
1 & & 1 & & 1 & & 1 & & 1 & & 1 \\
 & \sqrt{2} & & \sqrt{2} & & \sqrt{2} & & \sqrt{2} & & \sqrt{2} & \\
1 & & 1 & & 1 & & 1 & & 1 & & 1
\end{array}$$

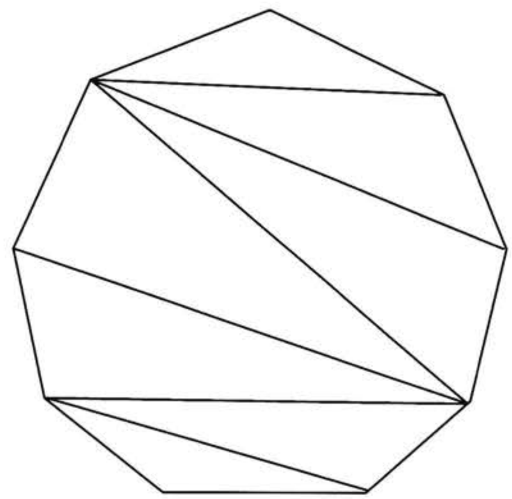
and

$$\begin{array}{ccccccccccc}
1 & & 1 & & 1 & & 1 & & 1 & & 1 \\
 & \sqrt{3} & & \sqrt{3} & & \sqrt{3} & & \sqrt{3} & & \sqrt{3} & \\
2 & & 2 & & 2 & & 2 & & 2 & & 2 \\
 & \sqrt{3} & & \sqrt{3} & & \sqrt{3} & & \sqrt{3} & & \sqrt{3} & \\
1 & & 1 & & 1 & & 1 & & 1 & & 1.
\end{array}$$

You can see that the numbers aren't all whole any more. What are they? Hint: The answer is connected with polygons.

The multiplicative frieze patterns in which all the entries are positive whole numbers are also connected with polygons. Can you see how?

Another hint: The polygon for our first example is this:



John Conway *is a professor of mathematics at Princeton University.*

Walker in a winter wonderland

But since it's almost summer, let's fake the snow, make some ice cream, and watch the smog roll in!

by Alexander Borovoy

JEARLD WALKER'S WONDERful book *The Flying Circus of Physics* has been translated into Russian and has become so popular in the Soviet Union that it sold out long ago, despite several print runs.[1] Answers to many of the questions posed in this book involve setting up clever experiments or carrying out interesting observations. In this article I'll follow up some of Walker's questions and supplement them with some curious stories. I'll also describe some additional preparatory experiments. All these questions and experiments have to something do with cold phenomena—a last gasp of winter before the lazy days of summer set in.

"Squeak, squeak"—Nanook is coming!

The great film director Sergey Eisenstein, famous for the classic film *The Battleship Potemkin*, is also the genius behind the film *Alexander Nevsky*, which is still popular in the Soviet Union though it was shot over

half a century ago. This epic was one of the first superproductions in the movie industry. It's renowned for the alternately rousing, suspenseful, and plaintive background music provided by Sergey Prokofiev and for its rendering of the climactic battle scene on Lake Chud. It was there that Russian forces, led by Nevsky, defeated the better trained and better equipped Teutonic knights in 1242.

Few people, however, are aware that the film's central episode, the battle on the ice, was actually shot at the height of summer! The film shows snow and ice on Lake Chud, while in fact it was as hot as 86°F. Here's how they did it. A large flat area was covered with a mixture of naphthalene and salt. When the actors walked over this "snow," it creaked just as real snow does when it's bitterly cold.

You can easily duplicate the "Nevsky solution" on a smaller scale by filling a plate with a smooth layer of granulated sugar or salt. Press a spoon against it—you'll hear a faint creak. Now wet the granules or melt them slightly on the stove—see if you can get them to creak some more.

Now I'll let Jearld Walker have the floor: "Sometimes snow crackles when you walk in it but only when the temperature is far enough below freezing. What causes the noise, and why does its production depend on temperature?"

How to put the "ice" in ice cream

When ice cream first appeared, its method of preparation was kept strictly secret. At many European courts, chefs tried using snow or ice to freeze the mixture of cream, sugar, and fruit juices—the mixture would be chilled, but it wouldn't freeze. They finally had to resort to what is now called "industrial espionage." And what did these culinary spies find out?

"When my grandmother makes

Art by Edward Nazarov

[1] "What's a flying circus?" you ask? The phrase dates back to World War I and meant airplanes flying in formation. As time went by, it came to include stunt flying in air shows (no doubt because vintage biplanes often take part in them). And who is Jearld Walker? He's a professor of physics at Cleveland State University. His book, published by John Wiley & Sons, is available in paperback—in English!—*Ed.*

homemade ice cream," writes Walker, "she packs ice around the ice cream container, and then she salts the ice." So the secret is plain old salt! "Why does she add the salt?" Walker asks.

Before trying to answer the question, maybe you can conduct a couple of simple experiments. Take some crushed ice from your refrigerator and gradually add salt, mixing it all together. (It would be nice—but not essential for our purposes—to make ice cream during the experiment!) You should be able to get the temperature of such a mixture down to –20°C.

Everyday sodium chloride is used most often for this kind of cooling. But more arcane substances will also do: KCl, $NaNO_3$, and so on. In fact, the champion temperature reducer is $CaCl_2$. If you mix 42 g of this salt with 100 g of crushed ice, you can work the temperature down to –55°C!

So how did your experiments turn out? And what's your answer to Walker's question?[2]

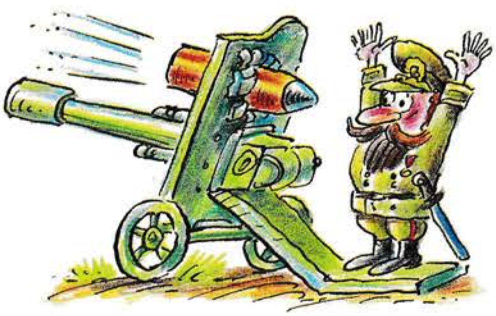
Skating bombs skirt state-of-the-art armor

Engineers often prescribe lubrication to reduce friction. Different materials play this role—sometimes, it seems, the most unexpected and inappropriate. By way of example, I'll tell you about a case where *steel* acted as a lubricant (though we can't exactly call it "cold, hard steel").

At the end of the last century, the English industrialist Harvey sent samples of new armor plates for warships to Russia. During testing the shells from the huge heavy guns, instead of breaking through the plates, crashed against the armor without damaging whatever was behind it. Then the Russians asked that the tests be repeated. In the second series of tests, the shells broke through the armor plates. And in a later series of tests, after certain improvements, the shells began to pierce holes in the armor.

Why did that happen? How had the design of the shell been modified?

The tips of the shells now had special caps made of soft steel. The cap would spread and melt upon impact.



This prevented the shell from cracking, and it served as a kind of lubrication for the shell after it broke into the armored plate. So, under tremendous pressure, steel can serve as lubrication. Admiral Makarov, a talented scientist and as well as a worthy seaman, invented this soft steel cap for artillery shells.

And now for something completely different. "When you are ice skating," Walker asks, "why do your skates slide along the ice surface?" That is to say, what's the lubrication in this case? (Maybe it's not so completely different after all.)

"Through the misty mist and the dusky dusk ..."

Damp winter days are always misty. We're used to thinking of the mist as gray, but it's not always so. Here's what the Russian writer Constantine Paustovsky had to say in his book *The Golden Rose*:

"The French artist Monet came to London and painted Westminster Abbey. Monet was working on an ordinary foggy London day. In Monet's picture the Gothic outlines of the abbey appear through the mist. The picture is now universally recognized as a masterpiece.

"But when the painting was exhibited, it caused confusion among the Londoners. They were astonished that Monet painted mist with a crimson hue, even though everybody knows that mist is gray.

"At first, Monet's impertinence caused indignation. But the indignant Londoners went outside, looked attentively at the mist, and saw for the first time that it was indeed crimson."

Walker also agrees with Monet. "If you've lived in a large city," he writes, "you almost certainly have spent part of your life in a haze. Why are such hazes brown?"

What causes a mist's coloring is not an easy question to answer. It's not by accident that I've saved it for the very end! Take a close look at the color of fog, smog, or haze in your town and in the countryside. Maybe you can even do some experimenting and come up with a method of producing "mists" of different colors. If you do, tell us about it—don't hog your fog! ◙

Alexander Borovoy, *a doctor of physics and mathematics, teaches in the School of Natural Sciences at the Kurchatov Atomic Energy Institute. He has recently worked on problems associated with the Chernobyl accident and was part of a scientific group that examines the damaged nuclear power plant.*

[2]Can you figure out how to lower the temperature of the ice–salt mixture without extracting heat from it?—*Ed.*

Summer study in New York and Tartu, Maryland and Moscow

Students from different countries mix it up academically and socially

JULY 1989 FOUND 15 SOVIET students winging their way to the United States, and a month later 47 American students flew off in exactly the opposite direction. They were all participants in the Science and Mathematics International Summer Institutes sponsored by *Kvant* magazine (Moscow), the International Educational Network (Washington, DC), and the American Association of Physics Teachers. The Soviet group basically consisted of winners of *Kvant*'s problem contest who also had an active command of English. The American students had to pass some challenging tests, but knowledge of Russian was not demanded.

Rather than describe the program in the abstract, we'll let two of the participants give their candid impressions. Ivan Arzhantsev was a student at the Physics and Mathematics School No. 45 in Kiev; Tania Edwards had recently graduated from the Washington International School and now attends Brown University.

OUR TRIP TO AMERICA was preceded by a two-week summer school where we had daily classes in physics, mathematics, and English (including practice in conversational English). This experience was not only useful but pleasant. We had interesting meetings with members of the *Kvant* editorial board, American scientists, and even a businessman.

Of 26 candidates for the trip, 15 students were chosen. The others had the chance to attend the Soviet–American school in Tartu [see the essay below by Tania Edwards—*Ed.*].

On July 11, after an 11-hour flight, we arrived in Washington, where we were met by representatives of the American Association of Physics Teachers. And late that night we were already at the international summer institute on Long Island in New York.

Classes were conducted in three subjects: mathematics, physics, and biology. Dr. Alexander Soifer of the University of Colorado taught mathematics; Dr. Edward Lozansky, director of the summer institute, taught physics along with Dr. Alexander Buzdin; and Dr. E. Trifonov of the Weizman Institute in Israel gave lectures in molecular biology.

The math classes were devoted to combinatoric geometry and graph theory. In addition to traditional problems, several were proposed that were of real scientific interest. And there was a material reward for those who could solve them.

Here are the problems. Let's take a convex quadrilateral of area 1 and a point inside it. The vertices of this quadrilateral and the internal point form 10 triangles. Of these, let's take the triangle with the smallest area. The task is to find the greatest possible value of this area. Solutions to this and a similar problem with a convex pentagon were valued at 20 dollars each. The next problem had to do with convex polygons. In class we had already shown that if we take 6 points inside or on the border of an arbitrary convex polygon of area 1, we can always select three such that the area of the triangle with vertices at these points is less than $1/4$ (condition 1). It had also been proved that, in an arbitrary convex polygon of unit area, we can position 4 points such that the area of any triangle with vertices at these points will be greater than $1/4$. Here's what we had to determine: For which convex polygons is 5 the least number of points needed to satisfy condition 1, and for which—6? A solution to this problem was valued at 50 dollars. (We didn't manage to completely solve either of these problems.)

In the physics classes we solved problems from many branches of elementary physics, and there was a physics tournament in which 12 teams competed.

At the lectures on molecular biology, Dr. Trifonov explained the geometric and chemical structure of DNA and described the latest experiments in studying the inner structure of the cell. The lectures were very interesting and informal. (The level of biological studies in American schools is generally higher than in ours, which cannot be said of the level of physics and math.)

We also heard a series of lectures on a wide variety of topics in physics, brought together under the general heading "Forces in Nature." They were given by the Nobel Prize–winner Sheldon Glashow, and the views of such a great specialist were of tremendous interest to all of us. Another Nobel Prize winner, Dr. Pronin, gave a lecture on methods of determining molecular structure, but we didn't enjoy this lecture nearly as much.

The final days of the summer institute were taken up with math and physics olympiads. The institute's organizers promised the winners free trips to the US or France next summer.[1]

In addition to completing a broad academic program, we also just relaxed a lot during those two weeks, taking part in tennis and chess matches, playing volleyball, and canoeing on the river. Students from four countries participated in the program: the USSR, the US, France, and Switzerland. We really got to know each other because the room assignments ensured an international mix.

During the summer school we drove into New York City, visited the Brookhaven National Laboratory, and hit the beaches of the Atlantic Ocean. We were greatly impressed by our visit to the suburban home of one of the American families, where for several hours we played volleyball and tennis and swam in their pool.

The kids from each country put on show one evening—an improvised performance illustrating something from the past or present of their country. It was my fate to play the role of Ivan the Terrible[2] ...

On the last day of the summer school prizes were awarded to the winners of all the olympiads, tournaments, and contests. A lot of amusing prizes and souvenirs were given away, and there was a little artistic contest in which Mikhail Kapustin garnered the most points by improvising on the piano.

[1]The winners were Soviet students in both instances: Ivar Martin in physics and Ivan Arzhantsev in math.—*Ed.*

[2]The Russian tsar who ruthlessly centralized state power in the 16th century.—*Ed.*

After school let out we took a three-day trip to Washington in vans, stopping at Princeton University, Philadelphia, and Baltimore on the way. But the thing I remember most is our visit to the Trump Plaza casino in Atlantic City, where wallets are emptied by the pitiless laws of probability in the midst of feverish eyes and the tinkle of coins. And finally, Maryland University (right outside Washington), where we spent the remaining 10 days. Seventeen American kids who are planning to visit the USSR came here for three days. We attended several lectures on physics and worked with computers. We took several trips into Washington and visited the National Gallery of Art, the Library of Congress, and NASA headquarters. A highlight was our visit to the Soviet embassy, where we met with Yuri Ossipyan, Vice President of the Academy of Sciences and editor in chief of *Kvant*.

Without a doubt, we got good language practice on this trip, we studied physics and mathematics pretty seriously, but mainly we saw something worth seeing—America!

—Ivan Arzhantsev

LAST SUMMER I HAD the incredible opportunity to participate in a unique program. For the month of August, together with about 40 other Americans, I went to study math and science at the University of Tartu in Estonia.

My experience proved to be far beyond the simply academic. I not only learned from some of the best professors in the USSR Academy of Sciences, I also had the opportunity to interact with Estonians and Russians who taught me about cultures very different from my own and who had a tremendous impact on how I perceive my country and those around me.

The program involved studying for three weeks and then traveling around Estonia, Leningrad, and Moscow for one week. At Tartu, we were received very warmly since we were the first Americans to ever stay there. Each one of us had Estonian buddies who showed us around the town, introduced us to their friends and family, told us about their life and traditions. We, in turn, taught them about America, gave them T-shirts, and shared our music.

In the morning we had math and physics classes taught primarily by Soviet scientists. It was very exciting to be able to learn things we were never exposed to in our schools, surrounded by Soviet classmates. In the afternoons we saw the city, went to hear Estonian folk music, saw museums, played volleyball, and painted cabins on the beach of the main lake as a souvenir to the city. On weekends we went to different places such as the beautiful old city of Tallinn, the capital of Estonia, where we stayed in an Olympic hotel. We saw forests, farms, and the highest point in the republic, ate potatoes around the clock, went to concerts, and experienced the public baths. We also went camping near the city of Pärnu and on the island of Saaremaa, where special permission was required to enter since it was a military base. Endless nights with the guitar by the fire made us oblivious to everything around.

Ann Garlitski of New York works on a physics experiment with Mati Pirn of Estonia at the Science and Mathematics International Summer Institute in New York. (Photo courtesy of Tatiana Lozansky)

Of course, these experiences were incredible, but what was most important is that we got to really see how the political changes, spurred by *glasnost*, were taking place all over the country. Contrary to our expectations, we were received by mayors and by well-known activists, one of whom is a current Estonian representative to the Congress of People's Deputies. We even had opportunities to ask them questions about their personal views and plans. Such contact was very exciting and led to numerous political discussions with the Soviet students.

On the negative side, in the camp itself we soon found ourselves right in the middle of the ethnic conflict between the Estonians and the Russians, who were also part of the program. The Russians were treated quite poorly. They lived separately from us, didn't go on any of the trips, and had no one cheering for them at the games. When we approached Estonians with questions, we received bitter responses: "They were the occupiers, they don't deserve to be treated well." Our friendship with the Russians, however, greatly improved in the second part of the program when we started to realize what was happening. It was especially good when we traveled to Russia—to the ancient fortress-city of Pskov, the beautiful palaces of Leningrad, and the museums and shows of Moscow. Like the Estonians, the Russians were eager to show us about their life and to learn about ours.

When it came time to leave, no one could believe that only a month had gone by. It felt like we learned immense amounts not only in classes but from our new friends that we couldn't say good-bye to. When we arrived in Paris, we walked around as if in a fairy-tale world, bewildered by the overflowing markets and the oblivious people walking by. When our plane finally landed in New York, we all burst into applause. We came home different—more objective and more knowledgeable about the world around us. And I think the other participants would agree that the feeling we all shared was one of gratitude for the country we live in.

—Tania Edwards

The American Regions Mathematics League

Teamwork is the key in this summer competition

by Mark Saul

EACH JUNE, OVER ONE THOUsand of North America's most able and interested young mathematicians gather for a celebration of their field of interest—the annual competition of the American Regions Mathematics League (ARML).

This competition is the largest onsite event of its kind in North America, drawing more than sixty teams of students from all over the United States and from Canada. The teams engage in a daylong sequence of events, comprising a variety of individual and team contest questions.

Students begin arriving at the campus of the host college on a Friday afternoon. For the past several years this has been Pennsylvania State University. Some years, telecommunications have allowed simultaneous competition at this site and at Duke University in Durham, North Carolina. Friday evening is devoted to recreational activities, including talks by noted mathematicians and teachers.

The competition begins in earnest on Saturday morning with the "power question." This event, unique to ARML, consists of a single complex question, often broken down into separate parts. The team must work together for an hour to produce a single answer paper, typically including examples, proofs, extensions, applications, and generalizations.

Team questions follow. Teams attack this set of ten "quick" problems by dividing up the work in any way they wish. This event is followed by a more traditional event, consisting of pairs of questions to be answered individually. The team score will be the sum of the individual scores.

A relay round follows a break for lunch. In this event, the team of fifteem members is divided into three or five subteams, working separately on chains of questions. The answer to each question forms part of the next question, and only the final answer is scored. Since an error in any one question can skew the final answer, success in this event demands a balanced team in addition to nimble individual minds.

At the conclusion of the competition, scores are tallied and prizes are awarded to high-scoring teams and individuals. Publishers and professional organizations donate books or plaques. An awards ceremony honors the winners.

Teams for ARML are formed in a variety of ways. Many large cities and suburban areas have longstanding traditions of these competitions, and teams are often formed as a result. Rural areas of a state often band together to send a team. Some teams represent a single state, and often a state chooses two or more statewide teams to represent it. Teams from Canadian provinces have recently joined those from the United States. Criteria for team membership are, for the most part, left up to local tradition.

The backgrounds of team coaches and chaperones vary as much as the teams themselves. Frequently, a teacher or math team coach assembles the team, but sometimes the coach is from the faculty of a local college or university. Many teams are accompanied by alumni of ARML, who have returned as coaches to rekindle the enthusiasm of their days on the team. Coaches give generously of their time—and sometimes of their money—in support of their teams.

An interesting by-product of the ARML on-site format, and particularly of the power question event, is the possibility of continued investigation even after the conclusion of the contest. ARML contest material is designed by the authors to be open-ended. Even the most humble relay question can often have an interesting extension. ARML contest questions have provided students with ideas for prize-winning research projects and fine expository papers.

But the most important source of stimulation at the ARML competition, and the one source that is difficult to find in other mathematical contests and events, lies in the interaction of the students themselves. Teammates share ideas and pool intellectual resources. Students on rival teams meet informally, exchanging solutions to problems and posing new challenges to each other. Some play the game of trying to find flaws in the day's contest problems. These interactions lead to the formation of a network of student mathematicians. Friendships started over lunch tables at ARML are often continued long after the day's end. For many students, the shared enthusiasm of the ARML competition becomes a central experience in their lives.

For more information on ARML, please write to Steven Adrian, ARML Executive Director, R.D. 5, Box 133, Kings Ridge Road, Mahopac, NY 10541 USA, or Barbara Rockow, ARML Corresponding Secretary, Bronx High School of Science, 75 West 205th Street, Bronx NY 10468 USA.

Mark Saul is the computer consultant/coordinator for Bronxville School in New York and president of the American Regions Mathematics League.

The 1989 ARML power question

A convex n-gon will be called "Pythagorean" if it has integer sides, it is cyclic, and its longest side is a diameter for its circumscribing circle. It shall be denoted by Pn or Pn:$(a,b,\ldots)$, where $a, b, \ldots$ are the lengths of its sides. We shall always use the letter d for its longest side. (Thus $P3$ is a Pythagorean triangle. Notice that it would be a right triangle.)

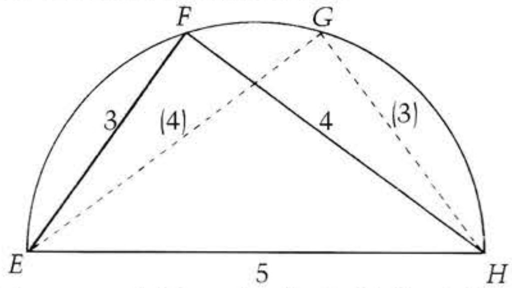
I. There is a theorem that states (in part): If a prime d is the hypotenuse of a Pythagorean triangle, then d^2 is the hypotenuse of two Pythagorean triangles, d^3 is the hypotenuse of three Pythagorean triangles, and so on.

I.A. Find two $P3$'s for which $d = 25$.

I.B. Find three $P3$'s for which $d = 125$.

II. Ptolemy's Theorem says: A convex quadrilateral is cyclic if and only if the product of its diagonals equals the sum of the products of the two pairs of opposite sides.

II.A. If the $P3$:(3,4,5) is reflected as shown in the figure, a quadrilateral $EFGH$ can be formed. (It will not be a $P4$, as FG is not an integer.) Multiplying each side by 5 produces a $P4$. Find the sides of this $P4$.



II.B. Find a $P4$ with two equal sides and with $d = 25$ that is different from the answer to part II.A. (Note: Two Pn's are *not* considered different if their sides are equal but in a different order.)

II.C. Show that a Pn must exist for all integers $n \geq 3$. (This may be done by describing how to create such a Pn.)

III.A. For the $P3$:(a,b,d), $d^2 = a^2 + b^2$. Prove that for the $P4$:(a,b,c,d), $d^2 > a^2 + b^2 + c^2$.

III.B. Given the $P4$:(a,b,c,d), prove that if $d > 2$, then d must be composite.

III.C. If all the diagonals of a Pn are integers, we will call it "super Pythagorean" and denote it by ${}^{S}Pn$.

III.C.1. Show that the area of any ${}^{S}P4$ must be an integer. (Hint: One approach might be to first show that the area of any ${}^{S}P4$ must be rational and its perimeter must be even.)

III.C.2. Assuming that the area of every ${}^{S}P3$ and every ${}^{S}P4$ is an integer, show that (for all $n > 4$) the area of every ${}^{S}Pn$ must be an integer. (You may do this part even if part III.C.1 has not been completed.)

SOLUTION ON PAGE 62

The 1989 ARML contest questions were created by Gilbert W. Kessler, Canarsie High School (retired), Brooklyn, NY; Harry Ruderman, Hunter College Campus School (retired), New York, NY; and Larry Zimmerman, Brooklyn Technical High School, Brooklyn, NY.

Bulletin Board

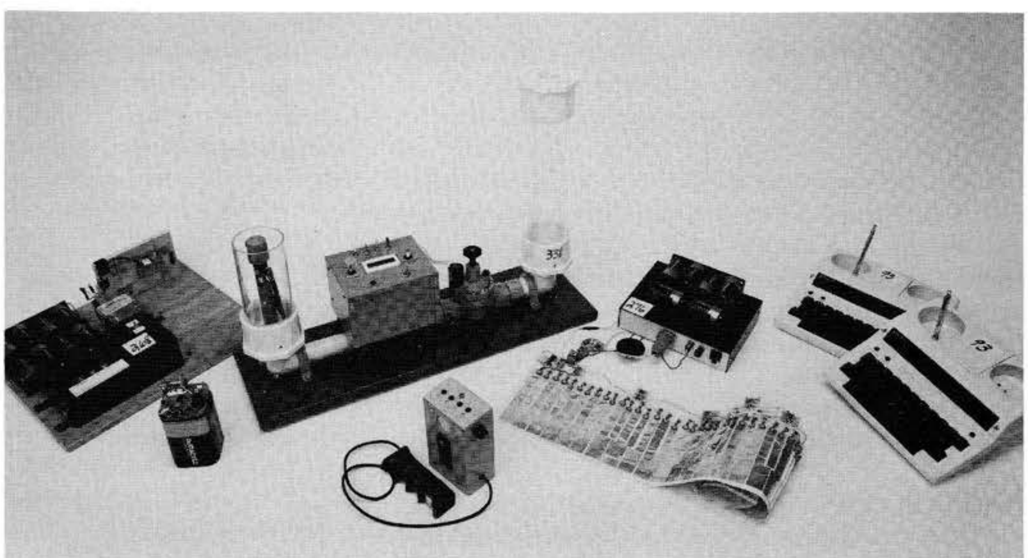
Battery-driven ideas earn scholarships

The eighth annual Duracell NSTA Scholarship Competition really got the creative juices flowing. To enter the competition, a student had to design and build a device that is educational, useful, or entertaining and is powered by one or more Duracell batteries. Jerry Pratt, a senior at Ashland High School in Wisconsin, was awarded a $10,000 scholarship for his "Knock-Out" Keyless Door Lock. You program the device's microcomputer by tapping out a rhythm (a favorite tune, the Morse code of someone's name, whatever), and the door can't be opened from the outside unless the encoded knock is repeated. Jerry says his Knock-Out lock is an improvement over other electrical locks because it's tamper-proof from the outside.

Five other students each received a $5,000 scholarship. Theodore Gielow, a senior at Newport Harbor High School in California, invented the High-Efficiency DC Light Dimmer, which decreases the brightness of a light bulb (for example, a night light on a camping trip) without wasting power by reducing the time during which current flows rather than cutting back the current. Albert Masaki Huntington, a senior at Pleasant Hill High School in Oregon, came up with the Portasynth, a compact, lightweight, highly efficient keyboard with 63 keys designed to be practical for trips into space. Matthew Jared Klam, a senior at the Wisconsin School for the Deaf, devised a Walkie-Talkie Device for the Deaf, which incorporates a portable screen and keyboard for sending and receiving messages. Robbie Glen Seibert, a junior at Grand Rapids High School in Minnesota, invented a Speech Leader—a talking compass that uses a directional sensor, speech processor chip, leveler, and alarm clock. Kurt Thorn, a sophomore at Shoreham–Wading River High School in New York, created the Multipurpose Angle Measurer for recording the degree of mobility in the joints of persons who are in physical therapy after an accident, for example, or who are arthritic. It's based on a pendulum hooked to a variable resistor.

Ten students were given $500 scholarships, and 25 students received $100 cash awards. The top winners explained their devices to an audience of educators and scientists at an awards luncheon during the recent NSTA annual convention in Atlanta, Georgia.

To find out how to enter the ninth annual Duracell NSTA Scholarship Competition, write to Katie Rapp, National Science Teachers Association, 1742 Connecticut Avenue NW, Washington, DC 20009, or call 202 328-5800.



Award-winning devices from the 1989 Duracell NSTA Scholarship Competition (left to right): "Knock-Out" Keyless Door Lock, High-Efficiency DC Light Dimmer, Environmentally Directed Watering Device, Speech Leader, Portasynth, Walkie-Talkie Device for the Deaf

Free 3-D report brings molecular puzzles to life

A colorful new report from the Howard Hughes Medical Institute, *Finding the Critical Shapes*, shows how the answers to many riddles of modern biology and medicine lie buried in the intricate, three-dimensional shapes of our bodies' molecules. Researchers in the field of "structural biology" increasingly depend on computers not only to do their mathematical computations but to help them see their results. This beautifully illustrated report includes a stereo viewer to help the reader visualize molecular shapes in three dimensions. This kind of view provides insights into the mechanisms of previously untamed diseases and helps scientists design new drugs to combat them.

For a free copy of *Finding the Critical Shapes*, write to the Howard Hughes Medical Institute, Communications Office, 6701 Rockledge Drive, Bethesda, MD 20817.

SOLUTIONS

Math

M6

For each of the three given circles, draw radii from its center to the points where it intersects the other two (the thin lines to points A, B, C, H in figure 1). Three rhombi appear (with common vertex H) whose sides are all equal to r. We can imagine them as representing three faces of a parallelepiped with common vertex H. Draw the three other faces—that is, three new rhombi with common vertex O. Their new sides (the thick lines OA, OB, OC in figure 1) are also equal to r. And this is precisely what has to be proved. (Notice that our reasoning also holds when B lies outside triangle AHC.) Figure 2 creates a more convenient view of the situation by breaking apart the parallelepiped and eliminating the distracting portions of figure 1.

We'll leave it to you to find other solutions to this problem and lots of interesting facts of triangle geometry related to it. You may have noticed, for example, that triangle ABC is congruent to the triangle with vertices at the centers of the three given circles; that B is the orthocenter of triangle AHC; that the points symmetrical to B relative to the sides of triangle AHC lie on its circumcircle; and that all four circumferences play equivalent roles—that is, every group of three circles has a common point.

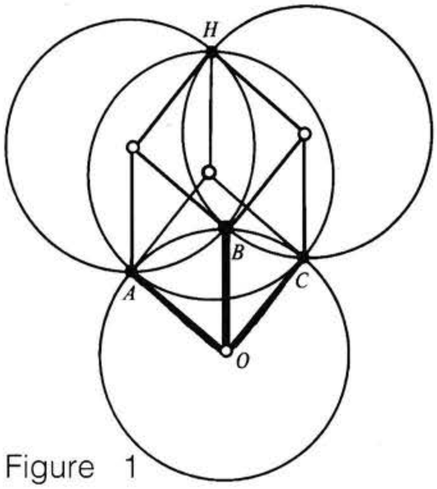


Figure 1

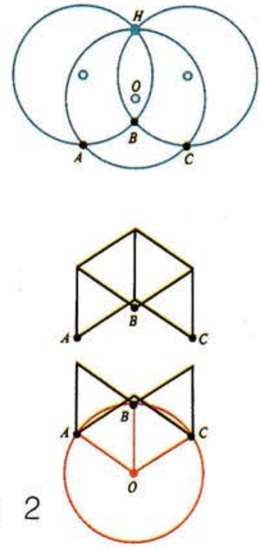


Figure 2

You can find a detailed discussion of this problem in the remarkable book *Mathematical Discovery* by George Polya.

M7

Imagine that the numbers $x_1, x_2, ..., x_n$ are written clockwise around a circle. Let k be the number of times the number 1 is followed (clockwise) by -1 in the sequence $x_1, x_2, ..., x_n, x_1$ (which we'll call the "basic sequence"). Notice that the number of reverse changes (from -1 to 1, clockwise) is the same. (Why?) The total number of negative 1's among n products x_1x_2, x_2x_3, ..., x_nx_1 (the "product sequence") is equal to the number of sign changes in the basic sequence, which is $2k$. But the sum of the members of the product sequence is 0 only if exactly half of the n products equals -1, and so $n = 4k$.

M8

The strip between the exterior polygon P_0 and the interior one P_1 can be cut into rectangles with altitude 1 (whose bases are the sides of P_1) and quadrilaterals left near the vertices of P_0 when the rectangles are removed (fig. 3). Take these quadrilaterals and fit them together by parallel translations so that their inner vertices (the ones originally on P_1) all coincide. We then obtain a polygon P circumscribed to a circle with radius 1. Its vertex angles are the same as those of P_1 (or P_0) and its sides are equal to the differences between the sides of P_0 and P_1.

If the sides of P_1 are proportional to the corresponding parallel sides of P_0 with a coefficient $k < 1$, then the sides of P are clearly proportional to those of P_0 with a coefficient $1 - k$. So P_0 and P_1 are similar to P, and all of them are circumscribed to a circle.

In this solution we've tacitly assumed that the similitude between P_0 and P_1 is the natural one (that is, the correspondence between the sides is the one induced by the outward motion). It can be proved that if any similitude exists, the natural one exists too. We'll give another solution, however, that doesn't rely on this assumption.

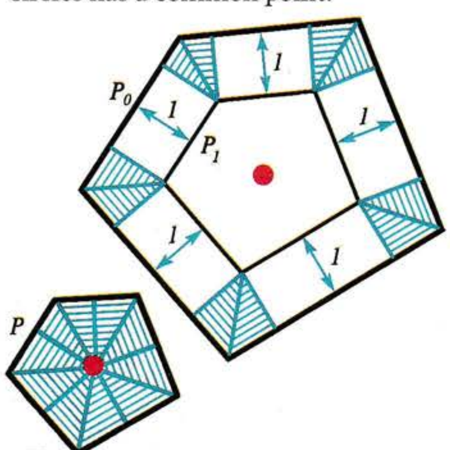


Figure 3

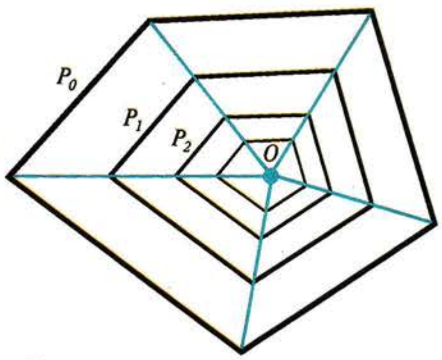


Figure 4

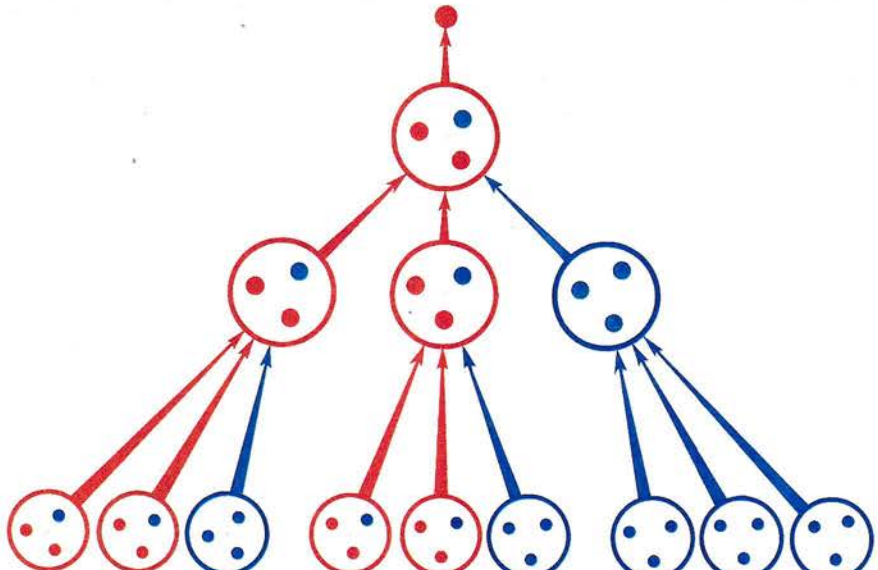


Figure 5

Let the inner polygon P_1 be the image of P_0 under a similitude with a coefficient $k < 1$. Then the image P_2 of P_1 under this similitude can be constructed from P_1 in the same way as P_1 was constructed from P_0 but with a k-fold reduction of all distances—that is, by moving the sides of P_1 the distance k toward the interior (fig. 4). Then the polygon P_3 (the image of P_2) can be constructed by moving the sides of P_2 the distance k^2, and so on. The intersection of all the polygons $P_0, P_1, P_2, \ldots$ is a single point O. Each side of P_0 moves the distances $1, k, k^2, \ldots$ toward O, so that the total distance to O is equal to the sum of the infinite geometric progression $1 + k + k^2 + \ldots = 1/(1 - k)$. It follows that P and P_1 are both circumscribed to circles with center O.

M9

Let's call Miraflores's supporters "reds" and the other voters "blues." The situation illustrated in figure 5, where eight red voters out of a total of 27 voters guarantee the victory of the reds in a 3-stage election, clarifies the main principle: in each group where the reds win, they have a minimal majority, whereas in each group won by the blues the vote is unanimous. Using this principle, we see that M red votes out of a total of N will win in an appropriately organized r-stage election if the ratio M/N can be represented as a product of r fractions each larger than 1/2.

Thus, $8/27 = (2/3)(2/3)(2/3)$ in the previous example. Since we have

$$\begin{aligned} &(3/5)^7 \cdot (9/16)^2 \\ = \ &3^{11}/(5^7 \cdot 2^8) \\ = \ &177{,}147/20{,}000{,}000, \end{aligned}$$

it follows that a well-organized 9-stage election will guarantee Miraflores's reelection with less than 200,000 red voters. An even smaller number of reds, $164{,}025 = 3^8 \cdot 25$, is sufficient in a 10-stage election, since

$$(3/5)^7 \cdot 3/4 \cdot (5/8)^2 = 164{,}025/20{,}000{,}000.$$

M10

You may have found it easy to guess and verify the answer: the portions of milk are 0, 1/7, 2/7, 3/7, 4/7, 5/7, 6/7 liter. When the dwarf represented by the solid circle in figure 6 has given his portion of milk to the others, the distribution turns out to be the same as the initial one but rotated by one-seventh of a full turn. So after seven iterations (repetitions), the original distribution is recovered.

But it's not so simple to prove the answer's uniqueness. Let $6x_k$ be the portion of milk that the kth dwarf pours out (x_k liters to each of the other dwarfs), where $k = 1, 2, \ldots, 7$. Without loss of generality we can assume $x_1 \geq x_k$ for all k, since all the dwarfs are equally nice. After seven iterations the first dwarf gets back all his milk; therefore,

$$6x_1 = x_2 + x_3 + \ldots + x_7.$$

Since $x_1 \geq x_k$, this is possible only if $x_1 = x_2 = \ldots = x_7$. Suppose x is the smallest portion of milk added; then by assumption

$$x + 2x + \ldots + 6x = 3,$$

so that $x = 1/7$, and the answer follows.

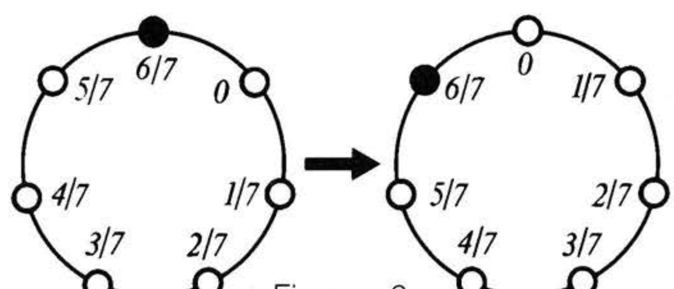


Figure 6

Physics

P6

In order to avoid a collision, the astronaut must change the velocity of the spaceship so that the angle between the initial direction toward the asteroid and the new heading is greater than the angle α_0 determined by the condition

$$\begin{aligned} \sin \alpha_0 &= \frac{d/2}{1 + d/2} \\ &= \frac{d}{2l + d} \\ &\cong 0.292 \end{aligned}$$

(see figure 7). After the additional velocity Δv is imparted to the ship, maximum deviation from the initial heading is provided if the vector $\mathbf{\Delta v}$ turns out to be perpendicular to the vector $\mathbf{v}_1 = \mathbf{v} + \mathbf{\Delta v}$ (fig. 8)—that is, if

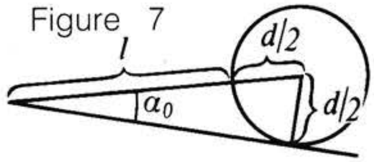
$$\sin \alpha = \Delta v/v = 0.3.$$

So by switching on the emergency engines, the astronaut can change the heading of the ship by an angle $\alpha > \alpha_0$, and there will be no collision with the asteroid.

P7

Let's consider the raft's motion in a frame of reference anchored to the current—that is, moving at a speed of $\mathbf{u}$. In this context the raft has an initial velocity $\mathbf{v}' = \mathbf{v} - \mathbf{u}$ and moves in a

Figure 7



straight line. Here the raft's velocity $\mathbf{v}'$ is reduced because of the force of water resistance. (If the resistance were absent, in time t the raft would be at point c with coordinates $x_c = ut$, $y_c = vt$.) The displacement of the raft relative to the riverbank in time t is a combination of its displacement $\mathbf{s}_r = \mathbf{v}'t$ relative to the water and the displacement of water $\mathbf{s}_w = \mathbf{u}t$ (fig. 9). During time $2t$ the water displacement will be twice as great. From point O let's measure off a segment with a length of $2s_w$ along the x-axis and draw a straight line parallel to $\mathbf{s}_r$ (through the point with coordinate $2s_w$). This line intersects the raft's trajectory at the point marked by the blue cross—the raft will be here at time $2t$ after the raft was launched.

Similar constructions can be used to find where the raft will be at times $3t$, $4t$, and so on.

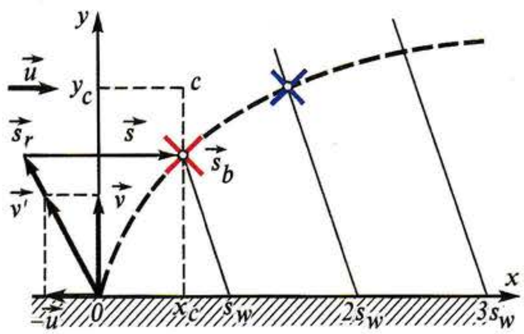


Figure 9

P8

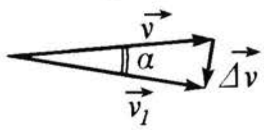
It's clear from the condition that the power from the heating element is equal to the energy leaving the water and entering the environment over time. (The temperature of the water doesn't change with time while the element is on.) So if the heating element is turned off, the energy transferred over time by the water will be 100 W. Recalling that heat flow $Q = Cm \cdot \Delta T$ and that $Q = Pt$ (where C is specific heat, m is mass, T is temperature, and P is power), we find that the water will cool down by one degree during the time

$$t = \frac{Cm \cdot \Delta T}{P} = \frac{(4.2 \cdot 10^3 \text{ J/kg} \cdot {}^\circ\text{C}) \cdot 1 \text{ kg} \cdot 1^\circ\text{C}}{100 \text{ W}} = 42 \text{ s}.$$

P9

From considerations of symmetry, it follows that the induction $\mathbf{B}_O$ of the field created in the center of the cube

Figure 8



(at point O) by the current I flowing in the circuit $ABCDA$ (fig. 10) is parallel to the y-axis. In fact, by rotating it 90° about the y-axis, circuit $ABCDA$ is transformed into itself, and so the vector $\mathbf{B}_O$ should also be transformed into itself. Therefore, by the corkscrew (or right-hand) rule the vector $\mathbf{B}_O$ must be directed along the y-axis. In projections on the x-, y-, and z-axes,

$$\mathbf{B}_O = (B_x, B_y, B_z) = (O, B_O, O).$$

To find the field $\mathbf{B}$ created by the current flowing in the circuit $ABCGHEA$ (fig. 11), we use the superposition principle. Notice that exactly the same distribution of current on the edges of the cube will be obtained if we take three circuits, $ABCDA$, $DCGHD$, and $ADHEA$, in each of which the current is I. The combination of these three circuits creates the required field B in the center of the cube. According to the superposition principle, B is equal to the vector sum of the three fields created by each of the circuits:

$$\begin{aligned}\mathbf{B} &= \mathbf{B}_{ABCDA} + \mathbf{B}_{DCGHD} + \mathbf{B}_{ADHEA} \\ &= (O, B_O, O) + (-B_O, O, O) + (O, O, B_O) \\ &= (-B_O, B_O, B_O).\end{aligned}$$

So the vector $\mathbf{B}$ is directed along the cube's principal diagonal **DF** and its value is $\mathbf{B}_O 3^{1/2}$.

The induction $\mathbf{B}$ can be found by simply adding together the fields of all six edges of the circuit, but this requires more cumbersome calculations.

P10

You might proceed in the following way. Draw two points on a sheet of paper very close to each other—for example, at the distance $l \cong 1.5$ mm. Moving the sheet away from your eyes, estimate the distance at which the points merge into one point. If you have normal eyesight, this distance should be $d \cong 2$ m. The angle $\alpha \sim l/d$ is the least angular distance between the points at which they are distinguished by the eye as two objects. This angle characterizes your eyesight.

The rails will seem to merge when the angular distance between them is α. If the rails are $l' \cong 1.5$ m apart, they seem to come together at the distance $d' \sim l'/\alpha \sim l'd/l \cong 2$ km.

Brainteasers

B6

Keeping the goat in the boat, take the dog and then the cabbage across the river. Take the goat back and leave it on the riverbank (all by itself). Take the two wolves across the river. Return with the dog. In the last crossing take the dog and the goat across.

B7

The problem is based on the fact that adding the mean value of a set of numbers to this set gives a set with the same mean value.

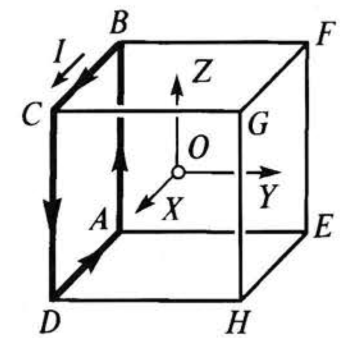


Figure 10

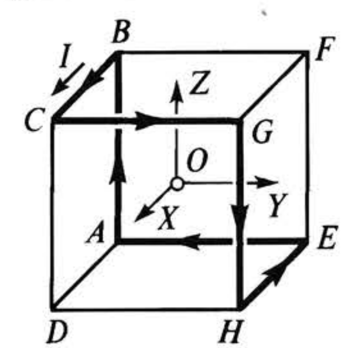


Figure 11

B8

$$\begin{array}{r} 285 \\ \underline{\times \quad 39} \\ 2565 \\ \underline{+\ 855\ \ } \\ 11115 \end{array}$$

B9

The diver can tell which way is up by watching the bubbles coming out of his breathing apparatus or by dropping a pebble.

B10

Combustion occurs when there is an influx of oxygen. Under ordinary conditions on Earth the influx of oxygen is due to convection: near the flame heated air, which is lighter, ascends together with the products of combustion; colder air, containing oxygen, takes their place. In the state of weightlessness there will be no convection, and the flame will die from a lack of oxygen.

Art by Edward Nazarov

Kaleidoscope

1. Particles of smoke take part in Brownian motion and move off little by little so that the density of the smoke decreases.

2. The Brownian motion of the fat droplets is weakened.

3. The number of collisions of a liquid's molecules with the surface of a particle is proportional to the surface area, whereas the mass of the particle is proportional to its volume. So the larger the particle, the harder it is for molecules to move it. A Brownian particle must also be small enough that collisions with molecules will be uncompensated.

4. The rate of diffusion increases as temperature increases.

5. Because of frequent collisions, the molecules move in zigzags. Their actual paths are much longer than their perceived change of position would imply.

6. Water vapor diffuses slowly through the film of lacquer so that the wood dries evenly throughout and the ball doesn't crack.

7. Since the atmosphere at those heights is very sparse, the number of molecules per unit of volume is too small to impart an appreciable amount of energy when they collide with a satellite.

8. If the Moon ever had an atmosphere, it disappeared over the eons of its existence. Amid the myriad molecules of atmosphere, there were always some whose velocity of thermal motion would achieve the escape velocity for the Moon.

9. In putting the pieces back together, it's practically (as opposed to theoretically) impossible to position the fracture surfaces at a distance where the forces of molecular attraction will be "noticeable."

10. Because of close contact over a long period of time, the atoms of the nut and bolt intermix along the boundary because of diffusion and the two pieces "lock."

11. Work to change the level of a liquid in a capillary is performed by the energy of molecular interaction.

12. The pressure would increase.

ARML power question

I.A. (7,24,25) and (15,20,25), the second of which is five times the sides of (3,4,5).

I.B. (35,120,125), (75,100,125), and (44,117,125). The first two come from the (7,24,25) and (3,4,5) triangles. A general approach would be to use the fact that the expressions $k(m^2 - n^2)$, $k(2mn)$, and $k(m^2 + n^2)$, where k, m, and n are positive integers with $m > n$, produce all Pythagorean triplets. Setting $m^2 + n^2 = 125$ leads to $m = 10$, $n = 5$, which produces (75,100,125), or $m = 11$, $n = 2$, which produces (44,117,125). Setting $m^2 + n^2 = 25$ leads to $m = 4$, $n = 3$, producing (7,24,25), for which we then use $k = 5$.

II.A. Ptolemy's Theorem produces $FG = 7/5$. The answer is (15,7,15,25) (in any order).

II.B. Using sides $(a,x,a,25)$, where each diagonal is $(625 - a^2)^{1/2}$; leads to $x = 25 - 2a^2/25$. Positive integral x's come from $a = 5$, 10, or 15. The last leads to the solution in II.A. The other possibilities are (5,23,5,25) or (10,17,10,25). (Either answer, in any order, is acceptable.)

II.C. Apply the theorem given in part I, for any appropriate d, to find $n - 2$ $P3$'s of hypotenuse d^{n-2}. Build these triangles in a semicircle of diameter d^{n-2}. Connecting successive points on the semicircle produces an n-gon. Successive applications of Ptolemy's Theorem to find each side of the n-gon shows each must be rational. Multiplying all sides by the least common denominator involved produces a Pn.

III.A. On diameter AD, draw the $P4$ $ABCD$, with $AB = a$, $BC = c$, $CD = c$, $AD = d$, $BD = e$, and angle $BCD = \theta$. The law of cosines shows that $d^2 = a^2 + e^2 = a^2 + b^2 + c^2 - 2bc \cos\theta$. Since θ must be obtuse, $\cos\theta$ is negative, so d^2 is greater than $a^2 + b^2 + c^2$. Notice that this easily extends to any Pn.

III.B. Using the $P4$ described in solution III.A, with $AC = f$, Ptolemy's Theorem yields $bd + ac = ef = [(d^2 - a^2)(d^2 - c^2)]^{1/2}$. Squaring both sides, simplifying, and dividing each term by d produces $b^2d + 2abc = d^3 - a^2d - c^2d$. Since all are integers, d must divide $2abc$. If d is a prime greater than 2, it must divide at least one of the a, b, c. If

$d|a$, for example, then $a \geq d$, which is impossible. So d is composite or equal to 2; therefore, d is composite.

III.C.1.(1). Using the $P4$ described in solution III.B, we first show that its area must be rational. The area of triangle $ABD = ae/2$, which is integral, since one of the legs must be even (see solution I.B); the area of triangle $BCD = (bc/2) \sin \theta$, which must be rational since $\sin \theta = e/d$ (by the extended law of sines); thus, the area of the $P4$ is rational.

III.C.1.(2). We next show that the perimeter of the $P4$ must be even. We first note that all primitive Pythagorean triplets (generated by $m^2 - n^2$, $2mn$, and $m^2 + n^2$, where m and n are of opposite parities) must yield an odd hypotenuse and exactly one even leg; if a Pythagorean triplet has an even hypotenuse, each leg must have *at least* the same degree of evenness as the hypotenuse. Now by Ptolemy's Theorem, $b = (ef - ac)/d$, and we are given that this is integral. There are now two basic possibilities: (1) d can be even (then both ef and ac will be of a higher degree of evenness, making b even and the perimeter even) or (2) d can be odd (then either e and f are even while a and c are odd, making b odd; or e and f are odd while a and c are even, making b odd; or e and f are of opposite parities, as are a and c, making b even; in each of these situations the perimeter will be even). Thus, the perimeter is even in every case.

III.C.1.(3). With s as the semiperimeter, Hero's Formula for an inscribed quadrilateral gives area $= [(s-a)(s-b)(s-c)(s-d)]^{1/2}$; since s will be integral, the area is the square root of an integer. But for the area to be rational also, it must actually be integral!

III.C.2. This is done by induction. We will just indicate the basic approach here: Given the ${}^{S}Pn\ ABCD \ldots TUV$, on diameter AV, let the area of triangle ABC be K, the area of $ACD \ldots TV$ be K', and the area of TUV be K''. We assume that the area of any ${}^{S}P(n-1)$ and ${}^{S}P(n-2)$ are integral (this is our "extended" induction assumption). Then $K + K'$ is integral, and K' is integral, so K must be integral. Since $K' + K''$ must also be integral, $K + K' + K''$ is integral.

CHECKMATE!

Symmetry on the chessboard

Sometimes it happens by chance, sometimes by design, but it always has a certain charm

by Yevgeny Gik

THE "SYMMETRY MOTIF" arises often both in the composition of chess problems and in actual play. This geometric theme is a lot of fun and seemingly inexhaustible.

Here's a game from a published collection of chess curiosities. It was played at the championship of the USSR in Moscow in 1931.

M. Botvinnik–N. Ryumin

1.	d2-d4	d7-d5
2.	c2-c4	c7-c6
3.	Ng1-f3	Ng8-f6
4.	e2-e3	e7-e6
5.	Bf1-d3	Nb8-d7
6.	0-0	Bf8-d6
7.	Nb1-d2	e6-e5
8.	e3-e4	0-0
9.	c4xd5	c6xd5
10.	e4xd5	e5xd4

The opponents had played out the Slavic Defense, and there was no desire to copy moves on black's part, let alone white's. But after ten moves the players were no doubt startled to see the perfect symmetry on the board, as well as the unusual concentration of pieces on the d-file. Nevertheless, the preference in a symmetrical position generally goes to the player whose move it is. In this case, Botvinnik skillfully makes use of this advantage, ably defending his d-pawn while attacking his opponent's.

11.	Nd2-e4!	Nf6xe4
12.	Bd3xe4	Nd7-c5

(12. ... Nd7-f6 would have been better.)

13.	Be4-c2	Bc8-g4
14.	Qd1xd4	Bg4xf3
15.	g2xf3	Rf8-e8
16.	Rf1-d1	Re8-e2
17.	Bc2-f5	g7-g6
18.	Bf5-h3	Nc5-d7
19.	Bc1-e3	Bd6-e5
20.	Qd4-c4	Re2xb2
21.	Ra1-c1	

(Commenting on this game, Botvinnik noted that 21. d5-d6 would have led to victory more quickly.)

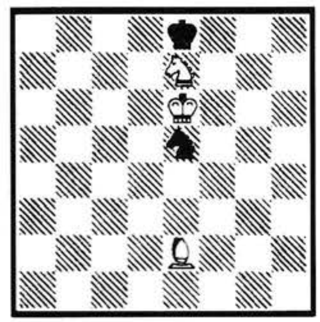
21.	...	Nd7-b6
22.	Qc4-e4	Qd8-d6?

(22. ... Be5-d6 would have been stronger—now it's all over.)

23.	f3-f4	Be5-g7
24.	Be3-c5	Qd6-d8
25.	Bc5-e7	Qd8-e8
26.	d5-d6	Qe8-b5
27.	d6-d7	Nb6xd7
28.	Bh3xd7	

A few moves later, black resigned.

The basic idea of the "symmetry–asymmetry" genre of chess problems is that an externally symmetrical position has an asymmetrical solution. Of course, there can't be absolute vertical symmetry on the chessboard because of the presence of an "extra" file. In fact, paradoxical situations arise precisely because of this file. The following examples are all problems based on this phenomenon.



G. Adamson, 1924. To win.

1. Ne7-d5! Ne5-d7

(Alternatives: 1. ... Ne5-f7 or Ne5-g6 2. Be2-h5; 1. ... Ne5-c6 2. Be2-b5.)

2. Ke6-d6!

(But not 2. Be2-b5? Ke8-d8! 3. Bb5xd7—stalemate.)

2. ... Ke8-d8!

(2. ... Nd7-b8 won't save black: 3. Be2-b5+ Ke8-d8 4. Nd5-b6 Nb8-c6! 5. Kd6xc6!; nor will 2. ... Nd7-f8: 3. Be2-h5+ Ke8-d8 4. Nd5-f6! Kd8-c8 5. Kd6-e7.)

3. Be2-d3

Here are some possible ways of working this position out: 3. ... Kd8-c8 4. Nd5-e7+ Kc8-d8 5. Ne7-c6+ Kd8-e8 6. Bd3-g6+; 3. ... Kd8-e8 4. Nd5-c7+ Ke8-d8 5. Nc7-e6+ Kd8-c8 6. Bd3-a6+; 3. ... Nd7-b8 4. Nd5-b6 Kd8-e8 5. Kd6-c7; 3. ... Nd7-f8 4. Nd5-f6 Kd8-c8 5. Kd6-e7! and the knight is caught. Why wouldn't 1. Ne7-f5 work? Let's see how it plays out: 1. ... Ne5-f7! 2. Ke6-f6 Ke8-f8 3. Be2-f3 Kf8-e8 4. Nf5-g7+ Ke8-f8 5. Ng7-e6+ Kf8-g8! As it turns out, the bishop needs one more file to check the king.

Let's look at another problem.

A. Seleznev, 1917—white king on e5, bishop on a8, and pawn on e6; black king on g4, knight on f4, and pawns on b4 and h4. To draw.

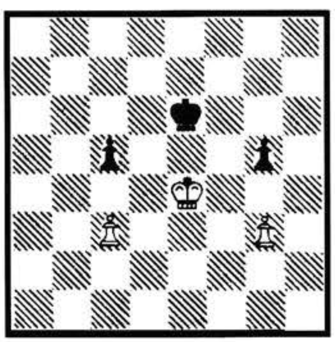
"Where's the symmetry here?" you're saying. It turns up after three moves:

1. Ba8-e4! Nf4xe6
2. Be4-f5+ Kg4-f3
3. Bf5xe6 Kf3-e3!

It seems white won't be able to save itself (one of the pawns will march on to become a queen). And yet...

4. Ke5-d6! Ke3-d4
5. Kd6-c6 Kd4-c3
6. Kc6-d5! b4-b3
7. Kd5-e4 b3-b2
8. Be6-a2 h4-h3
9. Ke4-f3

Both pawns are held back. If white's fourth move had been Ke5-f6?, the white bishop would not have found a position at the right corresponding to a2. It appears that this is the first chess problem in history with "symmetry–asymmetry."



Y. Kneppel, 1967. To win.

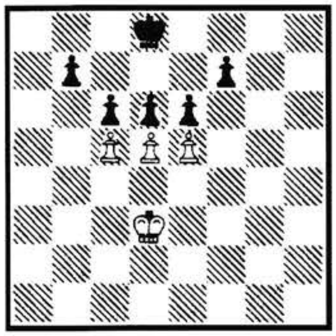
At first glance, it's not clear what significance the extreme left file has in this pawn problem. After all, the pawn isn't in a position to land on it.

1. c3-c4! Ke6-f6

(Alternatively: 1. ... Ke6-d6 2. Ke4-f5 g5-g4 3. Kf5xg4 Kd6-e5 4. Kg4-h5 Ke5-d4 5. g3-g4 Kd4xc4 6. g4-g5 Kc4-d3 7. g5-g6 c5-c4 8. g6-g7 c4-c3 9. g7-g8Q c3-c2 10. Qg8-g5.)

2. Ke4-d5 Kf6-f5
3. Kd5xc5 Kf5-g4
4. Kc5-d4 Kg4xg3
5. c4-c5 g5-g4
6. c5-c6 Kg3-f2
7. c6-c7 g4-g3
8. c7-c8Q g3-g2

The queen can easily manage the black pawn. What if the g-pawn is advanced first? 1. g3-g4? Ke6-d6! 2. Ke4-f5 (2. c3-c4 Kd6-e6) 2. ... Kd6-d5 3. Kf5xg5 Kd5-c4 4. Kg5-f4 Kc4xc3 5. g4-g5 c5-c4 6. g5-g6 Kc3-b2 7. g6-g7 c4-c3 8. g7-g8Q c3-c2. The white pawn has allowed black to get a draw (the stalemate idea)!



O. Riikhimaa, 1942. To draw.

In this example, the h-file will be used in a completely different way.

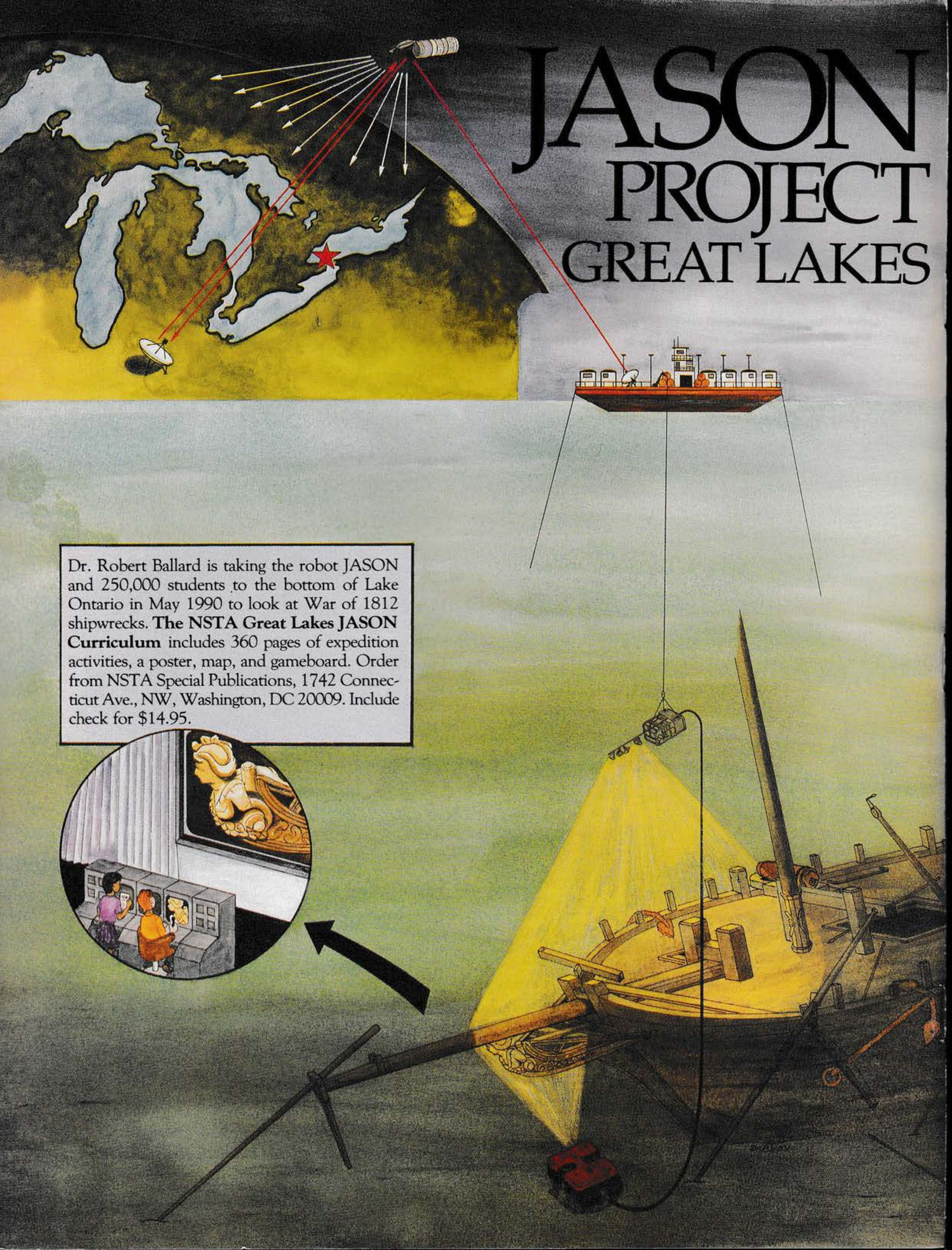
1. d5xe6! d6xe5

(After 1. ... d6xc5 2. e6xf7 Kd8-e7 3. e5-e6, the moves by black ultimately lead only to a stalemate.)

2. e6xf7 Kd8-e7
3. Kd3-e4 Ke7xf7
4. Ke4xe5 Kf7-e7
5. Ke5-f5! Ke7-d7
6. Kf5-e5 Kd7-c7
7. Ke5-d4 Kc7-b8
8. Kd4-c4 Kb8-a7
9. Kc4-b4 Ka7-a6
10. Kb4-a4

Draw.

Opening to the other side won't work, though: 1. d5xc6? d6xc5! 2. c6xb7 Kd8-c7 3. Kd3-c4 Kc7xb7 4. Kc4xc5 Kb7-c7 5. Kc5-b5 Kc7-d7 6. Kb5-c5 Kd7-e7 7. Kc4-d4 Ke7-f8 8. Kd5-e4 Kf8-g7 9. Ke4-f4 Kg7-h6! Black has now gained an advantage thanks to the maneuver on the "extra" file on the far right.


JASON
PROJECT
GREAT LAKES
Dr. Robert Ballard is taking the robot JASON and 250,000 students to the bottom of Lake Ontario in May 1990 to look at War of 1812 shipwrecks. **The NSTA Great Lakes JASON Curriculum** includes 360 pages of expedition activities, a poster, map, and gameboard. Order from NSTA Special Publications, 1742 Connecticut Ave., NW, Washington, DC 20009. Include check for $14.95.