QUANTUM

JULY/AUGUST 1994

$5.00

SPRINGER INTERNATIONAL

The Skater (Portrait of William Grant) (1782) by Gilbert Stuart

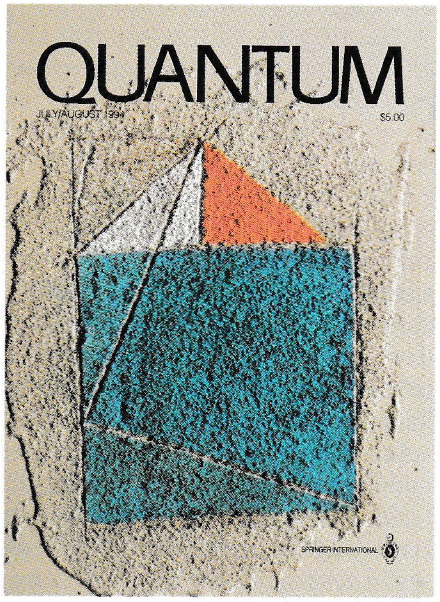
THIS GENTLEMAN CUTS A FINE FIGURE IN MORE ways than one. He also had the good fortune to be portrayed by the foremost American portrait painter of his day, Gilbert Stuart (1755–1828). Stuart painted the likeness of George Washington on numerous occasions, but he is perhaps best remembered for the painting he didn't finish. It was commissioned by Martha Washington in the final years of her husband's life and is familiar to all Americans as the image on the US one-dollar bill.

Like the portraits of Washington, this painting exudes strength and confidence. Mr. Grant manages to maintain a perfectly erect posture even though, from his track on the ice, he is executing a rather tight curve. His skating appears effortless, and he would be the last person to worry about what's happening beneath his blades. He knows it works, and that's enough for him. Let someone else read "The Friction and Pressure of Skating" on page 25. He'd rather just skate.

QUANTUM

JULY/AUGUST 1994

VOLUME 4, NUMBER 6



Cover art by Yury Vashchenko

When Pythagoras discovered his famous theorem, he was so overjoyed he sacrificed an ox to the gods—so the story goes. Tinkering with triangles and the theorem, another Pythagorean made a startling discovery. If the two arms of a right triangle are one unit long, the hypotenuse is $\sqrt{2}$. Now, 2 isn't a perfect square, so its square root isn't a whole number or a fraction made of whole numbers. It's *irrational*. Pythagoras was so pleased with the discovery he sacrificed another ox—so they say.

Morris Kline notes that the second story is suspect on two counts. First, "if all the legends telling of Pythagoras sacrificing an ox were true, he could not have had time for mathematics." More importantly, irrational numbers shook his philosophy to its core. The Pythagoreans believed that everything in nature can be reduced to whole numbers or their ratios. So another legend has the ring of truth: sailing to an untold destination, they tossed the hapless discoverer overboard.

Reprising its appearance in the January/February 1994 issue, the Pythagorean theorem pops up in "Suggestive Tilings," which begins on page 36.

FEATURES

DEPARTMENTS

PUBLISHER'S PAGE

Anticipating future things

Some thoughts on science education in 2044

AS WE CELEBRATE THE 50th Anniversary of the National Science Teachers Association, our thoughts naturally turn to the future. What will science education be like 50 years from now? What will the chronicle of NSTA's first *hundred* years have to say about the period 1995–2044? While it may be folly to prognosticate, we cannot function as educators without a vision of the future. What happens in classrooms fifty years hence will depend on what we do today. So this is a good opportunity to restate our goals for science education, and to take stock of science and society, as we enter the 21st century.

I believe that science itself—its basic concepts, principles, empirical laws, and fundamental theories—will change very little in the next fifty years. Scientific advances are slow compared to the pace of technological innovation. Science educators will take a great step forward when they learn to distinguish between science and technology, yet are able to use technology to teach science. The confusion of the two has long been with us. In 1883 Henry Augustus Rowland, addressing the American Association for the Advancement of Science, said "the proper course of one in my position is to consider what must be done to create a science of physics in the country, rather than to call telegrams, electric lights, and such conveniences by the name of science." While interactive CD-ROM and powerful computers (not to mention electric lights) are not science per se, they can be very useful in the classroom, freeing up the teacher to teach and motivating students to learn science based on relevance.

> "True, there remain among us some means of divination by the stars, by spirits, bodily traits, dreams, and the like—a notable example of the frenzied curiosity of our nature, which wastes its time anticipating future things, as if it did not have enough to do digesting the present."—Michel de Montaigne, "Of Prognostications"

What are the most important things young people in the 21st century will learn from science? The same things they should be learning today (only more so). First and foremost, they will still need to learn facts, names, definitions, concepts, empirical laws, theories, models, and universal laws of science, and to keep them from turning into one another (the "Law of Relativity," the "Theory of Universal Gravitation"). They will learn about countless applications of scientific principles to meet human needs or solve societal problems, without losing sight of the basic science. Second, they will understand that what makes a certain kind of activity *science* is its ability to predict. It is not enough to explain what happened—science must be able to say what *will* happen. Students will learn that in a true science, someone's explanation of a phenomenon is acceptable only if independent investigators can verify empirical facts or reproduce experimental results. Third, students will learn how to learn. In school they will discover reasons to keep educating themselves about science for the rest of their lives, and they will have picked up the intellectual tools required to do it. Fourth, they will develop a keen sense of skepticism. They will carefully examine statements from "authorities," whether they are the world's leading scientists, politicians, or clerics. They will draw on their own reasoning ability and scientific training to study issues and come to their own conclusions.

How will students acquire these abilities? Not by burrowing in a textbook. Not by listening to a teacher. Not by watching a video, no matter how well produced, in which every step in the development of a particular concept is explained with the greatest clarity. Not even by interacting with a fancy computer system that provides continuous feedback and monitors the student's progress. They will acquire these abilities by arduous concentration and hard work motivated by their desire to learn some basic aspect of science for reasons that are entirely their own. They must be guided to the science that *they* find essential, some piece of a puzzle of their own devising.

These puzzles and problems will be so varied that no classroom situation or "cooperative group" with a general set of "typical" problems will be appropriate. I foresee that learning science will become highly individualized. The social or group aspect of science education will involve communicating one's insights and helping one another in the individual struggle to make such discoveries.

So, where does the teacher fit in? The teacher will be a facilitator, a source of guidance. The teacher will not grade the student. There will be better means of assessing a student's grasp of science, and the student–teacher relationship will be the healthier for it. Students will know that the *only* thing the teacher can do is help them learn science, not judge their worth by assigning a grade.

If we do our job well, here in the late 20th century, one major impediment to learning will have been eliminated: the destruction of self-esteem in young people who are humiliated in the classroom for failing to grasp a concept or recall a fact. Perhaps the most damaging thing a teacher can do is to tell students, explicitly or implicitly, that they cannot learn science—that they lack the inherent ability. Unfortunately, this damaging impression is conveyed regularly to young people by well-meaning parents, teachers, counselors, and others in our society.

In the same address to the AAAS cited above, Henry Augustus Rowland said, "American science is a thing of the future"—meaning it was yet to be formed. But I would repeat his words almost verbatim: American science education is a thing of the future. But our only avenue is through the present. I believe we are living on the cusp between two worlds. One is a world of strife, chaos, and misery, fomented by ignorance and superstition. The other is a world of enlightenment, peace, and prosperity, based on rational discourse and universal moral principals. Education for all—especially science education—is the only thing that will keep us from falling into that abyss.

—Bill G. Aldridge

QUANTUM

THE MAGAZINE OF MATH AND SCIENCE

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

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

Quantum (ISSN 1048-8820) is published bimonthly by the National Science Teachers Association in cooperation with Springer-Verlag New York Inc. Volume 4 (6 issues) will be published in 1993–1994. *Quantum* contains authorized English-language translations from *Kvant*, a physics and mathematics magazine published by Quantum Bureau of the Russian Academy of Sciences, as well as original material in English. **Editorial offices:** NSTA, 1840 Wilson Boulevard, Arlington, VA 22201, telephone (703) 243-7100. **Production offices:** Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010-7858. Supervising Production Editor: Madeline R. Kraner.

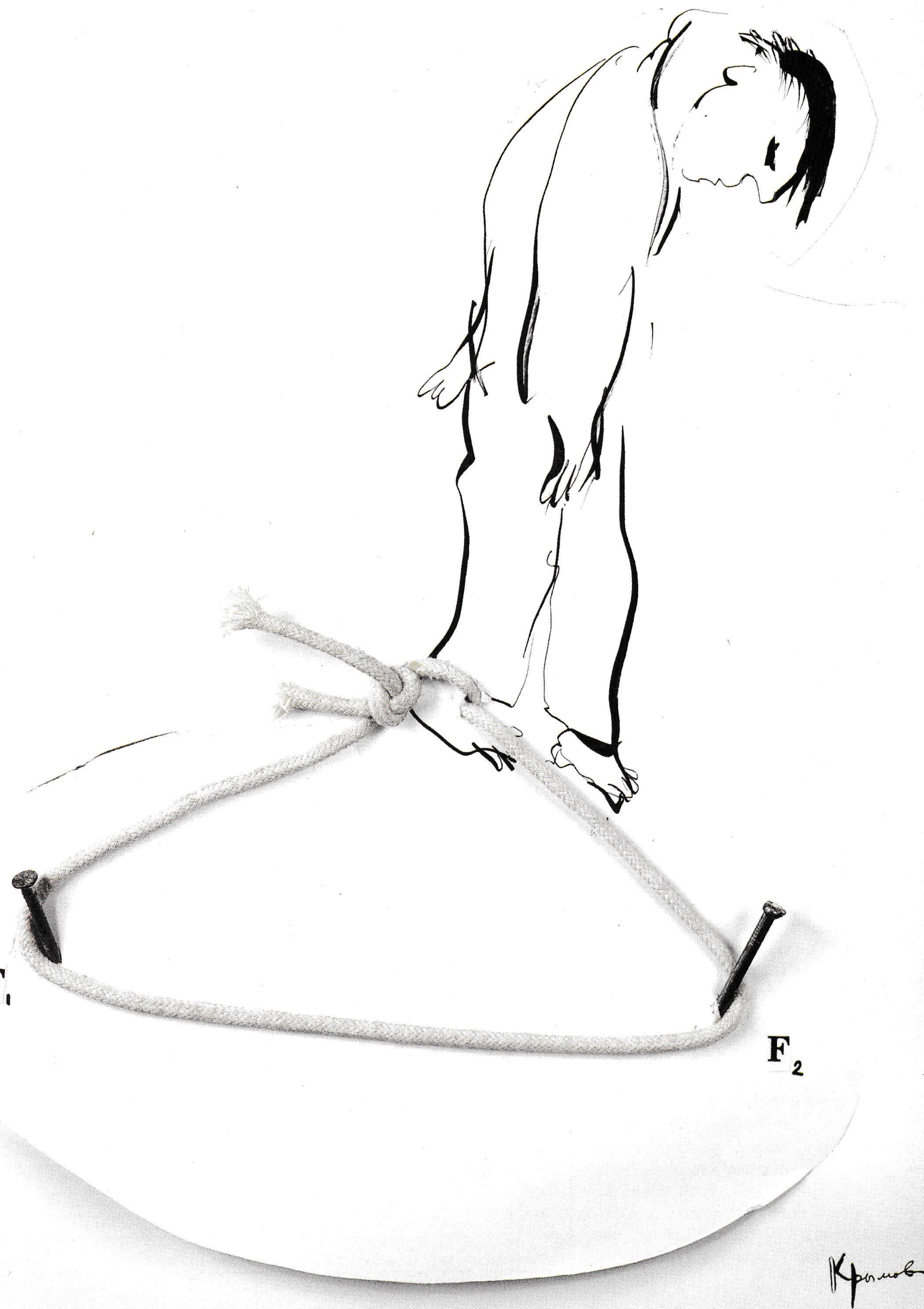
Second class postage paid at New York, NY, and additional mailing ofices. **Postmaster:** send address changes to: *Quantum* Springer-Verlag New York, Inc., Journal Fulfillment Services Department, P. O. Box 2485, Secaucus, NJ 07096-2485.

Subscription Information: North America

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

All Countries Outside North America

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


F_1
F_2
Крылова

Mathematics in perpetual motion

It really does keep going, and going, and going . . .

by Anatoly Savin

THE WORDS "PERPETUAL motion" (*perpetuum mobile* in Latin) are usually associated with a machine like the one you see in figure 1. According to the inventor's conception, this wheel with several balls rolling inside must not only rotate for an infinitely long time, it can also set in motion other machines—looms, lathes, and so on. By the time we graduate from high school we all know that such a machine can't work, because it would contradict the law of conservation of energy. Yet year after year academies of sciences the world over receive hundreds of new (and not new) designs for such machines. Physicists call them "perpetual engines of the first sort." So—are there "perpetual engines of the second sort"? Yes, there are. And that's what I'm going to discuss here.

Here's an idea for such an engine. Suppose we have two equally heated bodies. As it works, the engine transfers some heat from one body to the other without expenditure of energy, and then obtains kinetic energy by means of a heat engine that uses the temperature difference thus created. The work of the heat engine leads to a leveling of the temperatures of the two bodies, so the entire process can be started again: heat is transferred from one body to the other, work is extracted from the temperature drop—and so on to infinity.

A perpetual engine of the second sort doesn't contradict the law of conservation of energy. In this case, energy neither appears nor disappears—it merely passes from one body to the other and then is spent in the heat engine to perform some work. In so doing, the temperature of both bodies becomes lower than it was initially, but this loss can be compensated by the heat of the surroundings.

However, there is another law that prevents the construction of a perpetual engine of the second sort. The idea of this law first appeared in works by the prominent French physicist and engineer Sadi Carnot, a son of the outstanding figure in the French Revolution and well-known mathematician Lazare Carnot. Later it was developed by the English scientist William Thomson and the German physicist Rudolph Clausius. This law is called the second law of thermodynamics and reads as follows: *It is impossible to transfer heat from one body to another without expenditure of energy if the temperature of the first body is no higher than that of the second.*

This prohibition doesn't look all that convincing. Does it always work? The great English physicist James Clerk Maxwell conceived of a device, called Maxwell's demon, intended to refute this law. Imagine a box divided in half with a partition that has a small hole in it and a "demon" sitting near the hole. Fill the box with any gas and ask the demon to let only fast molecules from the left half of the box into the right half and leave all slow molecules in the left half. Since the temperature of a gas is characterized by the average

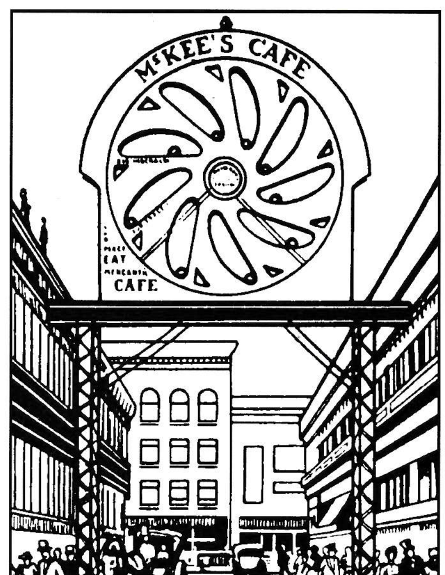


Figure 1

Art by Dmitry Krymov

molecular speed, the gas in the left half of the box will be cooled and the gas in the right half will be heated. A number of devices have been proposed that, according to their authors, could play the role of such a demon, but all of them failed to work because of certain unanticipated effects found in them.

However, where physics fails, mathematics may succeed. I want to suggest a device, free from demons and other evil spirits, that can be built even in a high school workshop. I'll present calculations that will clearly show that this device can transfer heat from one body to another if they were heated equally at the outset.

Before going into a description of my "perpetual engine," you'll have to swallow a certain amount of math so that you can convince yourself that the reasoning and constructions to follow are correct. We'll be dealing with the ellipse and its properties.

By definition, *an ellipse is the curve formed by all the points in the plane such that the sum of their distances to two fixed points F_1 and F_2 is constant.* Each of the points F_1 and F_2 is called a *focus*; the constant sum of the distances is usually denoted by $2a$.

This property is used by gardeners when they want to make oval flower beds. They drive a pair of sticks into the ground (at the foci), tie the ends of a rope to them, then take another stick with a sharp end and stretch the rope taut with this stick. If the stick is moved so that the rope remains taut, its sharp end traces an ellipse. The size and shape of the ellipse depend on the length of the rope and the distance between the foci. You can verify this on a sheet of paper using two pins and a pencil instead of sticks (fig. 2).

After you draw several ellipses you'll see that any ellipse is a closed convex curve that has a center of symmetry and two symmetry axes—the line F_1F_2 and the perpendicular bisector of the segment F_1F_2. Also, it's easy to see that the sum of the distances to the foci is less than $2a$ for points inside the ellipse and greater than $2a$ outside it.

This information suffices to prove an important and not so obvious property of the ellipse: the segments that join the foci of an ellipse to an arbitrary point M on it make equal angles with the line that touches the ellipse at M.

Comparing this property to the law of reflection of light—the angle of incidence equals the angle of reflection—we come up with the following formulation: a ray of light issuing from a focus of an ellipse after reflection from it hits the other focus.

This is the "optical property" of the ellipse. It can be observed in nature: there are caves with ellipsoidal domes where you can find two spots, far enough from each other, such that the voice of a person standing at one of these spots is heard at the other spot as if the speaker were just inches away. And some palaces and castles have halls intentionally designed to produce this effect.

Since the optical property of the ellipse plays a major role in what follows, I'll give its proof—it's rather short and simple.

Let l be the tangent to an ellipse at its point M (fig. 3) and let α and β be the angles between the line l and segments F_1M and MF_2 (F_1 and F_2 are the foci). Reflect F_1 about l into F_1', join $F_1'F_2$, and find the intersection point N of this segment with l. Suppose $N \neq M$; then N lies outside the ellipse (actually, the entire line l, except its point M, lies outside the ellipse, because any ellipse is a convex curve). Therefore $F_1N + NF_2 > 2a = F_1M + MF_2$. But by the construction, $F_1N = F_1'N$ and $F_1M = F_1'M$, so we get $F_1'N + NF_2 > F_1'M + MF_2$, or $F_1'M + MF_2 < F_1'F_2$, which contradicts the Triangle Inequality for triangle $F_1'F_2M$. Therefore, points N and M must coincide, so F_2MF_1' is a straight line—that is, $\alpha = \beta$.

Now let's bring our project to fruition. Take a sheet of good drawing paper, mark points F_1 and F_2 on it, and draw two ellipses with the foci at these points, using a longer string the first time and a shorter one the second time. Draw the perpendicular bisector to the segment F_1F_2 and erase a part of what we've drawn so as to obtain a "mushroom" like the one in figure 4.

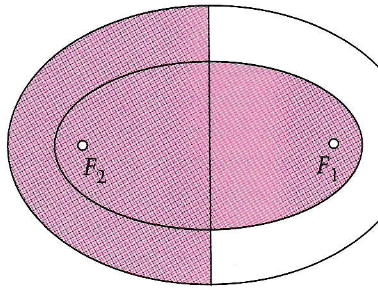


Figure 4

Now roll your "blueprint" up, put it in a cardboard tube, and go to the nearest metalworking shop. Ask that a tin shell be made in the shape obtained by rotating the curve in the blueprint about its symmetry axis F_1F_2 (fig. 5). The inside of the shell must be covered with a reflective coating. When the thing is ready, take it home. Now you are the

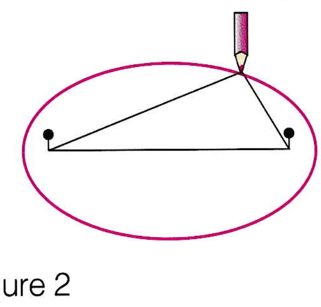


Figure 2

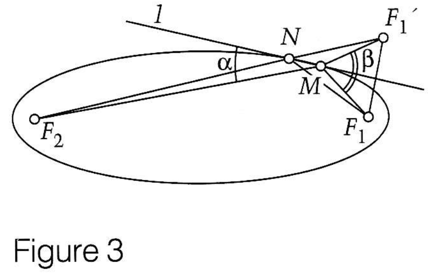


Figure 3

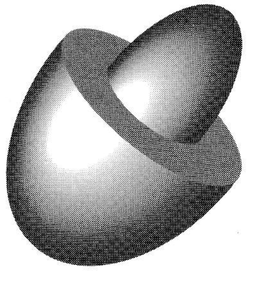


Figure 5

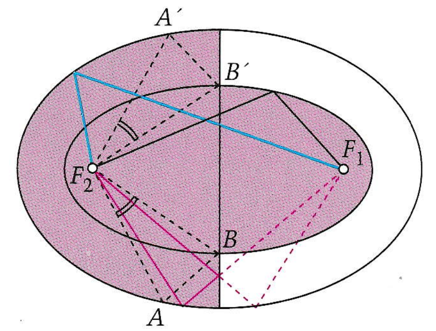


Figure 6

owner of a perpetual engine! You don't believe me? Well, I'll prove it mathematically.

Place two equally heated bodies at points F_1 and F_2. Notice that any ray (of light or heat) issuing from F_1 must arrive at F_2 (fig. 6). A ray propagated along the same trajectory from F_2 will arrive back at F_1, of course. But look at a ray aimed from F_2 at the vertical barrier. If there were no partition, this ray would rebound from the erased part of the ellipse and hit F_1. But now it will be reflected at some point R of the partition in the direction it would have had if it had propagated from F_1 to R (fig. 6). Therefore, it will come back to F_2! And, clearly, leaving point F_2 along the same trajectory in the opposite direction, we again come back to F_2. So there's a considerable number of rays that make "light loops." In the cross section of our "engine" shown in figure 6, these are all the rays within the angle AF_2B and the symmetric angle $A'F_2B'$.

So, the body at F_2 will be heated, while the body at F_1 will be cooled. We can even compute the eventual temperatures of the bodies after the process stabilizes.

Suppose the initial temperatures were T_0 K (or "kelvins," named after the English physicist W. Thomson mentioned above, on whom the title of Lord Kelvin was conferred for his outstanding scientific achievements). It's known that the rate of energy loss by radiation (the luminosity) is proportional to the fourth power of the temperature (in kelvins) of the radiating body. Therefore, if the temperature at

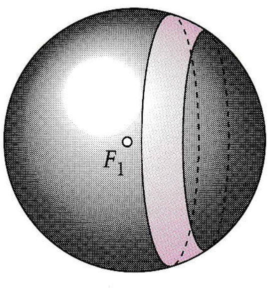


Figure 7

point F_1 is T_1, and at F_2 it's T_2, then the luminosity at F_1 and F_2 can be written as kT_1^4 and kT_2^4, respectively. All the rays from F_1 after one reflection arrive at F_2, but only some of the rays from F_2 arrive at F_1. How many? Consider a small sphere centered at F_2. The rays that return to F_2 cut a ring out of this sphere (fig. 7). If its area is A_1 and the total area of the sphere is A, than the amount of radiation that comes back to F_2 equals $A_1kT_2^4/A$, and the portion that arrives at F_1 is $(A - A_1)kT_2^4/A$.

In the steady state, the energy radiated from point F_1 equals the energy that arrives at this point in the same time interval. That is,

$$\left(A - A_1\right)\frac{kT_2^4}{A} = kT_1^4.$$

On the other hand, by the law of conservation of energy, the loss of energy at F_1 equals the gain of energy at F_2—that is, $C(T_1 - T_0) = C(T_0 - T_2)$, where C is the heat capacity of either body (we assume that the bodies are absolutely alike). The last equation gives $T_1 + T_2 = 2T_0$, so using the previous relationship and letting $b = [(A - A_1)/A]^{1/4}$, and noting that $b < 1$, we finally get

$$T_1 = \frac{2T_0 b}{1+b} < T_0,$$

$$T_2 = \frac{2T_0}{1+b} > T_0.$$

That's it! You can check the reasoning from the very beginning and be satisfied that it's perfectly correct. So, did we really refute the second law of thermodynamics? Unfortunately (or fortunately), no. We've made a mistake. But where? Think about it yourself, and then compare your answer with what's written below. I have a feeling that not every reader will be able to find the correct answer.

You were on the wrong track if you tried to find an error in the physics—for instance, in the fact that we ignored convection. Why? We can create a vacuum inside the shell. But what else is there in our little "engine" except math and physics? The error is hiding at the border, so to speak, between these two sciences—in the transition from the physical process to its mathematical model.

Recall that we dealt with two bodies placed at the foci. Thus, we neglected the sizes of the bodies. This is common in physics reasoning, and the phrase "a body is located at point M" doesn't cause anyone to protest. In many situations, this disregard of the sizes of the bodies is justifiable. For instance the motion of a body under given forces applied to its center of mass doesn't depend on its size and shape. Or, as Newton showed, a body consisting of a number of concentric homogeneous spherical layers attracts other bodies as the same mass concentrated at the center of this body. So to replace a body with a point is indeed a customary operation in physics.

But in our case this operation leads to an error. Let's see whether our reasoning remains valid if we consider balls of nonzero radius centered at the foci.

Consider three rays reflected from the ellipse at the same point M in its left half (fig. 8): a ray issuing from the foci F_2 and two rays, AM and BM, with endpoints at a distance r from F_2.

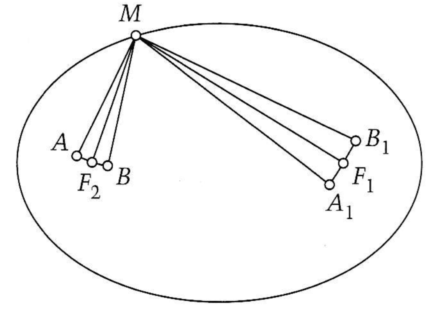


Figure 8

After reflection, the first ray will pass through the focus F_1, and the two other rays at the same distance R from F_1 (because $\angle AMF_2 = \angle F_2MB = \angle B_1MF_1 = \angle F_1MA_1$). By the similarity of right triangles AMF_2 and A_1MF_1, $r/R = AF_2/A_1F_1 = F_2M/F_1M$. Since M lies in the left part of the ellipse, $F_2M < F_1M$, and so $R > r$. This means that some of the rays from a ball of radius r centered at F_2 will neither come back to this ball after several reflections nor hit an equal ball at F_1—these rays will disperse. And this destroys all our reasoning and constructions.

Physics lovers who have read the article to this point can now rest assured that the second law of thermodynamics remains inviolable. And for math lovers, here are a few more curious facts about the behavior of rays reflected in an elliptical mirror.

If a ray emerges from a focus of an ellipse, then after the first reflection it passes through the second focus; if there is no body there to restrain it, it reflects again to return to the first focus, then arrives once again at the second focus; and so on. That's clear enough. But it's less clear though curious that with each reflection the trajectory approaches the line through the foci, and in the limit merges with the segment of this line inside the ellipse. If the first segment of the trajectory intersects with this segment not in a focus, it will never pass through a focus later. Not only that, if the first intersection occurs outside the segment between the foci, the same will be true for all subsequent intersections as well (fig. 9). And most surprising is the fact that there exists a smaller ellipse with the same foci such that the trajectory touches it after each reflection.

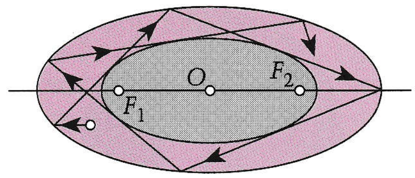


Figure 9

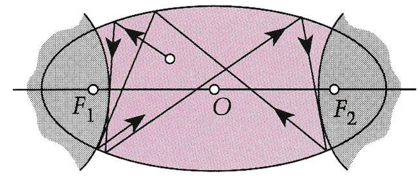


Figure 10

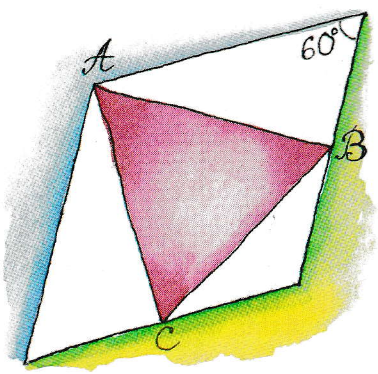
Similarly, a trajectory intersecting the segment between the foci (fig. 10) will do so after each reflection, and its segments will all touch a hyperbola with the same foci. (By definition, this is a curve formed by points such that the differences of their distances to the foci equal a given constant.)

These and other properties of reflections in a curve or curved surface are studied in a comparatively new branch of modern mathematics called the theory of mathematical billiards. As to the physical side of the matter, M. I. Feingold is at the blackboard in this issue (page 40), showing the effects of reflection in a parabolic mirror. ◘

Just for the fun of it!

B116

Band on parade. A huge military band was playing and marching in formation on the parade grounds. First the musicians formed a square, then they regrouped into a rectangle so that the number of rows increased by 5. How many musicians were there in the band? (S. Dvoryaninov)

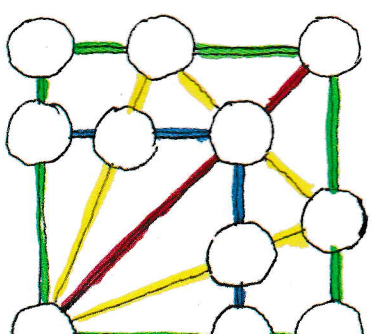


B117

Sixty degrees everywhere. A triangle ABC is inscribed in a rhombus with an acute angle of 60° as shown in the figure. One of the triangle's angles is also 60°. Prove that ABC is an equilateral triangle. (V. Proizvolov)

B118

Logic meets physics. There are three lamps in one room and three switches in another. Each switch is connected to its own lamp. How can you determine which lamp is connected to each switch if you're allowed to enter the room with the lamps only once? (A. Zilberman)



B119

Magic network. Write the numbers 1, 2, 3, ..., 11 in the circles in the figure such that the sum of the three numbers along each of the ten segments is the same. (M. Varga)

B120

History meets arithmetic. The height of a certain Egyptian pyramid in meters is greater than the product of two odd two-digit numbers, but less than the square of their half-sum. For which of the pharaohs was this pyramid built? (You'll need to refer to an encyclopedia for the heights of various pyramids.) (I. Akulich)

Art by Pavel Chernusky

ANSWERS, HINTS & SOLUTIONS ON PAGE 52

Mushrooms and X-ray astronomy

"Two things affect us most deeply: the stars above and the conscience within."—Folk wisdom

by Alexander Mitrofanov

THIS ARTICLE IS DEVOTED TO a modest but still glorious event. Thirty years ago—or, to be exact, on June 18, 1962—the first nonsolar source of X rays was discovered. This source is in the constellation of Scorpio, and according to the accepted terminology is now called Sco X-1.

The discovery was made quite unexpectedly during a rocket experiment by the American scientists Bruno Rossi and George Clark of the Massachusetts Institute of Technology and Ricardo Giacconi, F. Paolini, and Herb Gursky of American Science and Engineering, Inc. The experimenters had planned to do research in the X-ray band of the spectrum (in the region 0.2 nm $< \lambda <$ 0.8 nm) on the fluorescence of the lunar surface induced by the flow of fast particles coming from the hot solar corona (the solar wind). So-called "soft" X rays can't be detected at the Earth's surface because they are absorbed in the atmosphere. So devices were installed onboard the Aerobee-150 rocket, which was capable of lifting scientific equipment to an altitude of 200 km or more.

At that time the existence of detectable celestial X-ray sources other than the Sun and Moon was considered improbable.[1] Indeed, the distances to the stars, even the nearest ones, are so large that the $1/R^2$ decrease in the flux of X rays from a star at a distance R would negate all attempts at detecting X-ray radiation from stars like the Sun and those that are even hotter and bigger.[2]

It was estimated that the X-ray flux from hot stars should not be more than about 10^{-3} photon/cm$^2 \cdot$ s —far too faint to be detected by devices available at that time. Nevertheless, in 1962 two of the three photon detectors in the X-ray experiment showed a sharp increase in the photon counting rate.

During that historic flight the rocket rotated about its longitudinal axis, and the readings of the detectors were correlated with this rotation. Thus, despite the large angular view of the detectors, it was clear that the source of the X rays was located somewhere near the center of our galaxy. The exact direction to the source could not be established in that experiment, although evidently neither the Sun nor the Moon was connected in any way with the increase in X-ray photons. That very first experiment also showed the existence of a cosmic background of comparatively bright, continuous X-ray radiation.

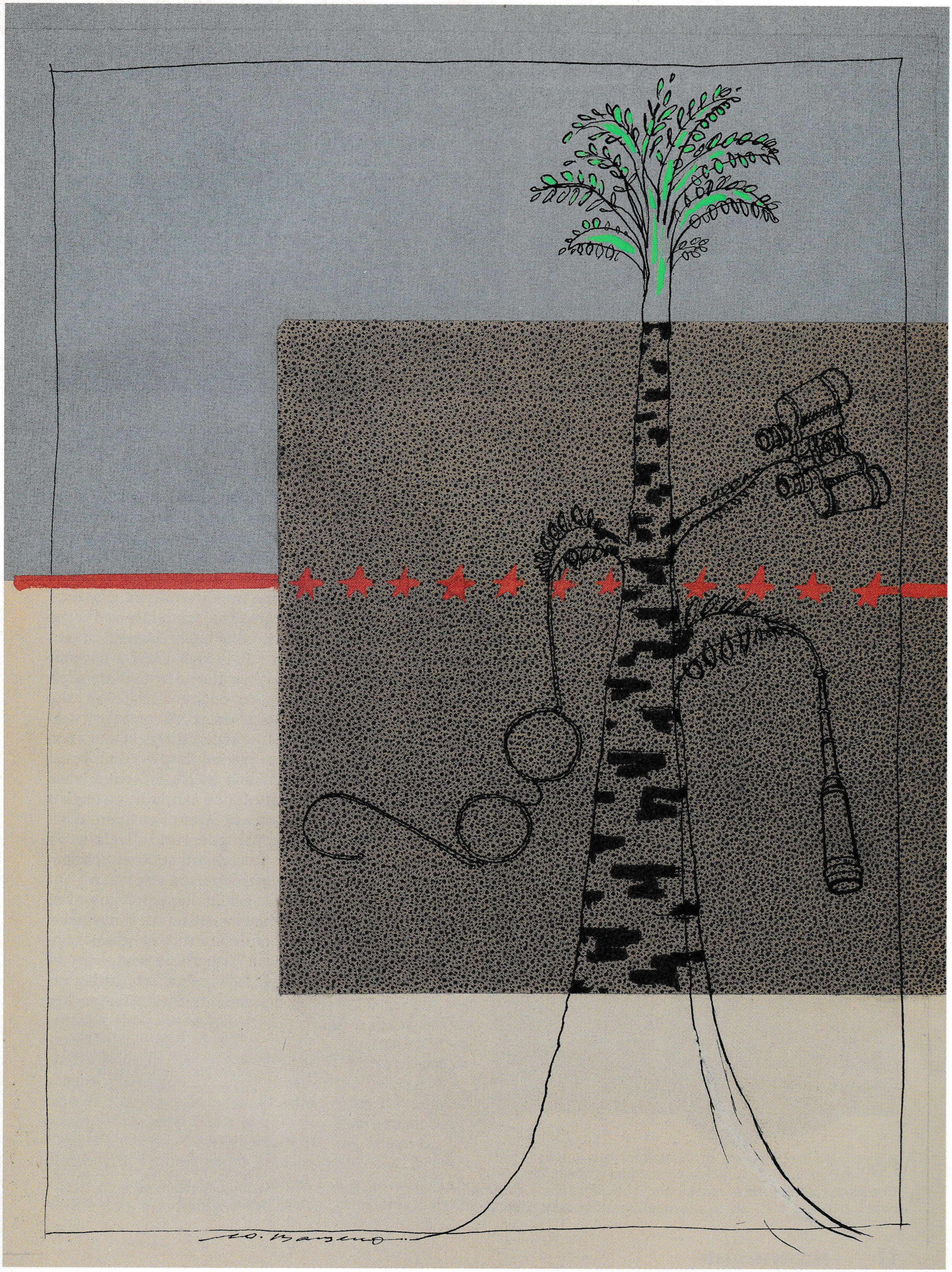
One year later the scientists again launched Aerobee, but this time with new equipment. The angular view of the X-ray detectors was decreased by means of a Roentgen collimator whose walls were impervious to X rays. The same source discovered in 1962 was detected again! But this time its coordinates on the celestial sphere were determined. In another part of the sky a second bright X-ray source was found, in the Crab Nebula (Tau X-1). There could no longer be any doubt: unusually bright X-ray sources exist outside our solar system (see figures 1 and 2 on page 12). This discovery revolutionized our view of the structure of the universe and stimulated the development of a new experimental science: X-ray astronomy. You can read more about this in many fine books on popular astronomy.

As for our story, we leave the sky along with the discovery that transformed astrophysics in our century and return to Earth, to the experimental equipment that makes such discoveries possible. Let's take a closer look at one of the tools of X-ray optics—the collimator, which is a component of modern X-ray or gamma-ray telescopes (figures 3 and 4). Its design is rather simple: it consists of a system of parallel metal plates, masks, slits, and little identical tubes working together to limit the angular

[1]X-ray solar radiation was discovered in the late 1940s by American scientists. In near-Earth orbit the flow of X-ray radiation from the "calm" Sun is about 10^6 photons/cm$^2 \cdot$ s. During solar bursts it increases many times over.

[2]Modern technology is capable of recording the X-ray radiation of the stellar coronas in ordinary stars—for example, Alpha Centauri, which is a star similar to our Sun.

Art by Yury Vashchenko



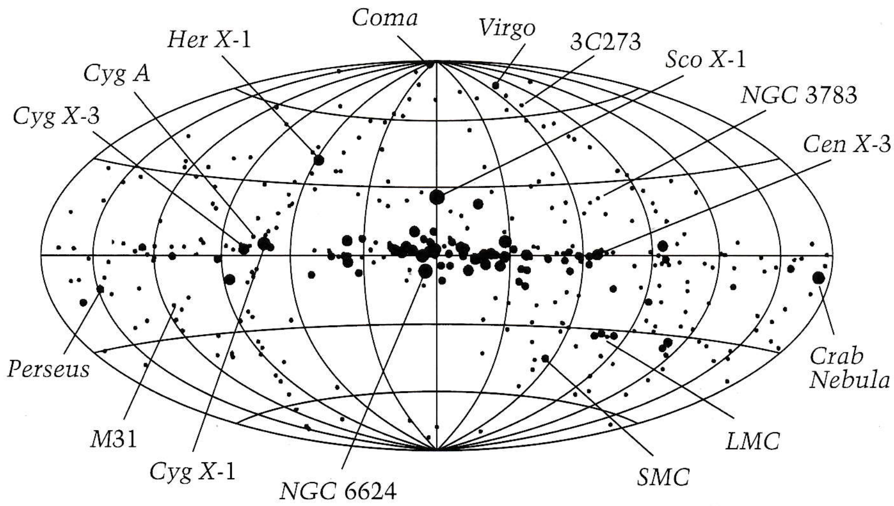


Figure 1

Map of the X-ray sky from data recorded by Uhuru, launched by Ricardo Giacconi's group on December 12, 1970. The size of a circle corresponds to the luminescence of the source. The source Sco X-1 recorded during the rocket flight in 1962 is among the brightest in the sky. It's located in the center of the map, near the galactic equator. This experiment used X-ray photon counters with plate collimators having an angular view of 5 × 5° and 5 × 0.5°.

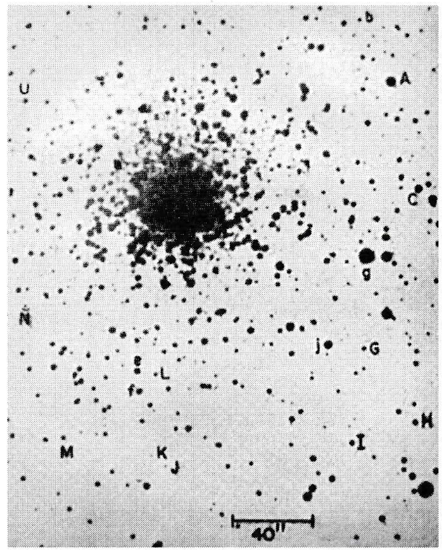


Figure 2

Circular cluster NGC 6624, which contains a bright X-ray source (see the map of the X-ray sky in figure 1).

view of a photon detector. The collimator's walls are usually "black"—that is, opaque for wavelengths in the spectral band under investigation. This ensures that objects located outside its view, as well as the continuous background noise, won't hinder its observation of sources located at a small angle to the telescope's axis.

Now, if you were told that binoculars or other optical devices grow in the forest, you of course wouldn't believe it. Nevertheless, you can find something in the forest that looks very much like an optical device—the X-ray collimator. I'm serious! A model of such a device—one that you could almost use "as is" in an experiment—grows on birches and some other trees. It's a polyporoid (*Fomes fomentarius*) from the family of tree fungi (*Polyporaceae*). You can find polyporoids clinging to dead birches and alder trees. Under birches infected by polyporoids, other mushrooms grow up in due time—specifically, honey agarics, a "secret" known to many mushroom hunters.

In its outward appearance and to the touch, the fruit of a polyporoid resembles natural cork. However, closer inspection shows that this "cork" consists of a huge number of cylindrical pores whose axes are directed perpendicular to the lower horizontal surface of the polyporoid. These tubular pores contain the spores. Each year the polyporoid grows a new porous layer, which is why the mushroom has a layered structure. You can tell how old the mushroom is by the number of projections on the outside. For our purposes it's noteworthy that along almost the entire surface of the polyporoid the pores don't grow randomly during each season, but are a continuation of the preceding years' tubes. Thus, when the spores leave the tubes, one can look through a thick, dense layer of polyporoid because the pores go all the way through. It's an X-ray collimator, don't you think?[3]

Find a large polyporoid in the forest and cut some layers of different thicknesses from it. With these samples you can do several interesting experiments. First, you need to

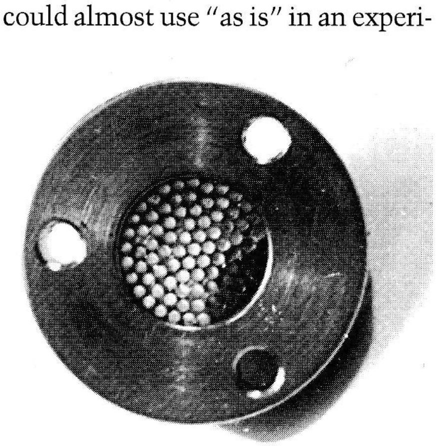


Figure 3

Collimator composed of metal capillary tubes.

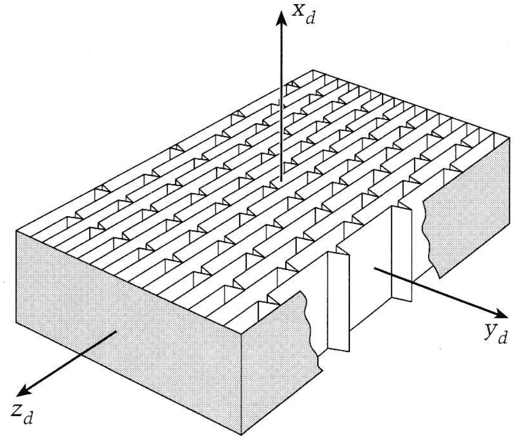


Figure 4

Plate X-ray collimator used by Rossi and his colleagues.

[3]It's worth noting that the empty pores of the polyporoid once served humanity in another capacity. Before matches were invented, polyporoids were used to prepare tinder—a material that would ignite and smolder when a spark landed on it. (The spark was produced by striking steel against flint.)

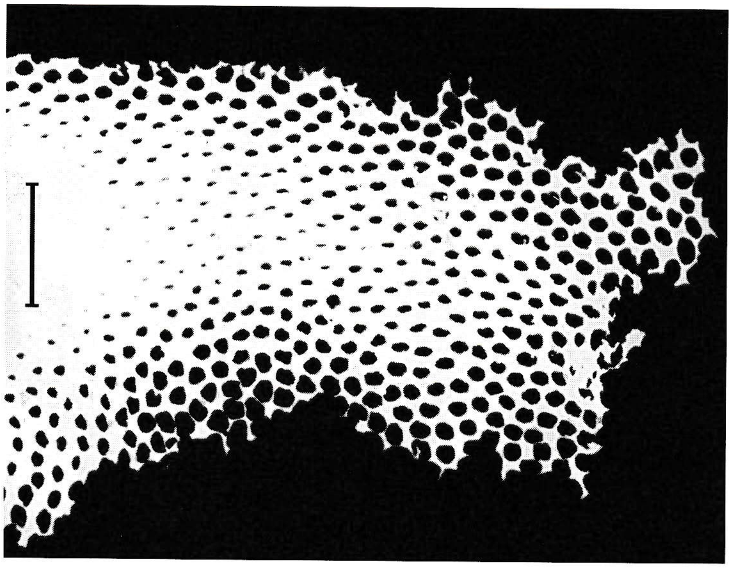


Figure 5
Cross section of a polyporoid layer.

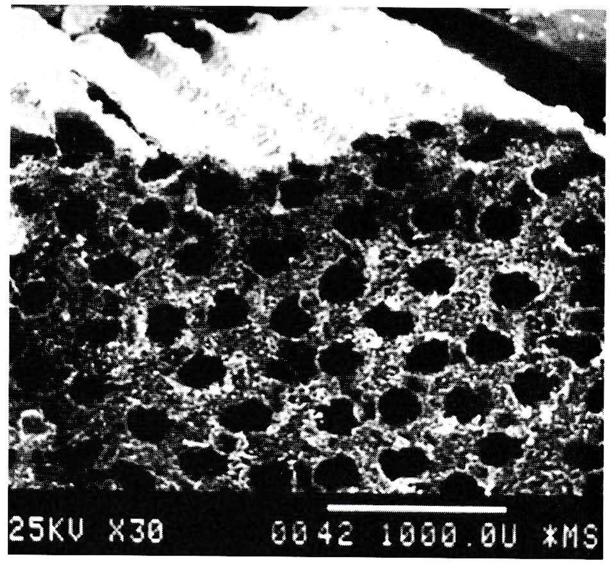


Figure 6
Electron micrograph of the surface of a polyporoid.

dry the layers of polyporoid under a press to prevent them from warping.

Experiment 1. As a first step, let's find the density of the pores—that is, the number of pores per unit area in a layer of a polyporoid. Then we'll try to measure the diameter D of the pores. These tasks are rather easy to do with a microscope, but you can still do them without one. We'll use an ordinary photographic enlarger or film projector. With a sharp knife or razor blade cut off a thin section of the fruit of a polyporoid, which should be oriented perpendicular to the pores—that is, parallel to the bottom surface of the mushroom. Then place the section in the enlarger (instead of a negative), focus the image, and make a magnified positive photograph of the section. However, it would be sufficient just to trace the outline of the pores on an ordinary sheet of paper. All that's left is to calibrate the magnification of the enlarger, count the number of pores in the photograph, and measure the diameter of the pores. By way of example, figure 5 shows the shadow projection of the cross section of a polyporoid with a 2-mm scaling bar. Figure 6 is a picture of the surface layer of a polyporoid obtained with a more complicated device—the scanning electron microscope. The scaling light bar here corresponds to 1 mm. These figures show that the average diameter D of the pores is about 1/3 mm. The deviation is not too large, although, strictly speaking, the pore's shape is far from being a perfect cylinder.

Try to figure out why we used a thin section of the mushroom for our measurements instead of the entire layer, even though the entire layer also lets the light through (the pore tubes go all the way through the mushroom).

Experiment 2. It's interesting to look at objects through a polyporoid layer. Turn on your desk lamp, place a porous layer of a polyporoid in the path of the light, and observe the hot filament through the pores. To do this experiment you need to practice getting the right direction and turning the layer by small angles. The pore-tubes are narrow, which results in a comparatively small sighting angle: $\alpha_{max} \sim D/L << 1$. Even for a layer only 3 mm thick, this is about 10°, and for a layer 3 cm thick, it's 1°. Figure 7 shows the hot filament of a tungsten lamp photographed through the porous layer of a polyporoid.

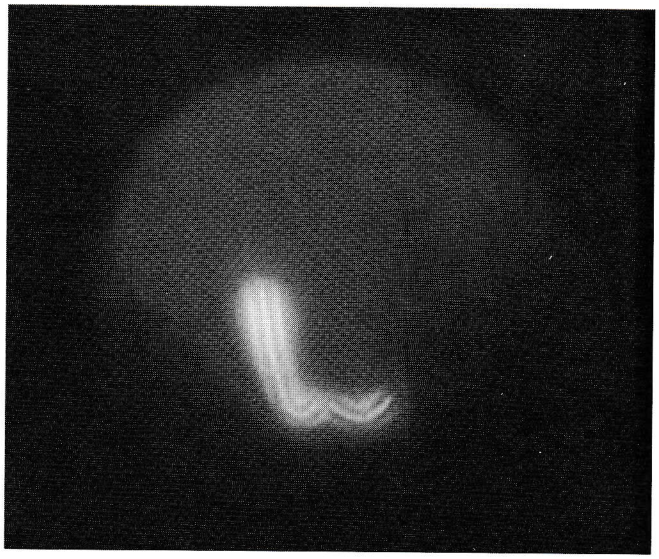


Figure 7
Hot filament of a tungsten lamp photographed through the porous layer of a polyporoid placed in front of the lens. Note that the image is somewhat blurred, which is the result of light diffraction in the pores.

After you do the second experiment, you'll probably conclude that the reflection of visible light from the walls was very small, as if the walls were black. Second, you'll observe some blurring of the filament image connected with the diffraction of light in narrow openings (which is what the pores are). For X-ray radiation the

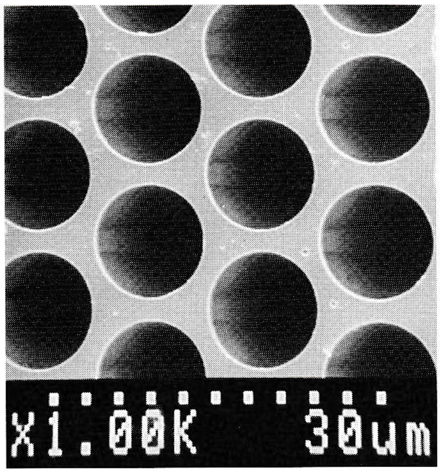


Figure 8

Electron micrograph of a portion of the surface of a glass collimator, based on the so-called microchannel plate. The main use of the microchannel plate in X-ray astronomy is to amplify the image. The system of empty glass capillaries can also be used to focus X rays (the Kumakhov lens).

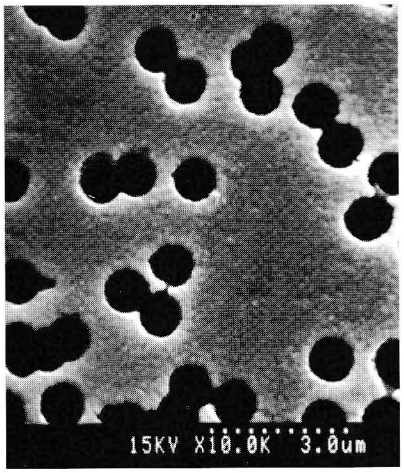


Figure 9

Polymer porous collimator.

role of the diffraction-induced blurring is not as significant (for the range of sizes used in the experiment) because the wavelength of an X-ray photon is far less than that of a photon of visible radiation.

It's also interesting to determine how the optical characteristics of a polyporoid layer depend on the parameters of the pore-tubes—that is, on the diameter D and length L. The optical properties of any collimator (a polyporoid layer included) are characterized first of all by the dependence of the transmittance T on the angle α between the collimator axis and the direction of a parallel beam. You may be tempted to try a third experiment to measure the function $T(\alpha)$, but try to find it theoretically instead. This task will be of particular interest to math aficionados. Here I'll merely give the result: see equation (1) in the box below, where T_0 is the maximum transmittance of the collimator (at $\alpha = 0$). When the angular view of the collimator is small—that is, $\alpha << 1$—this formula can be simplified, taking into account that for small α, $\cos\alpha \cong 1$, $\tan\alpha \cong \alpha$: see equation (2) in the box. Usually the formula for $T(\alpha)$ is written in this form for an X-ray or gamma collimator composed of identical narrow capillaries.

I could end my story here, but I really must say a few words about the technical solutions that humans have devised to collimate X-ray fluxes. Figures 8 and 9 show a glass microcollimator with channels 10 μm in diameter and a thin polymer film collimator for soft X-ray and vacuum ultraviolet radiation with randomly located cylindrical pores approximately 1 μm in diameter (which could be made even smaller). The polymer collimator is only 10–20 μm thick. The pores in the polymer films are obtained by irradiating the material with heavy ions in an accelerator and then processing the irradiated film chemically. These and other artificial porous structures have been used for a long time in electronics and optics, as well as in many technological processes for liquid and gas filtration. These materials, of course, have far outstripped the polyporoid in their technical characteristics and potential uses. But then, these fancy gadgets don't smell like mushrooms, and there are no honey agarics growing underneath . . . ◘

(1) $$T(\alpha) = T_0 \frac{2\cos\alpha}{\pi}\left[\arcsin\sqrt{1-\left(\frac{L}{D}\tan\alpha\right)^2} - \frac{L}{D}\tan\alpha\sqrt{1-\left(\frac{L}{D}\tan\alpha\right)^2}\right],$$

where $$-\arctan\frac{D}{L} \le \alpha \le \arctan\frac{D}{L}$$

(2) $$T(\alpha) = T_0 \frac{2}{\pi}\left[\arccos\left|\frac{\alpha}{\theta}\right| - \left|\frac{\alpha}{\theta}\right|\left(1-\left(\frac{\alpha}{\theta}\right)^2\right)^{1/2}\right],$$

where $$\theta = \frac{D}{L}$$


2

Through the decimal point

Where the quadratic $x^2 = x$ has four roots

by A. B. Zhiglevich and N. N. Petrov[1]

WE WILL TRY TO PERSUADE YOU THAT THE number of roots of the equation $x^2 = x$ is not two, as everybody thinks, but four, as stated in the title of this article. But first we offer the following undisputed fact. If a natural number (in decimal notation) ends in any of the digits 0, 1, 5, or 6, its square ends in the same digit. This also applies to the two-digit endings 00, 01, 25, and 76—that is, if a number ends in one of these pairs of digits, its square does too (for instance, $176^2 = 30{,}976$; $225^2 = 50{,}625$; and so on). A similar property holds for the three-digit endings 000, 001, 625, and 376. At the same time, no other one-, two-, and three- digit endings possess this property. What's so special about these combinations of numbers? The answer is given by the following theorem.

THEOREM. *For any natural k there are exactly four sets of k digits (00...00, 00...01, and two more, ending in 5 and 6, respectively) such that if a natural number ends in one of these sets of digits, the square of this number ends in the same set of digits.*

To eliminate any doubts, we'll give two proofs of this theorem. But first let's try to understand better what we have to prove.

We're looking for an integer x, $0 \le x < 10^k$, such that for any integer $a \ge 0$ the square of the number $10^k a + x$ (which is the general form of a number whose k-digit ending coincides with x) has the form $10^k b + x$, where b is an integer. But $(10^k a + x)^2 = 10^k(10^k a^2 + 2ax) + x^2$, so our condition simply amounts to the fact that $x^2 - x$ is divisible by 10^k. Thus we've arrived at a very simple problem: to find all integers x, $0 \le x < 10^k$, such that $x^2 - x$ is divisible by 10^k. We must make sure that this problem has four solutions, and that these solutions have the properties specified in the theorem.

First proof. The difference $x^2 - x = x(x - 1)$ must be divisible by $10^k = 2^k 5^k$. But the numbers x and $x - 1$ can't both be divisible by 2 and can't both be divisible by 5. So we're left with four possibilities:

(1) x is divisible by 10^k;
(2) $x - 1$ is divisible by 10^k;
(3) x is divisible by 2^k and $x - 1$ is divisible by 5^k;
(4) x is divisible by 5^k and $x - 1$ is divisible by 2^k.

Since $x < 10^k$, the first possibility can occur only if $x = 0$; similarly, the second possibility corresponds to $x = 1$.

In case (3) the number x (and in case (4) the number $x - 1$) is one of the numbers $a \cdot 2^k$, and since $x > x - 1 \ge 5^k$, we know that $0 \le a \le 5^k - 1$. All these numbers yield different remainders when divided by 5^k, because if the numbers $2^k b$ and $2^k c$ give the same remainders upon division by 5^k, then their difference $2^k(c - b)$ is divisible by 5^k, and so by 10^k as well, which is impossible for $0 < c - b < 5^k$. But there are only 5^k possible remainders, so each of them is represented once among our numbers $2^k a$. This means that exactly one of these numbers—say, $2^k a_1$—has the remainder 1, so that for $x = 2^k a_1$, the number $x - 1$ is divisible by 5^k. Similarly, exactly one of these numbers—say, $2^k a_2$—has the remainder $5^k - 1$ when divided by 5^k, and so, for $x - 1 = 2^k a_2$, the number x is divisible by 5^k.

This conclusively shows that in each of our four cases there is one and only one x satisfying our conditions, so the problem has exactly four solutions. It's also clear that in case (3) the number x ends in 6, and in case (4) it ends in 5. The proof is complete.

Second proof. We'll take it for granted that the problem has no more than four solutions. (We know this from the first proof, but it can be proved directly as well.) Also, we'll use induction over k: we'll assume that we've already found the numbers x_{k-1} and y_{k-1} ending in 5 and 6, respectively, such that $0 \le x_{k-1} < 10^{k-1}$, $0 \le y_{k-1} < 10^{k-1}$, and the differences $x_{k-1}^2 - x_{k-1}$ and $y_{k-1}^2 - y_{k-1}$ are divisible by 10^{k-1} (we create a basis for the induction—that is, the existence of x_1 and y_1— by setting $x_1 = 5$,

[1]Some material by V. Denisenko was incorporated in this article with his permission.—*Ed.*

Art by Pavel Chernusky

$y_1 = 6$). Now we have to find the numbers x_k and y_k with similar properties.

To find x_k, let's square x_{k-1} and take the last k digits of this square, so that $x_{k-1}^2 = 10^k a + x_k$. Let's show that $x_k^2 - x_k$ is divisible by 10^k:

$$\begin{aligned} x_k^2 - x_k &= (x_{k-1}^2 - 10^k a)^2 - (x_{k-1}^2 - 10^k a) \\ &= x_{k-1}^4 - 2ax_{k-1}^2 \cdot 10^k + 10^{2k}a^2 - x_{k-1}^2 + 10^k a \\ &= (x_{k-1}^2 - x_{k-1})(x_{k-1}^2 + x_{k-1}) + 10^k(10^k a^2 + a - 2ax_{k-1}^2). \end{aligned}$$

But $x_{k-1}^2 - x_{k-1}$ is divisible by 10^{k-1}, and $x_{k-1}^2 + x_{k-1}$ is divisible by 10 (since both x_{k-1} and x_{k-1}^2 end in 5). Therefore, both terms in the last expression are divisible by 10^k.

The construction of y_k is a bit more complicated: we have to take the last k digits of y_{k-1}^5. The divisibility of $y_k^2 - y_k$ by 10^k is proved in almost the same way:

$$\begin{aligned} y_k^2 - y_k &= (y_{k-1}^5 - 10^k b)^2 - (y_{k-1}^5 - 10^k b) \\ &= (y_{k-1}^2 - y_{k-1})(y_{k-1}^8 + y_{k-1}^7 + y_{k-1}^6 + y_{k-1}^5 + y_{k-1}^4) + 10^k c, \end{aligned}$$

where $c = 10^k b^2 + b - 2by_{k-1}^2$. The term $y_{k-1}^8 + \dots + y_{k-1}^4$ is divisible by 10 because it's the sum of five numbers ending in 6, and the difference $y_{k-1}^2 - y_{k-1}$ is divisible by 10^{k-1} by the induction hypothesis. So $y_k^2 - y_k$ is divisible by 10^k, and we're done.

Implications

From the first proof, we can see that the sum of the third and fourth numbers we've found is equal to $10^k + 1$ (indeed, 25 + 76 = 101, 625 + 376 = 1,001).

Exercise 1. Prove that this is true for all k.

The second proof demonstrates one amazing fact: the four k-digit endings that are preserved under squaring are obtained from the respective $(k - 1)$-digit endings preserved under squaring merely by adding one digit on the left! Indeed, the last $k - 1$ digits of x_k coincide with those of x_{k-1}^2, and so they constitute x_{k-1}; only the kth digit from the right in x_k is new. In a similar argument for the y's we must use the fact that y_{k-1} comprises the last $k - 1$ digits of y_{k-1}^5: this follows from the identity $y_{k-1}^5 - y_{k-1} = (y_{k-1}^2 - y_{k-1})(y_{k-1} + 1)(y_{k-1}^2 + 1)$ and the divisibility of $y_{k-1}^2 - y_{k-1}$ by 10^{k-1}. You may have noticed this already in our numerical examples:

0—00—000—...,
1—01—001—...,
5—25—625—...,
6—76—376—....

Thus, in the sequence 5, 25, ..., x_{k-1}, x_k, ... , the terms are built up by adding digits on the left end; this is also true of the sequence 6, 76, ... , y_{k-1}, y_k, ... If we don't interrupt this process, it will yield two infinite "numbers"; with ten digits written out they are

$$X = \dots 8212890625,$$
$$Y = \dots 1787109376.$$

We cautiously put the word "numbers" in quotation marks, but, as we'll see in a while, these infinite sequences can quite legitimately be granted the status of genuine, though somewhat unusual, numbers.

A new kind of number

We need a name for our infinite-to-the-left sequences of digits $\dots a_4a_3a_2a_1$. Let's call them *supernumbers*. If all the digits in such a sequence starting from a certain place are zeros, we'll say that this supernumber is an ordinary number—for instance,

$$\dots 000132 = 132.$$

Thus, among supernumbers one can find all ordinary nonnegative integers, but other "numbers" as well. Supernumbers can be added and multiplied using the ordinary digit-by-digit rules (fig. 1). They resemble infinite decimals, as if they were reflected through the decimal point, except that because of the "carry," the rules for performing operations are not reflected. This makes them an essentially new algebraic object. However, for *ordinary* supernumbers these operations are our usual addition and multiplication. It's interesting that supernumbers can be subtracted from one another in any order, also in the usual digit-by-digit way—see figure 2. (The relation "greater/smaller" can't be introduced for supernumbers so as to agree with the algebraic operations as it does for ordinary numbers.) In particular, any supernumber can be subtracted from zero (the supernumber ...000). Therefore, all ordinary negative numbers can also be found among supernumbers. For example, figure 3 illustrates the equality −132 = ...999868. We can see that "ordinary negative integers" are simply supernumbers with an infinite row of 9's to the left.

Addition and multiplication of supernumbers, and subtraction as well, have the usual properties of these operations: $a + b = b + a$, $ab = ba$, $a + (b + c) = (a + b) + c$, $a(bc) = (ab)c$, $a(b + c) = ab + ac$, and so on. In particular, for any supernumber x we can calculate the supernumber $x^2 - x$, and our theorem immediately tells us that *the*

```
a      ......6847
      +......4219
      -----------
       ......1066

b      ......6847
      ×......4219
     ------------
       ......1623
      ......6847
     ......3694
    ...... 7388
    ..........
    -------------
    .........7493
```

Figure 1

```
  ......6847
 -......4219
 -----------
  ......2628

  ......4219
 -......6847
 -----------
  ......7372
```

Figure 2

```
  .....00000
- .....00132
  ----------
  .....99868
```

Figure 3

```
      ......0625
    × ......9376
    ------------
      ......3750
     ......4375
    ......1875
   ......5625
  ..........
  ------------
  .........0000
```

Figure 4

equation $x^2 = x$ has exactly four solution in supernumbers: 0, 1, X and Y.

And that's the result we've been waiting for.

How is that possible?

If the operations on supernumbers obey the usual rules, the usual proof of the fact that the equation $x^2 - x = 0$ has two solutions must work equally well with supernumbers. Let's see if they really do. As before, $x^2 - x = x(x - 1)$—there's nothing to rule out this factorization. Therefore, if $x^2 - x = 0$, then either $x = 0$ or $x - 1 = 0$. But why? How do we know that the product ab of supernumbers a and b is zero only if one of them is zero? In fact, that's wrong! We *know perfectly well* that $XY = 0$!

Indeed, x_k is divisible by 5^k and y_k is divisible by 2^k, so the product x_ky_k is divisible by 10^k— that is, it ends in k zeros. So multiplying X and Y digit by digit, we'll get only zeros in the result. To dissolve any possible doubts you may have, look at figure 4.

If $a \neq 0$, $b \neq 0$, but $ab = 0$, then a and b are called *zero divisors.* There are no zero divisors among ordinary numbers, but we can find them among supernumbers—and that's the gist of the matter.

By the way, now we can verify once again that X and Y satisfy our equation: since $X + Y = 1$ (remember, $x_k + y_k = 10^k + 1$— see exercise 1), $X(X - 1) = -X(1 - X) = -XY = 0$, and, similarly, $-Y(Y - 1) = 0$.

The equation $x^m = x$

If $x^2 = x$, then $x^3 = x^2 \cdot x = x \cdot x = x^2 = x$, $x^4 = x^3 \cdot x = x \cdot x = x^2 = x$, and so on: $x^5 = x$, $x^6 = x$, ... So our four solutions to $x^2 = x$ satisfy $x^m = x$ for any m. But does the equation $x^m = x$ have any other roots?

Let's start with $m = 3$. Besides 0, 1, X, and Y, the equation $x^3 = x$ has a solution $X - Y$:

$$(X - Y)^3 = X^3 - Y^3 - 3XY(X - Y) = X - Y,$$

because $XY = 0$; alternatively, we could prove that $(X - Y)^2 = 1$, which also implies $(X - Y)^3 = X - Y$. In addition, we can reverse the sign of any root of the equation $x^3 = x$ and this will yield four other solutions: -1, $-X$, $-Y$, $Y - X$. All in all, we've found nine solutions:

$$0, 1, -1, X, -X, Y, -Y, X - Y, Y - X.$$

Exercise 2. Prove that the equation $x^3 = x$ has no other solutions.

The case $m = 4$ (and in general the case of any even m—see the theorem below) is not as interesting.

Exercise 3. Prove that the equations $x^4 = x$ and $x^2 = x$ have the same solutions.

In the case $m = 5$ new solutions emerge. To describe them, it will suffice to display one of them. Consider a sequence z_k in which $z_1 = 2$ and any term z_k consists of the last k digits of z_{k-1}^5:

$$z_1 = 2, z_2 = 32, z_3 = 432, \ldots .$$

Exercise 4. Prove that z_k is obtained from z_{k-1} by adding one digit on the left, and that $z_k^5 - z_k$ is divisible by 10^k for all k.

The numbers z_k define the supernumber

$$Z = \ldots 9879186432.$$

Exercise 5. The equation $x^5 - x$ has fifteen solutions: the nine solutions of $x^3 - x$ (see exercise 2) and the supernumbers Z, $-Z$, $X - Z$, $Z - X$, $X + Z$, $-X - Z$.

For greater values of m new solutions don't appear. The reader may try to prove the following theorem.

THEOREM. *If m is even, the equation $x^m = x$ has the same solutions (in supernumbers) as $x^2 = x$. If m has the form $4n - 1$, this equation has the same solutions as $x^3 - x$. And if $m = 4n + 1$, then $x^m = x$ has the same solutions as $x^5 - x$.*

Here is one more statement concerning the equation $x^m = 1$: If m is odd, then $x^m = 1$ has only one solution $x = 1$. If m is even but not divisible by 4, then $x^m = 1$ has four solutions: $1, -1, X - Y, Y - X$. If m is a multiple of four, then $x^m = 1$ has eight solutions: $1, -1, X - Y, Y - X, X - Z, Z - X, X + Z, -X - Z$.

Two more equations

Exercise 6. For what (ordinary) numbers m does the equation $mx = 1$ have a solution in supernumbers? (That is, what ordinary fractions can be found among supernumbers?)

Exercise 7. Is the statement of Fermat's Last Theorem—"if n is an (ordinary) integer greater than 2, then the equation $x^n + y^n = z^n$ has no roots"—true for supernumbers x, y, z?

Conclusion

If the notion of supernumbers rang a bell for you, don't be surprised. They are known to mathematicians as "integer 10-adic numbers." All the troubles—or, to put it more positively, surprises—that we've encountered are due to the fact that 10 is a composite number. If we had considered p-adic numbers with a prime p (which are defined the same way except that we'd have to write them using the number system with the base p), our equation $x^2 = x$ would have had two solutions, as it should. Actually, if our number system had a prime base, we wouldn't have had enough material for an article like this.

The superproblem of space flight

The Tsiolkovsky formula and the resurrection of the fathers

by Albert Stasenko

WHY A "SUPERPROBLEM"? As if there weren't enough problems foisted on poor, defenseless students. Like this one . . .

Imagine that an astronaut, building a structure in the boundless expanse of outer space, swallowed a nut and needs to be taken as quickly as possible from the spacecraft to the space station traveling parallel to it with the same speed at a distance of 100 km. The maximum acceleration the astronaut can endure is $4g$, where g is the acceleration due to gravity. The question would naturally arise for the captain of the spacecraft: what is the *minimum* amount of fuel the ambulance rocket needs to take the victim to the doctors *as quickly as possible*? The speed at which the combustion products are ejected from the nozzle of the engine is constant and equals 2 km/s.

> "The heavenly worlds are the future home of the fathers, as the heavenly spaces can be accessible only to those who are resurrected and resurrecting: the study of heavenly spaces is a preparation for these inhabitants."—N. F. Fyodorov, *The Philosophy of the Common Cause*

Now the usual thoughts occur to the captain. Let the rocket have a velocity v and mass m at a given moment. We mentally divide the rocket into two parts (fig. 1): one part that in the time Δt will move backward (the combustion products), its mass being ΔM; and another part with a mass $m - \Delta M$ that will move forward with the astronaut, attaining the velocity $v + \Delta v$. Denote the velocity of the ejected gas relative to the rocket as u_0; then its velocity relative to the starting point will be $(v + \Delta v) - u_0$. As the separation of the two parts results from the action of internal forces, the total momentum of the rocket and the combustion products does not change (in the coordinate system of the space station and the spacecraft, which move parallel to one another with the same velocity relative to the stars):

$$(m - \Delta M)(v + \Delta v) + \Delta M(v + \Delta v - u_0) = mv.$$

After some algebraic manipulation (the captain had learned to do it in his head way back when he was on Earth), the law of conservation of momentum takes the following form:

$$m\Delta v = u_0 \Delta M. \qquad (1)$$

Taking into account that the ejected mass is equal to the decrease

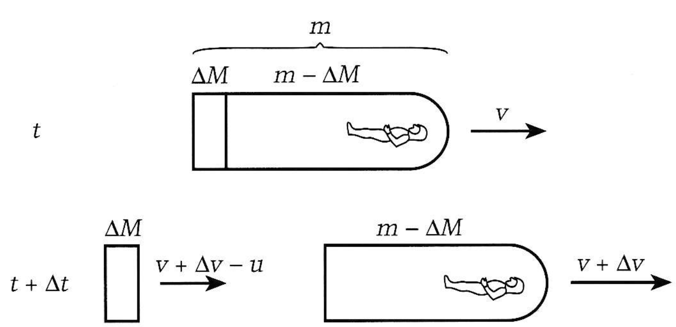


Figure 1

Art by Yury Vashchenko

in the rocket's mass ($\Delta M = -\Delta m$), we can rewrite equation (1) as

$$\frac{\Delta v}{u_0} = -\frac{\Delta m}{m}. \qquad (2)$$

"Well, well," you may be thinking, "this imaginary captain wasn't such a good student, was he? Why, elementary integration of equation (2) gives us the Tsiolkovsky formula

$$\frac{v}{u_0} = \ln\frac{m_0}{m} \qquad (3)$$

(in which we assume that at $t = 0$ the rocket's mass was m_0).

Rest assured, the captain knew equation (3)—otherwise he wouldn't have passed the examinations to become an astronaut. But he had been ordered to take the victim as quickly as possible, yet with an acceleration of no more than 4g, which means that during the flight the acceleration must be constant—that is, $a = a_{max} = 4g$. This means we have to go back to equation (2)—the differential version of equation (3)—and, dividing by Δt, set $\Delta v/\Delta t = 4g$ = constant. Inserting into equation (3) the required velocity change $v = 4gt$, we get the equation for the change in the rocket's mass that must be fed into the onboard computer:

$$\frac{m}{m_0} = e^{-\frac{4gt}{u_0}}. \qquad (4)$$

Traditionally another condition (constant rate of fuel consumption) is examined—that is, μ = constant (kg/s), which means that $m(t) = m_0 - \mu t$. But this leads to a continuously increasing acceleration, which can violate the condition of our problem. So it wasn't for nothing that all these thoughts flashed through the captain's mind!

Not much remains for us to do. The ambulance rocket must come to the space station with zero velocity—any small but finite velocity at contact will give an infinitely large acceleration, which is forbidden. In our problem the trip is made with uniform acceleration. Thus, the motion of the rocket can be depicted as a time-dependent graph (fig. 2). It shows that in the middle of the trajectory, at time $t = \tau/2$ (where τ is the total travel time) the propulsion force must reverse directions. From the well-known laws for motion with uniform acceleration, we obtain

$$\frac{L}{2} = \frac{1}{2}a\left(\frac{\tau}{2}\right)^2 = \frac{g\tau^2}{2},$$

from which we get

$$\tau = \sqrt{\frac{L}{g}}.$$

The final mass of the rocket will be

$$\frac{m(\tau)}{m_0} = e^{-\frac{4g}{u_0}\sqrt{\frac{L}{g}}} = e^{-\frac{4\cdot 10}{2\cdot 10^3}\sqrt{\frac{10^5}{10}}} = \frac{1}{e^2}.$$

Thus, the initial amount of fuel that made up part of the rocket's mass can be obtained from the equation

$$\frac{m_0 - m(\tau)}{m_0} = 1 - \frac{1}{e^2}$$

(where it is assumed that all the fuel is used up the moment the rocket docks with the space station).

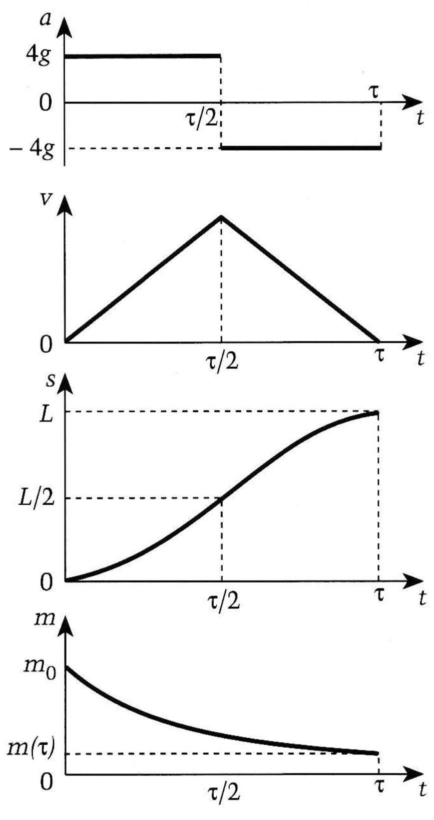


Figure 2

Now here's another problem. Instead of a nut, the astronaut swallows a sandwich (which isn't as harmful) because he suddenly sees an enemy spacecraft approaching with uniform velocity v_0 = 2 km/s. How near should he let the spacecraft approach before he launches a missile if he wants to get the most out of his missile (that is, deliver maximal kinetic energy to the target)? The missile's control devices can bear an acceleration of no more than 100g; the velocity of the gas ejected is the same as in the previous problem (u_0 = 2 km/s).

It's clear that the missile must travel with the maximum acceleration $a = a_{max} = 100g$, and there are no reasons to change the direction of the reactive force, so the relative velocity of the missile and the target changes according to the equation $v = v_0 + at$, and the distance between them $s = s_0 - v_0t - \frac{1}{2}at^2$.

The formula for kinetic energy can be obtained from equation (4):

$$K = \frac{mv^2}{2} = \frac{m_0}{2}e^{-\frac{at}{u_0}}\left(v_0 + at\right)^2.$$

It's clear that this product of a decreasing exponent and an increasing parabola has a maximum. This maximum can be found from the graph of the kinetic energy as a function of time (fig. 3). Those readers who know calculus can set the first derivative equal to zero: $dK/dt = 0$, from which we get

$$-\frac{(v_0 + at)^2}{u_0} + 2(v_0 + at) = 0.$$

This equation has two roots. One corresponds to $v_0 + at = 0 \Rightarrow t_1 = -(v_0/a) < 0$; since t_1 is negative, it relates to the past and so does not interest us. (The spacecraft are getting closer, which means that $v_0 > 0$.) The second root is $t_2 = \tau = (2u_0 - v_0)/a$, which, under the conditions of our problem, yields

$$\tau = \frac{2\cdot 2 - 2}{10^3}10^3 = 2 \text{ s}.$$

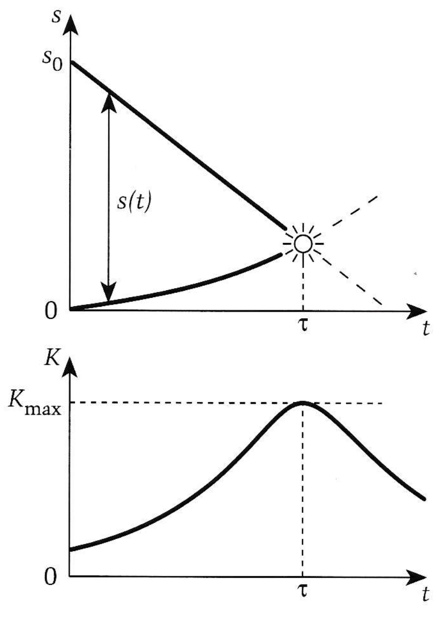


Figure 3

This root could well be negative, if $v_0 > 2u_0$—that is, if the spacecraft are approaching so quickly that the missile doesn't have enough time to reach a velocity corresponding to the maximum kinetic energy:

$$K_{max} = \frac{2m_0 u_0^2 e^{\frac{v_0}{u_0}}}{e^2}.$$

Finally, we obtain the required distance s_0 by substituting τ in the expression for $s(t)$:

$$s(\tau) = 0,$$

$$s_0 = \frac{4u_0^2 - v_0^2}{2a} = 6 \text{ km}.$$

Thus, we have imagined and solved two problems using equation (3), yet how different they are in their aims! One is for help, the other for harm, and both use the most advanced technology. Are human beings destined to tread a vicious circle of creation and destruction forever, despite our ever growing knowledge? Were our greatest minds right to be pessimistic?

"Perfected methods and unclear purposes—these are, to my mind, the characteristics of our time." (Albert Einstein, *The Common Language of Science*)

"Our winged thoughts turn into poultry." (Henry David Thoreau)

"Adhering to old practice, we again consider science as merely a new way to obtain the same old things, bread and arable land. We yoke Pegasus to the plow." (Pierre Teilhard de Chardin)

Still, there is some hope. First of all, there is the very existence of flight, which is somewhat more than just a transition from point *A* to point *B*. Do you remember what you felt the first time you flew, watching clouds beneath and the abyss above, full of stars and the shining Sun? (Physicists have wanted to fly from time immemorial, it seems, beginning with Daedalus—surely he was a physicist?—and his disobedient son Icarus.) For example, in 1783 Jacques Alexander Caesar Charles, an expert in the laws of gases, constructed a balloon just after the Montgolfier brothers. Charles filled his with hydrogen and made several flights, including one over the Alps. In 1804 Joseph Louis Gay-Lussac made two flights in a balloon and reached an altitude of 7 km. In 1887 Dmitry Mendeleyev, who devised the periodic table of the elements, made a solo flight in a balloon to observe a solar eclipse (he was in his forties at the time).

"People who have come out . . . from the infinite universe, appearing as members of our society at this time and at this place, are all brothers and sisters who share the same origin—Infinity—and who share the same parents—the natural environment. We share also our future destinies with these brothers and sisters, for we all return to nature and the infinite universe—our parents and our origin."—Michio Kushi, *The Book of Macrobiotics*

But the person who discovered the remarkable equation (3) above—what was *he* thinking about? Indeed, why would a provincial school teacher be drawn to outer space from the green pastures and forests, neglecting day-to-day life and spending most of his income on experiments and private publication of his works, which were rejected at the time by almost everyone? We can't ascribe it all to a maniacal striving for worldly glory—everything in his work is so right and so fruitful. His motivation must lie deeper.

In his diaries he complained that people considered him "just a one-sided mechanic, not a thinker. A rocket for me is only a means, a method for penetrating into outer space." But why? And then, step by step, a suspicion arises that every great thinker tries to solve some *superproblem* that lies far beyond common sense. The discovery of such a superproblem is a most important factor in the history of science and a most useful element in teaching.

In the history of Russian astronautics one can discern a chain of shining personalities and memorable events—logically connected links amid the jumble of persons and occurrences. Let's begin with N. F. Fyodorov (1828–1903), the illegitimate son of Prince Gagarin and the librarian of the Rumyantsev Museum. Fyodorov was "a modest, unpretentious philosopher in old, shabby, but still tidy clothes" who published almost none of his meditations during his lifetime. What occupied his mind? Nothing less than physical, corporal immortality, and not only that of *future* generations whose science will have reached the required level (this would have been understandable but selfish on their part), but also of *past* generations—in a word, the resurrection of the fathers. He saw this as the main purpose of science, and he even demanded that universities be built in cemeteries so that students would not waste their time on trifles, but think about the victory of life over death.

Just as our modern terminology divides all science into "basic" and "applied," Fyodorov seemed to be an advocate of practical, applied Christianity. In this he was in absolute accord with the tenets of the Orthodox Faith: "I hope for the resurrection of the dead and the life of the world to come."

What was he thinking when he went up to the mountains of the Pamirs? That there the heavens are nearer, the Sun and the Moon are absolutely white, and the stars are colored but do not twinkle? Or that, according to the ancient Aryan tradition, every *mahatma* (teacher) must spend time in the Himalayas?

"Between the multitude of dead generations and the plurality of worlds an expedient relation is possible, to create all the inhabitants of all the worlds from a single blood and the ashes alone of Earth . . . The Earth that swallowed countless multitudes of generations, moved and guided by the heavenly filial love and knowledge, will return those it swallowed and will populate the celestial, now soulless . . . starry worlds with them. It will be a great and wonderful but not miraculous day, as the resurrection will result not from a miracle, but from knowledge and work in common." (N. F. Fyodorov, *The Philosophy of the Common Cause*)

The next link: for three years (1873–1876) this Russian philosopher guided the self-education of K. E. Tsiolkovsky. Surely Fyodorov's ideas were accepted by his pupil as a stimulus for a concerted effort—not the resurrection of the fathers, however, which was and still is impossible, but preparation for the next step: solving the problem of where to put the resurrected as well as future generations of sons. The answer was obvious: upward, toward weightlessness, toward the sea of light, the ocean of energy—into outer space!

And now comes the link connecting the physicist and thinker Tsiolkovsky with the rocket designer Sergey P. Korolyov—the practical link in this chain. Their collaboration led to the first artificial satellite, or "sputnik" in Russian (1957); then the first human being in space (1961) (was it only a coincidence that the first cosmonaut, Yury Gagarin, had the same surname as N. F. Fyodorov's father?); then the first humans on the Moon (1969); and the list goes on.

Thus, physics and technology, to their honor, managed to overcome the Earth's gravitation.

Of course, the tremendous and noble goal of defeating death not only in the future but in the past is far from being reached. (According to Fyodorov, the basic idea occurred to him in 1851, whereupon he assigned the task to human reason to fulfill.) Our current thinking says this is most likely a Utopia—or, perhaps, the sole prerogative of the Almighty. But what noble "spiritual springs" may underlie the simplest formulas in physics—for example, equation (3)!

AT THE BLACKBOARD I

The friction and pressure of skating

Also glaciers, pressure cookers, and the Carnot theorem

by Alexey Chernoutsan

IF YOU WERE TO ASK A TEN-year-old child why ice skates glide so easily on the ice, the answer you'll probably get is the simple and obvious one: "The skates rub against the ice and they make a thin watery film, and this helps the skates slide along the ice." A student more experienced in physics, however, would find this answer too simple and rather boring. "No, no, no," the budding physicist will say, "it's not a matter of friction but of the pressure of the skates on the ice. As the pressure increases, the melting point of ice becomes less than 0°C, so the ice melts under the skates." There is some merit in this explanation—the melting point of ice actually does decrease with an increase in external pressure. However, physics is a quantitative science. If we want to find out whether this physical phenomenon has any bearing on ice skating, we need to produce the appropriate numerical calculations.

First of all, what is the melting point of ice, and why is it interesting? As you may know, when the temperature increases to this point, it can't be raised any further, and any additional heat goes into melting the ice. If no heat is provided from outside, the ice and water coexist in thermal equilibrium. Thus, the melting point is the equilibrium temperature of water and ice at a given pressure. For example, it's equal to 0°C when the pressure is 1 atm.

Art by Sergey Barkhin

How much will the temperature increase when the pressure increases to 1.01 atm?

Surprisingly, it is the Carnot theorem that helps us calculate the shift in the melting point of ice. Yes! The very theorem that deals with the maximum efficiency of heat engines. "What has this to do with the melting point of ice?" you may ask. The point is, Nicolas Carnot proved that the maximum efficiency of a cyclic heat engine of any design does not depend on what this engine uses as its working substance—an ideal gas, melting ice, or a soapy film. Only one condition must be met: the engine must acquire heat at a temperature T_1 and release it at a temperature T_2 (there must be no heat exchange during the intermediate steps). The efficiency of such an ideal engine (known as a Carnot engine) is attained by a slow, reversible process and, regardless of the working substance used, equals

$$\eta = \frac{W}{Q_1} = \frac{T_1 - T_2}{T_1}, \tag{1}$$

where W is the work performed by the engine per cycle and Q_1 is the amount of heat obtained at the temperature T_1.

Let's consider our imaginary Carnot engine to be a vertical cylinder with a piston (fig. 1a). Inside the cylinder we put ice of mass m under pressure P_1 = 1 atm and temperature 0°C (T_1 = 273 K). The pressure is stabilized by a weight set on the piston. To underscore the state of equilibrium between the water and the ice, the figure shows a small amount of water at the corner of the cylinder.

Now let's describe, step by step, what goes on in this kind of Carnot engine during one complete cycle.

1. Let's set the cylinder on a thermal reservoir at a constant temperature T_1 and transmit heat energy $Q_1 = Lm$ to the system necessary to melt all the ice (L is the latent heat of fusion). As a result, the piston sinks a little (fig. 1b), because the volume of ice $V_i = m/\rho_i$ is larger than that of water $V_w = m/\rho_w$. This melting stage is represented by the line 1–2 on the graph, where the coordinates are pressure P and volume V (fig. 2).

2. Now we take the cylinder away from the thermal reservoir, isolate it thermally (fig. 1c), and then increase the pressure very slowly until it is equal to $P_1 + \Delta P$ = 1.01 atm. (This can be done by pouring sand slowly onto the piston.) This will result in a decrease in the temperature down to $T_2 = T_1 - \Delta T$, which is equal to the melting point of ice at a pressure of 1.01 atm.

3. Now we put the cylinder on a thermal reservoir at a temperature T_2 and remove heat until the water freezes again (fig. 1d). In figure 2 this stage is shown by the line 3–4.

4. All that's left is to thermally isolate the cylinder and slowly remove all the sand from the piston. This brings us back to the initial state.

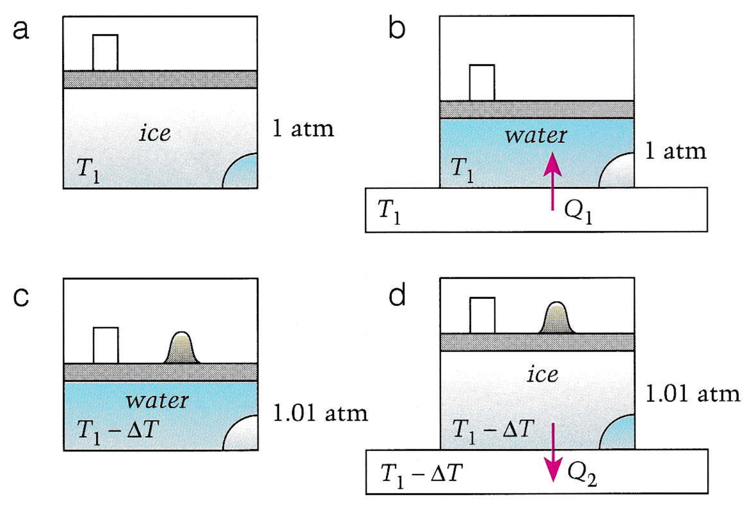


Figure 1

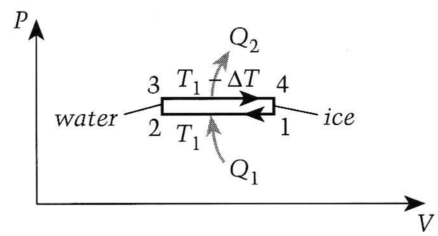


Figure 2

Now let's do some calculations. The work performed during the cycle can be found from the graph—it's equal to the area outlined by the cycle:

$$W = \Delta P(V_i - V_w) = \Delta P\left(\frac{m}{\rho_i} - \frac{m}{\rho_w}\right).$$

The amount of heat obtained from the heater is

$$Q_1 = Lm.$$

Hence, from the Carnot theorem (1) we get

$$\frac{\Delta P\left(\frac{m}{\rho_i} - \frac{m}{\rho_w}\right)}{Lm} = \frac{\Delta T}{T_1},$$

or, for any arbitrary temperature $T_1 = T$,

$$\Delta T = \Delta P \frac{T}{L}\left(\frac{1}{\rho_i} - \frac{1}{\rho_w}\right). \tag{2}$$

This is known as the Clapeyron–Clausius equation. Substituting numerical data in this equation gives us $\Delta T = 9.2 \cdot 10^{-5}$ K for ΔP = 0.01 atm. The effect is clearly very small. To change the melting point by say, 1 K, we need a pressure of about 133 atm. Now we can get back to our skating.

The pressure produced by an ice skater can be estimated as $P = mg/S \cong 600\ \text{N}/2\ \text{cm}^2$ = 30 atm. The corresponding shift in the melting point of ice is about 0.3 K, which is surely too little on a cold winter's day. So the "naive" little kid was right after all:

a lubricant is formed primarily by friction. Then what kind of role do the skates play? It turns out we need them after all! Without going into the details of "skating physics," I'll point out one obvious fact: far less ice needs to be melted to lubricate the blade of a skate (whose surface area is rather small) than it takes to lubricate the sole of a boot.

Are there any other phenomena where an substantial shift in the melting point of ice occurs? Most certainly. For example, consider the way a huge glacier overcomes obstacles as it creeps down a slope. Right where the glacier comes up against a boulder or outcrop, a great amount of pressure builds up, which causes the ice to melt. In a way the glacier flows around the stone and lets it pass through the ice. When the pressure drops, the water in the glacier freezes again.

"This is all very interesting," you may say, "but is that all you have to show for yourself?" Of course not. Let's take a closer look at our result. Now we can calculate the change in the equilibrium temperature of two phases—liquid (water) and solid (ice)—caused by a variation in external pressure. What's most remarkable is that we can apply the result to any other pair of phases, provided they are in thermal equilibrium—for example, liquid–vapor, metal–molten metal, solid–vapor, and so on. In other words, the Clapeyron–Clausius equation holds not only for melting, but for any process involving the transition of a substance from one phase to another (evaporation, sublimation, and so on). For these processes we need to modify equation (2) with the corresponding values for the densities of the substance in the two phases and the latent heat for the phase transition.

By way of example, let's look at the transition of water to vapor. As you know, vapor that is at equilibrium with water is called "saturated." The relation between the temperature of the saturated vapor and its pressure is used to calculate atmospheric humidity, the dew point, and so on. In addition, the

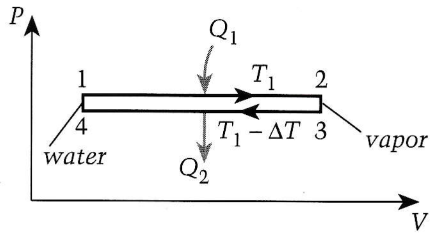


Figure 3

temperature of saturated vapor (that is, the equilibrium temperature in the water–vapor system) determines the boiling point of water at a given external pressure. Thus, at a pressure of 1 atm, the boiling point is 100°C (373 K). Coming at it from the other direction, we know that the pressure of saturated vapor increases with temperature. This phenomenon constitutes the working principle of the pressure cooker, which cooks food at a higher temperature and greater pressure.

What's the difference between the melting of ice and the evaporation of water? Why does the equilibrium temperature drop with an increase in pressure in the one case, and rise in the other? There's another factor involved: volume. When ice absorbs heat and melts, the volume of the system decreases (the density of water is greater than that of ice), but when water absorbs heat and evaporates, the volume of the system increases (the density of saturated vapor is less than that of water). However, in both cases the graph for the cycle must move clockwise in the P–V coordinate system—otherwise the work performed by the Carnot engine would be negative. Compare the trajectories of both plots (figures 2 and 3) and you can see why for one of them a lower temperature corresponds to a greater pressure, while the opposite occurs in the other case. In addition, try to find the sequence of actions you need to perform with a vessel containing water and vapor in order to obtain the Carnot cycle shown in the graph.

To put a cap on our story, let's calculate the shift in the boiling point caused by a pressure increase from 1.00 to 1.01 atm. We replace the latent heat of fusion for ice in equation (2) with the latent heat of vaporization L_v, and the density of ice with the density of saturated vapor ρ_v:

$$\Delta T = \Delta P \frac{T}{L_v}\left(\frac{1}{\rho_v} - \frac{1}{\rho_w}\right).$$

The density of saturated water vapor at $T = 373$ K and $P = 1$ atm can be found from the Clapeyron–Mendeleyev equation:

$$\rho_v = \frac{PM}{RT} \cong 0.58 \text{ kg/m}^3.$$

Substituting numerical data, we get $\Delta T \cong 0.28$ K for $\Delta P = 0.01$ atm.

As you can see, in this case the phenomenon is quite pronounced: to increase the boiling temperature by 1 K, we need to increase the pressure by a mere 0.035 atm, which is perfectly feasible even under ordinary conditions.

KALEIDOSCOPE

Mathematical hopscotch

Let your mind do the jumping

THE RULES OF THIS GAME ARE PRETTY SIMPLE. IN SOME OF THE boxes below you'll find a problem with three possible answers. Where you go next depends on which answer you choose. Other boxes have hints or correct answers and directions where to go next—you either go back to the previous problem or advance to a new one. If you find yourself hopping backward, don't be upset—there's something to be gained even from a roundabout trip! You can keep track of your path in the margins.

If you know a little programming, you can easily turn this into an educational computer game for others to enjoy. Perhaps you can even add a few favorite problems of your own.

1

Try to cross a triangular pyramid with a plane equidistant from all its four vertices. Have you found such a plane? In how many ways can such a plane be drawn?
Three: go to 6.
Four: go to 16.
Seven: go to 11.

7

Exactly! The point is neither the intersection of the angle bisectors of the given triangle nor the intersection of the medians. It's the point of intersection of the angle bisectors of the triangle whose vertices are the midpoints of the given one. Go to 13.

13

If we add 30 to an integer a, or subtract 30 from a, we get a perfect square in either case. The number of such integers a is
infinite: go to 22.
one: go to 2.
two: go to 10.

19

No! This is the smallest of the three numbers. Go back to 5.

20

This answer is wrong. It's true that for an equilateral triangle this center coincides with the point of intersection of bisectors (medians, heights, and so on), but in an isosceles triangle with a small base and large height, for instance, the point of intersection of the bisectors will be near the base, while the perimeter's center of mass will be close to the midpoints of the long sides. Go back to 17.

2

You're close to the truth. If we take 34 instead of 30, the number a would indeed be unique—namely, $a = 290$. Go back to 13.

8

The equation $x^2 = 2^x$ has
one solution: go to 12.
two solutions: go to 15.
three solutions: go to 18.

You're absolutely right! One is neither a prime nor a composite number. Of course, this is just a matter of agreement among mathematicians. At one time the number 1 was considered prime, but later it was considered more convenient to think of it as neither prime nor composite. Go to 17.

No. A prime must have exactly two dif-

2

3

Sorry, you're wrong. The centers of mass of a triangle and its perimeter are generally distinct. Go back to 17.

4

The number 1 is
prime: go to 21.
composite: go to 23.
neither: go to 14.

5

Which of the numbers 2^{121}, 9^{33}, 7^{44} is the greatest?
2^{121}: go to 9.
9^{33}: go to 19.
7^{44}: go to 24.

6

You've found the least obvious positions of the planes. Look for something simpler. Go back to 1.

9

This number isn't the smallest of the three, but it's not the greatest either. Go back to 5.

10

Right! The two numbers are 34 and 226. Indeed, if $a + 30 = n^2$, and $a - 30 = m^2$, then $60 = (n - m)(n + m)$. Both factors on the right are of the same parity, which is possible only in two cases: $n + m = 30$, $n - m = 2$ and $n + m = 10$, $n - m = 6$. This yields two solutions: $n = 16$, $m = 14$ and $n = 8$, $m = 2$, with a equal to 226 and 34, respectively. Go to 5.

11

You're right! Indeed, three of the planes divide the vertices into pairs, and the other four separate one vertex from the remaining three. Go to 8.

12

That's wrong. Take your time! There is at least one other solution, even an integer one. Go back to 8.

14

15

Not bad, but you didn't find them all. Try to sketch the graphs of $y = x^2$ and $y = 2^x$. Inspect them and go back to 8.

16

You haven't examined all possible positions of the planes. In addition to the case you've considered, there's another one, less obvious. Go back to 1.

17

The center of mass of a wire triangle lies at
the point of intersection of its medians: go to 3.
the point of intersection of its angle bisectors: go to 20.
neither of these two points: go to 7.

18

Your guess is right! But can you prove that there are exactly three solutions? These are the numbers 2, 4, and –0.76664. . . Go to 4.

21

No. A prime must have exactly two different divisors, unity and itself, while for the number one they are equal. Go back to 4.

22

You're mistaken. Notice that as n increases the difference between two consecutive squares $(n + 1)^2$ and n^2, equal to $2n + 1$, increases as well, and for $n > 30$ it becomes greater than 60. Therefore, there are only finitely many such numbers. Go back to 13.

23

Well, how could you think it's composite? By definition, a composite number must have at least two divisors distinct from 1. Go back to 4.

24

You're absolutely right! The largest of the three is 7^{44}. To verify this, extract the eleventh root of the given numbers: we get $2^{11} = 2,048$, $9^3 = 729$, and $7^4 = 2,401$.

Congratulations—you made it! We hope you traveled along the shortest path: 1—11—8—18—4—14—17—7—13—10—5—24. ◘

—Compiled by Anatoly Savin

Constructing triangles from three given parts

Of the 186 problems, 28 are still looking for a solution!

by George Berzsenyi

AT THE RECENT ANNUAL MAA/AMS meeting in Cincinnati, I enjoyed a wonderful evening with my mathematician friends, Stanley Rabinowitz (the series editor of *Indexes to Mathematical Problems*), Curtis Cooper and Robert Kennedy (the coordinating editor and problems editor of the excellent *Missouri Journal of Mathematical Sciences*), and Leroy (Roy) Meyers (who served as the problems editor of the *Mathematics Magazine* for several years). Since Roy and I serve on Stan's editorial board, part of the conversation was about books to be published by Stan's company, MathPro Press, in the near future. These include the *Leningrad Mathematical Olympiads, 1987–1991*; the *Problems and Solutions from the Mathematical Visitor, 1877–1896*, the *NYSML–ARML Contests, 1989–1994*, and two more volumes of the *Index to Mathematical Problems*, covering the years 1975–1979 and 1985–1989. All of these should be of great interest to my readers.

Since both Stan and Roy are deeply interested in geometry, our conversation led to some problems in that area, and I learned that Roy was an outstanding expert on the constructibility of triangles from given data. More precisely, he found that there are 186 nonisomorphic problems resulting from choosing three pieces of data from the following list of 18 parts of a triangle:

sides	a, b, c
angles	α, β, γ
altitudes	h_a, h_b, h_c
medians	m_a, m_b, m_c
angle bisectors	t_a, t_b, t_c
circumradius	R
inradius	r
semiperimeter	s

(For the sake of brevity, I have omitted the terms "length of" and "measure of" in the list above. I will also assume that the notation is self-explanatory and/or familiar to all of my readers.)

My first challenge to my readers is to **reconstruct the 186 problems mentioned above**. As a partial aid, I will retain the numbering given to the list of problems by Roy; his list is a variation of one provided earlier by Alfred Posamentier and William Wernick in their *Advanced Geometric Constructions* (Dale Seymour Publications, 1988). The interested reader may wish to consult chapter 3 of this book for a more thorough introduction to the topic.

Basically, the problems fall into four categories:

1. *Redundant* triples, in which any two of the three given parts will determine the third. Of the 186 problems, only (α, β, γ), (α, β, h_c), (a, α, R) fall into this group.
2. *Unsolvable* problems, which do not allow for the construction of a triangle by Euclidean tools (that is, compass and straightedge). There are 27 such triples.
3. *Solvable* problems (by Euclidean tools). There are 128 such problems.
4. *Unresolved* problems. These are listed below, retaining the numbering given to them by Roy in a preprint he recently sent me.

I wish to take this opportunity to thank him for sharing with me and my readers his wonderful findings.

72. a, m_b, t_a	131. a, m_b, r
81. h_a, m_a, t_b	135. h_a, m_b, r
82. h_a, m_b, t_a	138. a, t_b, r
83. h_a, m_b, t_b	142. h_a, t_b, r
84. h_a, m_b, t_c	143. m_a, t_a, r
86. m_a, m_b, t_c	144. m_a, t_b, r
88. a, t_b, t_c	149. m_a, R, r
89. α, t_a, t_c	150. t_a, R, r
90. α, t_b, t_c	165. h_a, m_b, s
110. h_a, m_b, R	172. h_a, t_b, s
117. h_a, t_b, R	173. m_a, t_a, s
118. m_a, t_a, R	174. m_a, t_b, s
119. m_a, t_b, R	179. m_a, R, s
120. t_a, t_b, R	180. t_a, R, s

To prove the unsolvability of some of the problems, Roy found

CONTINUED ON PAGE 55

Challenges in physics and math

Math

M116

Average side and diagonal compared. Prove that the arithmetic mean of the side lengths of an arbitrary convex polygon is less than the arithmetic mean of the lengths of its diagonals. (V. Lev)

M117

With a neighbor on the left. In how many ways can the numbers 1, 2, …, n be permuted so that any number i ($1 \le i \le n$) not in the leftmost place has at least one of its "neighbors" $i - 1$ and $i + 1$ in one of the places on its left side? (A. Anjans)

M118

Choosing the double-greatest. In a rectangular array of different real numbers with m rows and n columns, some numbers are underlined—namely, the k greatest numbers in each column ($k \le m$) and the l greatest numbers in each row ($l \le n$). Prove that at least kl numbers are underlined twice. (S. Konyagin)

M119

A condition for regularity. The base $A_1A_2...A_n$ of an n-sided pyramid $PA_1A_2...A_n$ has congruent sides $A_1A_2 = A_2A_3 = ... A_nA_1$; the angles PA_1A_2, PA_2A_3, …, PA_nA_1 are also congruent (fig. 1). Prove that the pyramid is regular—that is, its base is a regular n-gon and its altitude falls on the center of the base. (V. Senderov, V. Dubrovsky)

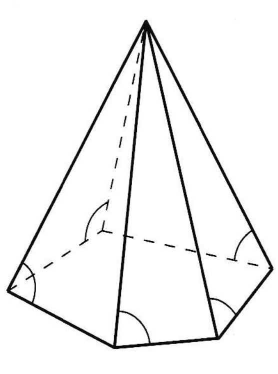


Figure 1

M120

The steepest parabola. Find the smallest positive number a such that for any quadratic function $f(x)$ satisfying $|f(x)| \le 1$ on the interval $0 \le x \le 1$ the inequality $|f'(1)| \le a$ holds. (V. Pikulin)

Physics

P116

Footprints on the water. A water bug can walk on the surface of water without sinking because of surface tension. What do its "footprints" look like on the surface of a calmly flowing river if the bug doesn't move relative to the shore? (S. Krotov, A. Stasenko)

P117

Rod under a dome. For what values of the coefficient of friction can a solid rod of length l with rubber tips remain in a horizontal position under a dome of radius R (fig. 2)? (D. Grigoryev)

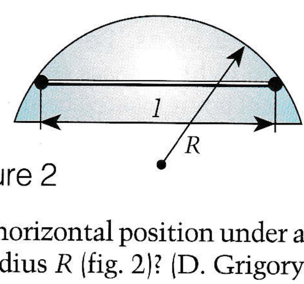


Figure 2

P118

Tea in a thermos. In order to create effective thermal isolation, the air in the space between the inner and outer shells of a thermos bottle was pumped out, resulting in a pressure of $P = 10^{-5}$ atm at room temperature. The volume of the flask is 1 l, its surface area is $S = 600$ cm^2. Estimate the time it takes tea to cool from 90°C to 70°C. The specific heat of water $c = 4.2 \cdot 10^3$ J/(kg · K) and the molar gas constant $R = 8.3$ J/(K · mol). Neglect the heat loss through the cap of the thermos. (A. Stasenko)

P119

Particle near a wire. A charged particle moves with kinetic energy K past a long, uniformly charged wire. The particle travels in the plane perpendicular to the wire and deviates from its initial path by a small angle α (fig. 3). Find α assuming that the particle's charge is e and the charge of a unit length of the wire is λ. At a distance R from the wire, the electric field $E = \lambda/2\pi\varepsilon_0 R$. (V. Mozhayev)

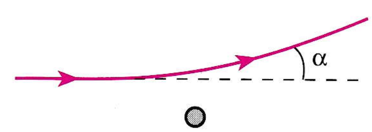


Figure 3

P120

Light through a lens. A light beam entering a thin lens intersects its principal axis at an angle $\alpha = 4°$ at a distance $d = 12$ cm in front of the lens and exits at an angle $\beta = 8°$ (relative to the principal axis). Find the focal length of the lens. (V. Deryabkin)

ANSWERS, HINTS & SOLUTIONS ON PAGE 49

Mirror full of water

"The empty mirror. If you could really understand that, there would be nothing left to look for."—van de Wetering

by Arthur Eisenkraft and Larry D. Kirkpatrick

"IT'S DONE WITH MIRrors." Whether we attend magic shows and see the magician push sharp swords through a box containing the "lovely assistant" or ride the "Haunted Mansion" and see the ghost flying through the room at Disneyland, we are often surprised and pleased by clever manipulations of images.

In this contest problem, we'll look at the image produced by a concave mirror filled with water. Because our confidence in a physics solution increases if different approaches to the problem yield the same result and there are many ways of obtaining the position of the image, we will want to discover as many of them as possible. Perhaps you will come up with a solution that is fundamentally different from the ones we expect.

Texts on geometrical optics often begin by showing that the reflection of light from plane mirrors follows the principle that the angle of incidence is equal to the angle of reflection. If the mirror is curved, this behavior still holds, but the geometry of the parabolic mirror is such that all parallel rays come to a focus for a concave mirror, or appear to diverge from the focus in the case of a convex mirror. For a spherical mirror, the spherical surface approximates the parabolic curve and parallel rays near the axis also come together at (or diverge from) the focus.

The relationship between the image and object is given by the mirror formula

$$\frac{1}{s}+\frac{1}{s'}=\frac{1}{f},$$

where s and s' are the distances of the object and image from the surface of the mirror and f is the focal length of the mirror. The focal length is often stamped on the mirror and is equal to one half of the radius of the spherical surface from which the mirror is made. The focal length can be measured by shining a beam of parallel light onto a concave mirror and measuring the distance from the surface of the mirror to the point where the beam is brought to a focus. For a convex mirror the light appears to diverge from a focal point located behind the mirror. Both of these points can be determined by drawing several rays parallel to the axis of the mirror, using the law of reflection at the surface, and locating where the rays cross.

To make effective use of the mirror formula we must remind ourselves of a number of conventions. The distance s is positive if the object is located in front of the mirror. This will always be the case for real objects, but the "object" could be an image produced by another optical device. In this case, the object could be located behind the mirror and s would be negative. If the image is located in front of the mirror, the image distance s' is positive; if the image is behind the mirror, s' is negative. Finally, f is positive for a concave mirror and negative for a convex mirror.

As an example, consider an object located a distance $3f$ in front of a concave mirror:

$$\frac{1}{s'}=\frac{1}{f}-\frac{1}{s}=\frac{1}{f}-\frac{1}{3}f=\frac{2}{3}f.$$

Therefore, the image is located a distance $3f/2$ in front of the mirror. This can also be shown with a diagram that traces the rays. Convince yourself that the image would be $3f/4$ behind the mirror if we use a convex mirror instead of the concave mirror.

The mirror formula also works for lenses if we adopt the following conventions: s is positive if the object is located in front of the lens, negative if the object is located behind the lens; s' is positive if the image is located behind the lens,

Art by Tomas Bunk


LENS
IMAGE
MIRROR
WATER
LOOK
OBJECT
FOPT
KHNP
STPX
FLTZY
KLPBH
RIP
TOMAS BUNK 94

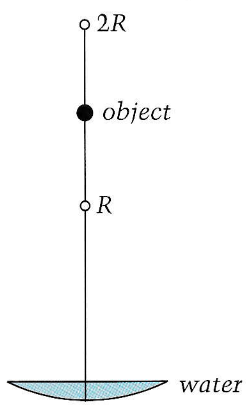
negative if the image is located in front of the lens. Converging lenses (thicker in the center than at the edges) have positive f, while diverging lenses (thinner at the center) have negative f.

Another useful relationship is the lens maker's formula. For the special case when one of the surfaces is planar, it tells us that

$$\frac{1}{f} = \frac{n-1}{R},$$

where n is the index of refraction of the lens material and R is the radius of the curved surface.

Now that we have completed this very brief review, let's take a look at our problem. A concave mirror of radius R resting face up on a table top has been filled with a small amount of water (index of refraction $n = 4/3$) as shown in the figure below.



A small object is located a distance $d = 3R/2$ from the mirror along the optic axis. Where is the image located? In the spirit of the "thin lens approximation" often used in such problems, we will neglect the thickness of the water.

A. Let's begin by using a technique used by eye doctors. Often the doctor will place a lens in front of your glasses to show you how the new lenses will work. This works because the effective focal length f' of two (or more) lenses (or mirrors) in close proximity is given by

$$\frac{1}{f'} = \sum \frac{1}{f_i}$$

—that is, the focal lengths add as reciprocals. Therefore, the mirror–water combination can be replaced by a mirror with an effective focal length and you can use the mirror formula given above. Does the water lens appear in the sum once or twice? Use the other methods to check yourself.

B. You can also obtain the effective focal length by tracing a ray parallel to the optic axis as it enters the water and bends according to Snell's law, reflects from the mirror surface, and exits the water again. Don't forget to make suitable approximations.

C. Our third method makes use of the observation that images formed by one optical element act as objects for subsequent optical elements. Begin by finding the location of the image formed by the air–water interface. Use this image as the object for the mirror (without the water) and find the new image location. Then find the image of this image formed by the water–air interface when the light exits the water. This is the final image produced by the combination.

D. The trickiest method treats the combination as a water lens, a mirror, and a water lens in combination. Find the location of the image produced by each element and then use it as the object for the next element. This is tricky because it's very easy to make mistakes with the sign conventions.

Please send your solutions to *Quantum*, 1840 Wilson Boulevard, Arlington VA 22201-3000 within a month of receipt of this issue. The best solutions will be noted in this space and their authors will receive special certificates from *Quantum*.

Stop on red, go on green . . .

Quantum readers were asked to determine when it's safe to go through a yellow light and when it's safe to apply the brakes at the yellow light. We hope many of our readers thought about the problem as they waited at an intersection for the light to change.

An excellent solution was submitted by Ophir Yoktan of Israel. Unfortunately, Yoktan provided no biographical information and so we don't know if Yoktan is a professor or a student. Irrespective of that, the solution presented here closely follows Yoktan's submission.

A. (a) In the "go zone" a person will be able to continue at the traveling speed and get through the intersection within the time that the yellow light is illuminated. This depends on the velocity of the car v_0, the yellow light time t_y, the width of the intersection w, and the length of the car l. (The go zone is quite different for a stretch limo and a compact car.) This gives us

$$d_g < v_0 t_y - w - l.$$

(b) In the "stop zone" a person will be able to stop before the intersection. It depends on the velocity of the car v_0, the acceleration a of the car while braking (a negative number), and the reaction time of the driver t_r:

$$d_s > v_0 t_r - \frac{v_0^2}{2a}.$$

(c) Whether we have a "dilemma zone" or an "overlap zone" depends on the difference between the go zone and the stop zone. If the zone is defined as the go zone minus the stop zone, a negative value will indicate a dilemma zone and a positive value will indicate an overlap zone:

$$\text{zone} = v_0 t_y - w - l - v_0 t_r + \frac{v_0^2}{2a}$$

$$= \frac{1}{2a} v_0^2 + \left(t_y - t_r\right) v_0 - (w + l).$$

B. We can find the conditions for which there will always be a dilemma zone by requiring that the zone be negative:

$$0 > \frac{1}{2a} v_0^2 + \left(t_y - t_r\right) v_0 - (w + l).$$

See the equation in the box on the next page. A dilemma zone will always exist if the terms within the radical sign are negative. This occurs

$$v_{1,2} = \frac{-(t_y - t_r) \pm \sqrt{(t_y - t_r)^2 - 4 \cdot \frac{1}{2a} \cdot (w + l)}}{2 \cdot \frac{1}{2a}}$$

$$= a\left[(t_r - t_y) \pm \sqrt{(t_y - t_r)^2 + \frac{2(w + l)}{a}}\right]$$

if the response time is greater than the yellow light time. (In this unrealistic case the yellow light time does not allow any decision making.) This also occurs when the term containing the acceleration is small in comparison to the difference in the yellow light and reaction times.

If the radical is positive, we will always get two positive values for v_0. Calling the smaller root v_1 and the larger root v_2, we see that there will always be a dilemma zone if $v_0 > v_2$ or $v_0 < v_1$ and an overlap zone if $v_1 < v_0 < v_2$. Physically it's easy to understand why a high speed can produce a dilemma zone. Why can a low speed produce a dilemma zone?

Another way of viewing the problem is to realize that people will usually be traveling at a typical speed for this type of road and we wish to set the yellow light time to make it a safe intersection. In this case, we should solve the equation for the yellow light time t_y. The dilemma zone exists for the following values of t_y:

$$t_y < t_r + \frac{(w + l)}{v_0} - \frac{v_0}{2a}.$$

An overlap zone exists when t_y is larger than this value.

This new equation allows us to set the yellow light time at an intersection, since we can assume that the velocity, the braking acceleration, the response time, and $(w + l)$ are constants.

C. If the car is traveling downhill, there is an acceleration equal to $g \sin \alpha$, where α is the slope. If we let $g' = g \sin \alpha$, then

$$\text{go zone} = v_0 t_y + \frac{1}{2} g' t_y^2 - w - l,$$

$$\text{stop zone} = v_0 t_r + \frac{1}{2} g' t_r^2 - \frac{(v_0 + g' t_r)^2}{2(a + g')}.$$

The overlap zone is, once again, the difference between the go zone and stop zone. The go zone increases as a result of the hill (if the driver allows the car to accelerate down the hill—this is not safe driving!); the stop zone also increases (if the driver loses some braking acceleration as a result of the hill—this is not usually true). We can see from these equations that if the acceleration a is due to a heavier foot on the gas pedal, the go zone does increase, as one might expect. This increase in the go zone is much smaller than we might anticipate, however, and the acceleration would lead to bigger problems if an accident were to occur at the higher speeds. ◘

Suggestive tilings

Another crack at theorems by Napoleon, Pythagoras, and Pick

by Vladimir Dubrovsky

OPEN AN OLD HIGH SCHOOL geometry textbook—you'll find hardly any mention of tilings. This isn't surprising: traditional geometry has been developing over thousands of years, whereas this mathematical notion has become a subject of study for mathematicians comparatively recently. Nowadays, due to their increasing role in modern geometry and, in no small measure, their inherent attractiveness, they are gradually penetrating the school curriculum. And, not to be outdone, *Quantum* has published a number of articles on this topic (see the references below).

Because tilings have a lot of interesting properties, they are usually studied by themselves. In this article, I want to show how they can be applied to solving traditional, even classical, geometry problems that originally had nothing to do with tilings. The solutions we'll discuss are not the only ones possible, nor are they always the shortest. But they certainly are beautiful, and I hope you'll enjoy them as much as I did.

We'll start with the two simplest kinds of tilings, shown in figure 1. These are the parallelogram and triangle tilings. We obtain the first by cutting the plane along two sets of equidistant parallel lines, and the second by additionally cutting the parallelograms of the first tiling along parallel diagonals. Clearly, a parallelogram of any size and shape can be used as a sample tile for a tiling of the whole plane. The same is true for any triangle.

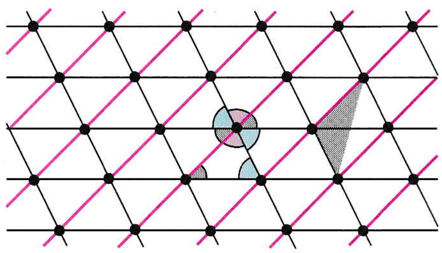


Figure 1

Even the simple triangular tiling can be used to understand a fundamental geometric theorem. Look at figure 1 again. Do you see what theorem I mean? I'm sure you do. Yes, it's the theorem about the sum of the angles of a triangle. In the figure, the angles of six triangular tiles fit around their common vertex leaving no gaps; since each angle of a tile occurs twice among these six angles, the sum of the angles of one tile is 180°.

So, staring at this tiling can help us discover a basic fact about geometry. This "power of suggestion" manifests itself in most of the "tiling solutions" we'll discuss below. Also, our simple example explains one of the reasons why tilings turn out to be useful in finding certain geometric facts and their proofs: when a figure is surrounded by its copies in a tiling, its parts come together to make visible relationships that originally were obscure.

There is one more feature, of a more technical character, that makes tilings useful. This requires a more detailed explanation.

Measuring areas by counting

Consider a parallelogram tiling and an arbitrary figure on the plane. The area of the figure is approximately equal to the number of parallelograms contained in it times the area of one parallelogram. The finer the tiling, the more exact this equality is. Here we think of the figure as fixed and of the tiling as getting finer. But we can also think of the tiling as fixed and the figure as being dilated. Then, assuming the tiles are of unit area, the area of our figure is approximately equal to the number of tiles contained in it, and the relative error gets smaller as the figure gets bigger.

Instead of counting parallelogram tiles, we can count their vertices. The set of all vertices of parallelograms in a tiling is called a *grid*; the vertices themselves are its *nodes*. Each node of a grid is the left bottom vertex of one and only one parallelogram of the associated tiling. With this correspondence in mind, we can see that the number of nodes in a figure is not less than the number of parallelograms contained in it and not greater than the number of parallelograms that have common points with it. So, taking the area of one parallelogram of our tiling for the unit area (which will be assumed throughout the article unless otherwise noted), we can say that the area of a figure is approximately equal to the number of nodes it covers.

This approximation gets more and more accurate as the figure is scaled up, but we can never achieve an absolute equality for all figures. However, for certain figures there are formulas that allow us to find the exact values of areas by counting the nodes they cover. One such formula is given by Pick's theorem, which says that the area of a polygon whose vertices are nodes of a grid is equal to $i + b/2 - 1$, where i and b are the numbers of nodes inside and on the border of the polygon, respectively. (For details and a proof, see the article "Chopping Up Pick's Theorem" in the January/February issue.) However, we'll use another formula below. If, on a parallelogram grid, we superimpose a new tiling, made of congruent copies of some figure, this formula gives the area of one of these tiles.

Consider first the simplest case: suppose each tile of the tiling in question contains the same number n of nodes of our initial grid in its interior and no nodes on its border. Then the area t of a tile will simply be equal to n.

Indeed, let's take a circle of a big radius R. Suppose it contains N tiles of the tiling in question. Then its area πR^2 is approximately equal to Nt (this is true for a parallelogram tiling and, of course, for any other tiling as well). This means that $Nt/\pi R^2 \to 1$ as R grows to infinity. On the other hand, as we've seen, the area of the circle is approximately equal to the number of nodes in it, which, in turn, is approximately equal to Nn. So $Nn/\pi R^2 \to 1$ as well. Therefore, $Nn/Nt = n/t \to 1$ as R grows to infinity. But this is possible only if $t = n$, since n and t do not depend on R.

A similar argument can be applied when tiles have nodes on their borders. But in this case, when we calculate the number n of nodes covered by a tile, we must count a border node with a factor $1/k$, where k is the number of tiles it belongs to. Then, adding the numbers n over all the N tiles in a big circle, we'll count such a node k times (with every tile containing it). Every time it is counted it gives a contribution of $1/k$, so its total contribution will be 1. The number of nodes in the circle will again be (approximately) Nn, leading to the desired equality $t = n$.

By way of example, let's look again at our initial parallelogram tiling (fig. 1). Each parallelogram tile covers four nodes, but each node is covered by four tiles, so the average number of nodes per tile is $4 \cdot 1/4 = 1$, which is the area of a tile. A more interesting example is given by an arbitrary triangle with its vertices on the grid such that, other than the vertices, there are no other nodes inside or on the border of the triangle (see the shaded triangle in figure 1). Let's show that the area of any such triangle is 1/2 (that is, one half of the area of one grid parallelogram, taken, as we assumed, to be the unit area).

Tile the plane with congruent copies of the given triangle as shown in figure 2. Then all these tiles will have their vertices on the grid and won't have any other nodes inside them or on their sides. (We can show this by using the fact that the grid is taken into itself under any translation by vector $\overrightarrow{AB}$, where A and B are arbitrary nodes.) Now, every triangular tile covers three nodes, and every node belongs to six tiles, so the average number of nodes per tile is $3 \cdot 1/6 = 1/2$, and we're done. By the way, now you can try to prove Pick's Theorem by cutting an arbitrary polygon with vertices on the grid into triangles of the sort considered above and counting the number of these triangles.

This method of calculating the area of a tile can be applied when the tiling and the grid are related so that any isometry that carries one tile into another maps the entire grid onto itself. This ensures that the pattern of nodes in all tiles is the same.

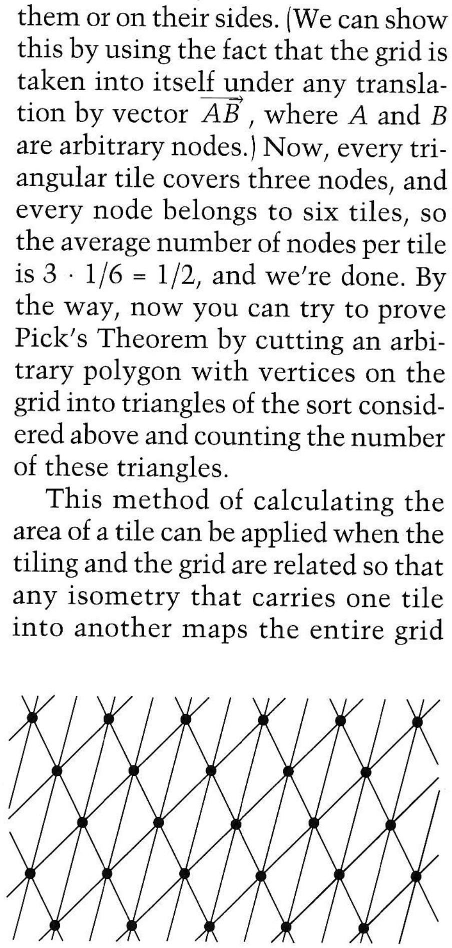
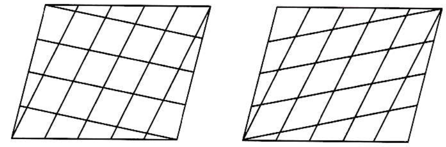


Figure 2



a b

Figure 3

Exercise 1. Two opposite sides of a parallelogram of unit area are divided into n equal parts, two other sides are divided into m equal parts. The points of division are joined in two different ways as shown in figures 3a and 3b. Find the areas of the small parallelograms thus obtained.

In this problem the grid is, in effect, given by the condition. Sometimes, as in the following exercise, you have to *create* a suitable grid.

Exercise 2. Points A_1, B_1, C_1 are given on the sides BC, CA, AB of a triangle ABC such that $BA_1 : A_1C = CB_1 : B_1A = AC_1 : C_1B = 1 : 2$. The triangle ABC is of unit area. Find the area of the triangle formed by the lines AA_1, BB_1, CC_1.

Pythagoras revisited

One of most beautiful applications of tilings is the proof of the Pythagorean Theorem illustrated in figure 4, which adds one more item to the collection of proofs in "The Good Old Pythagorean Theorem" (January/February 1994).

Take two squares—a small one and a bigger one (they'll be the squares constructed on the legs of a

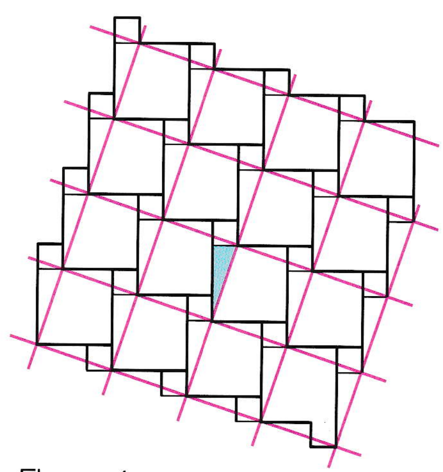


Figure 4

right triangle) and tile the entire plane with copies of these squares as shown in figure 4. (To prove strictly that this is really possible, and that our figure isn't merely an optical illusion, we can make *b*-shaped tiles out of bigger and smaller squares, make infinite bands of these tiles fitting the bottom left area of each tile into the "notch" on another, and then pave the plane with these bands without gaps.) Now mark the bottom left corner of each bigger square red (which is at the same time the bottom right corner of a certain smaller square). The red points form a square grid (we can prove this using the symmetries of our tiling). The side length of each red square is equal to the length of the hypotenuse of a right triangle whose legs are the sides of small and big square tiles—look at the blue triangle in the figure. Now it remains to notice that each *b*-shaped tile covers two red nodes and each red node is covered by two such tiles, and so the area of this tile (the sum of the areas of the squares on the legs of our right triangle) equals $2 \cdot 1/2 = 1$—that is, it equals the area of any red square (the square on the hypotenuse).

Notice that our tiling also shows how to cut a *b*-shaped tile (a pair of "leg" squares) into pieces that can be rearranged to form a red grid square: simply cut a *b*-tile along the red lines. Thus, we get one more proof of the famous theorem. Similarly, most of the problems about areas in this article can be solved both by counting nodes or by the cut-and-paste method—you can choose whichever you like better.

Quadrilaterals and hexagons

The more diverse tilings we deal with, the more diverse the results we're likely to get. Let's see what results can be derived from a tiling with congruent quadrilateral tiles of arbitrary shape.

To obtain such a tiling, consider a parallelogram tiling (the red lines in figure 5) and mark a point in each parallelogram in the same position with respect to the parallelogram (see the top part of figure 5); the marked points constitute a grid con-

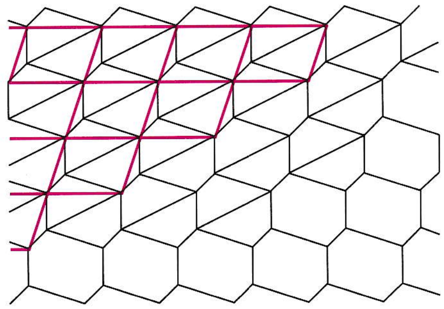


Figure 5

gruent to the grid of vertices of the tiling. Join each marked point to the vertices of its parallelogram and erase the red lines to produce a tiling that consists of congruent convex quadrilaterals. Clearly, a convex quadrilateral of any shape can be used as a sample tile for such a tiling: we can begin with a tiling by parallelograms whose sides are equal and parallel to the diagonals of the given quadrilateral, mark appropriate points in these parallelograms, and proceed as above. (By the way, this construction works also with nonconvex quadrilaterals—we merely have to join each marked point to the vertices of a certain parallelogram *not containing* this point, say, *n* rows above and *m* rows to the right of the parallelogram to which the point belongs. You might like to draw such tilings yourself and see whether the facts we discuss below remain valid for them.)

Every two adjacent quadrilateral tiles in our tiling are symmetric about the midpoint of their common side, because their corresponding sides are parallel and congruent. Red lines (sides of parallelograms) cut these tiles into pairs of triangles, and we see that the four triangles obtained from two adjacent quadrilaterals can be shifted to fill one red parallelogram. So the area of any of our quadrilaterals is half that of the parallelogram. Check this by counting nodes! Another simple consequence is that for any point *P* in a parallelogram *ABCD* the sum of the areas of triangles *PAB* and *PCD* is equal to the sum of the areas of *PBC* and *PDA*—each pair of triangles is obtained from the same quadrilateral by cutting it along different diagonals.

This property of pairs of triangles is almost obvious. But its generalization given in the next exercise will certainly require some serious thought.

Exercise 3. (V. Proizvolov) A small square lies inside a big square. Their vertices are joined to form four quadrilaterals as shown in figure 6. Prove that the sum of the blue areas in this figure is equal to the sum of the pink areas.

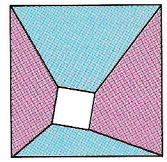


Figure 6

The following exercises demonstrate that the range of applications of the tiling technique is wider than one might think judging by the problems considered above.

Exercise 4. Prove that a midline of a quadrilateral—that is, a segment joining the midpoints of two opposite sides—is no longer than the half-sum of the other two sides and is equal to this half-sum only if the last two sides are parallel.

Exercise 5. In ancient Egypt the area of a quadrilateral was calculated as the product of the half-sums of its opposite sides. Prove that this formula yields a correct result only for rectangles.

Exercise 6. (I. Goldsheid). Let *P* be an arbitrary point in a rectangle *ABCD*. Prove that the area of *ABCD* is not greater than $PA \cdot PC + PB \cdot PD$. (Hint: if you want to solve this problem using tilings, you'll have to modify the construction of a quadrilateral tiling described above).

Let's turn back to figure 5. If we erase the common side of two adjacent quadrilateral tiles in this figure and all the sides of quadrilaterals parallel to it, we'll get a tiling of the plane with hexagons. These hexagons have symmetry centers (the midpoints of erased sides) and their opposite sides are parallel and congruent to each other. Of course, any centrally symmetric hexagon can serve as a sample tile for such a tiling.

Exercise 7. Three alternate vertices of a centrally symmetric hexagon are joined to form a triangle. Prove that the area of this triangle is half that of the hexagon.

Napoleon's problem

The example I'm going to present now, at the end of the article, is certainly the most remarkable and surprising. It's a rather well-known theorem often associated with the name of Napoleon Bonaparte, and it has already appeared in *Quantum* (see "Botanical Geometry" in the September/October 1990 issue). This is what it says.

THEOREM. *Let ABC be an arbitrary triangle. Let ABC_1, BCA_1, and CAB_1 be equilateral triangles constructed externally on the sides of ABC, and P, Q, and R their centers (fig. 7). Then PQR is also an equilateral triangle.*

Figure 8 shows that the triangles considered in this theorem can be embedded in a certain tiling. We can view it as a hexagonal tiling like those we considered above, in which every hexagon is subdivided into three equilateral triangles congruent to ABC_1, BCA_1, and CAB_1, and three triangles congruent to ABC. To make sure this tiling really exists, construct triangles B_1EA and C_1AD congruent to ABC (such that the vertices correspond, as implied by the notation: $\angle EB_1A = \angle AC_1D = \angle BAC$, and so on). Then $B_1E = AB = BC_1$, and $\angle EB_1C + \angle B_1CB + \angle CBC_1 = (\angle EB_1A + 60°) + (60° + \angle ACB) + (\angle CBA + 60°) = 180° + (\angle BAC + \angle ACB + \angle CBA) = 360°$, which means that B_1E is parallel to BC_1 (why?). Similarly, CB_1 is congruent and parallel to C_1D. It follows that B_1EC_1B and CB_1DC_1 are parallelograms with a common diagonal B_1C_1, so their diagonals BE, B_1C_1, and CD have a common midpoint, and, therefore, the hexagon CBC_1DEB_1 is symmetric about this midpoint, which allows us to tile the plane with its copies. Notice that $AE = AD = DE$ (because all three segments are congruent to BC), so ADE is an equilateral triangle congruent to BCA_1.

Now consider the "Napoleon triangle" PQR together with all the triangles constructed in the same way: by joining the centers of colored equilateral triangles (in our figure, with red lines). All these triangles are congruent, because half of them can be viewed as obtained by exactly the same construction as PQR—by joining the centers of the three equilateral triangles on the sides of a certain (white) triangle congruent to ABC, and any of the remaining triangles has three sides respectively congruent to the sides of any triangle of the first sort. Now, every node of the red-line grid is a common vertex of six such triangles (congruent to PQR). In all six triangles, it's not too difficult to see that the angles at this vertex correspond to each other (since they are opposite corresponding sides of the triangles). Therefore, each of these angles is 60°. And this is true for all the angles of all these triangles. So they are indeed equilateral!

The theorem is proved, but as a reward for our toil (which wasn't all that arduous, was it?) we get a neat formula for the area of the triangle PQR. We can view the nodes of the red tiling as a parallelogram grid, so that each equilateral triangle has area 1/2. Each of our hexagonal tiles contains three nodes of the red grid inside it and no nodes on its border, so its area is 3 "red grid units." Thus, the area of the hexagon is six times that of PQR. On the other hand, the hexagon consists of three copies of triangle ABC and three equilateral triangles constructed on its sides. So we come up with the following formula:

$$\text{area}(PQR) = \frac{1}{2}\text{area}(ABC) + \frac{\sqrt{3}}{24}\left(a^2 + b^2 + c^2\right),$$

where $a = BC$, $b = CA$, $c = AB$. ◘

ANSWERS, HINTS & SOLUTIONS ON PAGE 53

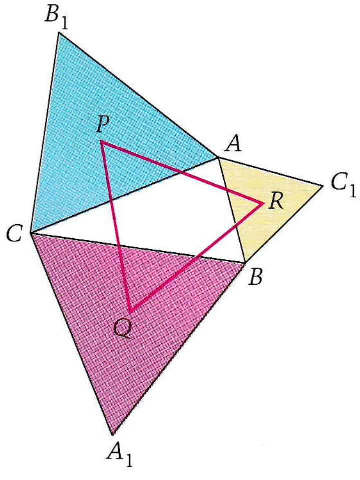


Figure 7

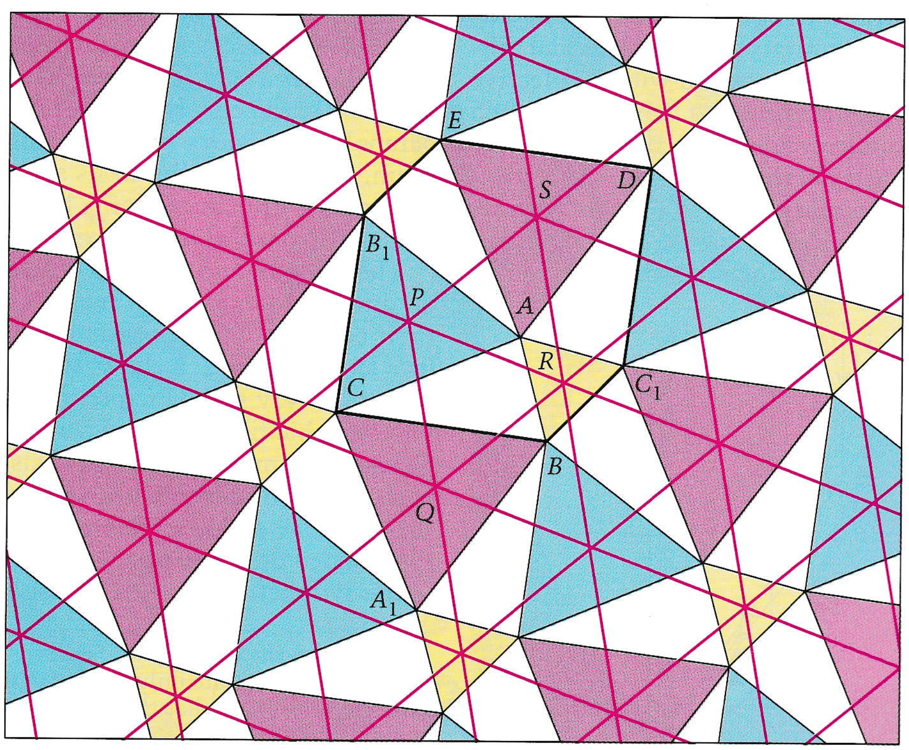


Figure 8

The amazing paraboloid

Double reflection and redistribution of energy

by M. I. Feingold

THE "PHYSICAL" DEFINITION of the optical properties of a parabola can be based on the following property: a beam of light that strikes a parabola parallel to its axis of symmetry passes through its focus after reflection. From the principle of reversibility of light it follows that a beam coming from the focus of a parabola will travel parallel to its symmetry axis after reflection. In this article we'll look at some purely "physical" features of light reflection from a paraboloid—that is, the surface formed by revolving a parabola about its axis of symmetry. The paraboloidal mirror is a paraboloid with a reflective interior surface. If light falls on such a mirror parallel to the axis of symmetry of the paraboloid, after reflection all the beams pass through its focus as if they were being collected there. On the other hand, rays emanating from a point source at the focus will propagate as a parallel beam after reflection on the mirror surface.

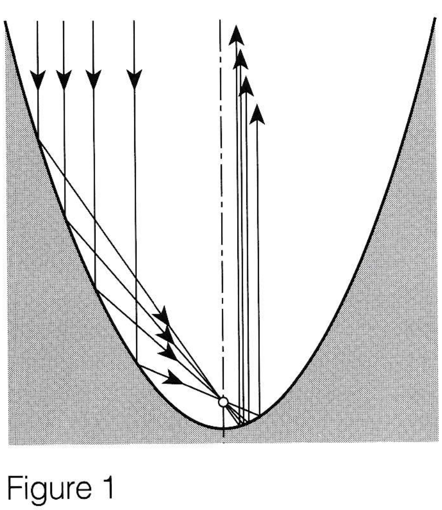


Figure 1

I should point out that both of these effects are the result of only one reflection of the rays from the paraboloidal surface. If a paraboloid is rather deep, most of the entering rays will be reflected twice (fig. 1). After the first reflection each beam, having passed through the focus, will again be reflected from the opposite side of the paraboloid. In other words, the focus becomes a kind of point source of light.[1] But the rays of such a source leave the paraboloid as a parallel beam. Thus, we come to the conclusion that the paraboloid converts the incoming beam, which is parallel to its symmetry axis, into an outgoing beam that is also parallel to this axis.

The incident and reflected beams do differ, however, with regard to their energies. To understand this, let's look at figure 2. It shows the results of a very simple experiment. Photographic film is placed perpendicular to the mirror's symmetry axis, with the photosensitive layer facing the reflective surface. The mirror is illuminated by a light beam parallel to the symmetry axis with homogeneous energy distribution over the cross section—that is, an equal amount of energy passes per unit time through a unit area placed at any location perpendicular to the beam. The incoming light would strike the film homogeneously. However, in our experiment the film is struck by reflected light, which produces the result shown in figure 2.

[1] In reality the designation "point source" is appropriate only to a certain extent, in the sense that the dimensions are small. Later we'll see how this source differs from the point source.

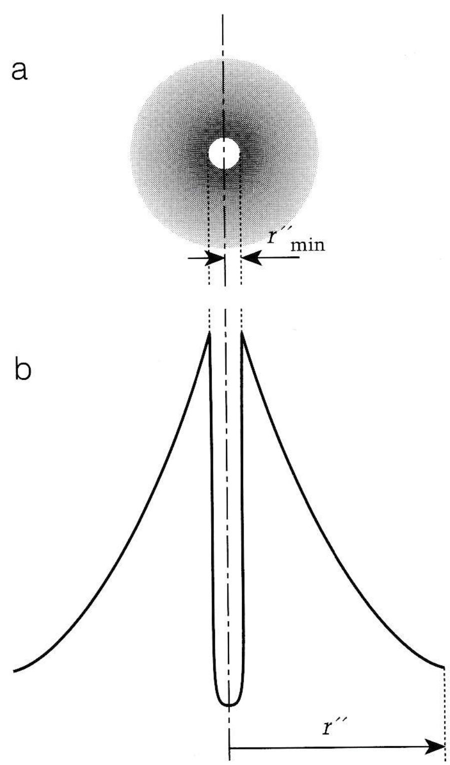
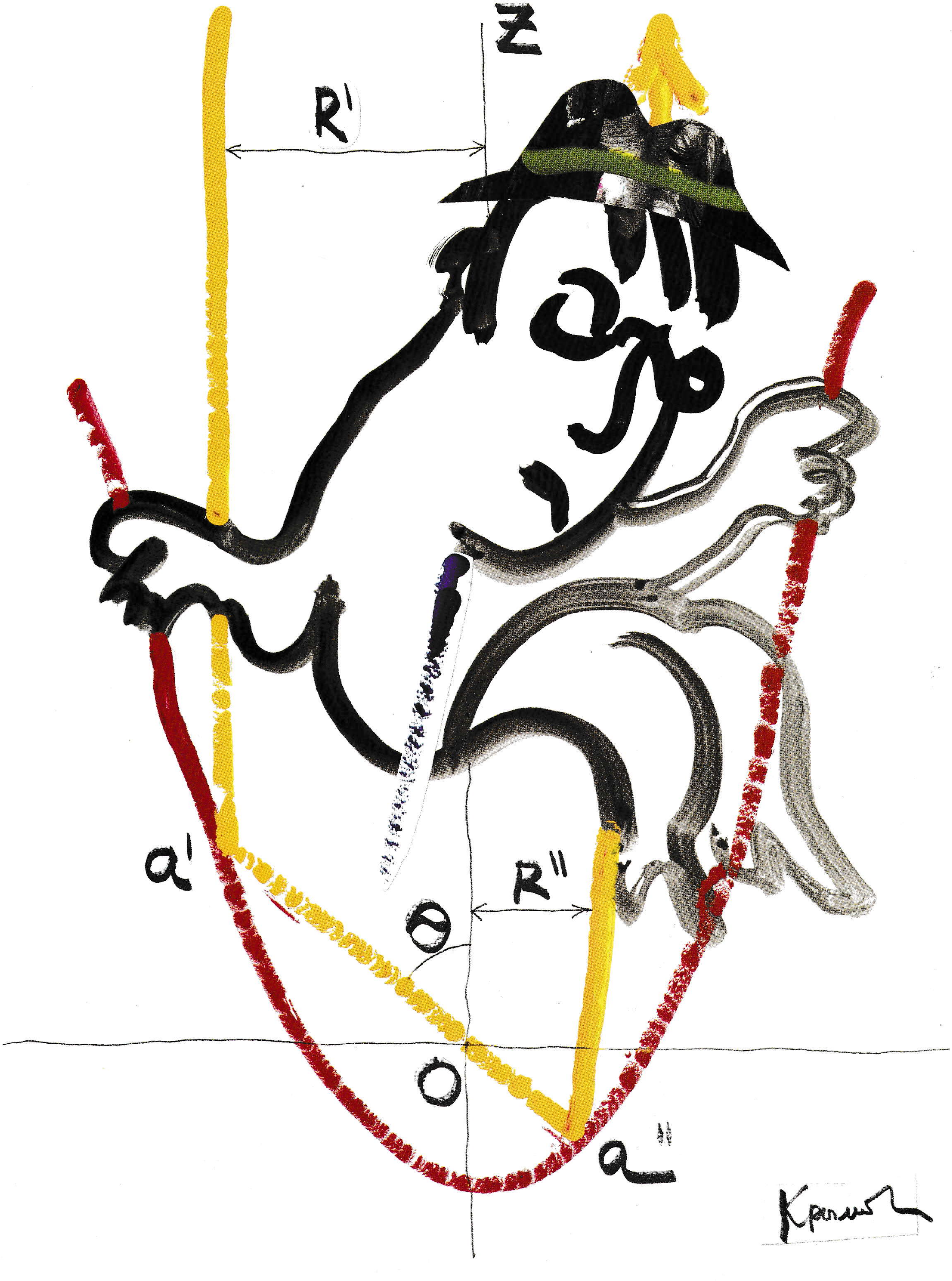


Figure 2

(a) Intensity distribution over the cross section of the outgoing beam; (b) plot of the intensity as a function of distance from the beam's axis.

Art by Dmitry Krymov


Z
R'
a'
R''
θ
O
a''
Krasnik

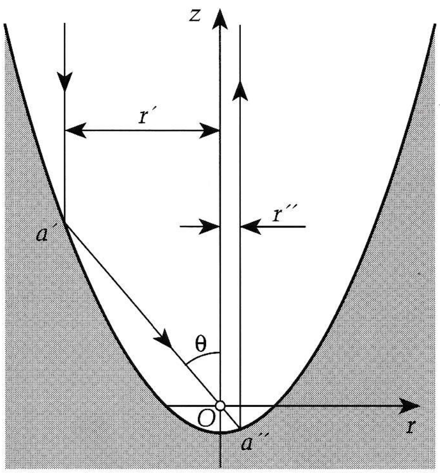


Figure 3

At first we explain this fact as follows. Let OZ be the symmetry axis, and let r', r'' be the distances between the axis and the corresponding incoming and outgoing rays (fig. 3). The farther from the axis OZ the incident beam falls, the nearer to the axis it will exit—that is, the larger r' is, the smaller r'' is. Thus, the energy of the peripheral areas of the incoming beam is shifted toward the axis of the outgoing beam. If the incident beam is rather wide (that is, the radius of the mirror is large), this phenomenon leads to a substantial concentration of energy near the axis. As a result, the energy emerges from the mirror practically in the form of a narrow pencil of light, and the wider the incoming beam, the narrower the light pencil.

Thus, we come to the following conclusion: a paraboloid can concentrate energy not only at the focus but also along the symmetry axis.

We can calculate this effect. Let's consider the parabola whose revolution about the axis OZ forms the surface of the paraboloid. If the origin of the coordinates is placed at the focus, the equation for the parabola will be

$$z = \frac{r^2}{4f} - f, \qquad (1)$$

where f is the focal length, r and z are the abscissa and ordinate of an arbitrary point of the parabola. The incoming beam meeting the parabola at the point a' at a distance r' from the symmetry axis will pass through the focus after reflection and again strike the parabola at the point a'' at a distance r'' from the axis OZ (fig. 3). Let's denote by θ the angle between OZ and the intermediate segment $a'a''$ of our beam. The equation for the line $a'a''$ is as follows:

$$z = -r \cot\theta. \qquad (2)$$

The points a' and a'' belong simultaneously to the parabola defined by equation (1) and the line defined by equation (2). Thus, we can equate the right-hand sides of equations (1) and (2):

$$r^2 + 4fr \cot\theta - 4f^2 = 0. \qquad (3)$$

Solving this equation we get

$$r_1 = -2f \cot\frac{\theta}{2}, \quad r_2 = 2f \tan\frac{\theta}{2}. \qquad (4)$$

The roots r_1 and r_2 are the values of the coordinate r for the points a' and a''. Thus, $r' = |r_1|$, $r'' = r_2$. It follows from equation (4) that

$$r'r'' = 4f^2. \qquad (5)$$

Equation (5) can be obtained by means of the Viète theorem without solving equation (3).This result supports the aforementioned qualitative conclusion that the distances of the incoming and outgoing rays from the axis OZ are inversely proportional.

Equation (5) can be "read" in the following way. Consider the plane A perpendicular to the axis OZ. Let every beam, either coming into or going out of the paraboloid parallel to OZ, cut this plane and leave a trace on it in the form of a dot (fig. 4). According to equation (5), double reflection from the paraboloid corresponds to a conversion of the points in plane A that transforms any segment $Oa' = |r_1|$ into another segment $Oa'' = |4f^2/r_1|$. This transformation was the centerpiece of a previous article in *Quantum*, "Making the Crooked Straight" (November/December 1990). It's known as the inversion relative to a circle of radius $2f$ with center O.[2] That article also described devices that put this transformation to practical use—the "inversors" of Peaucellier and Hart.

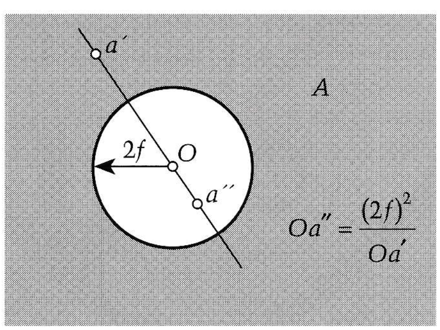


Figure 4

Equations (4) and (5) contain all the information we need about the redistribution of energy in a beam of incident light. Let's try to extract it.

Within the incident beam let's single out a narrow ring of width $\Delta r'$ and inner radius r' ($\Delta r' << r'$) lying in the plane perpendicular to the axis of the paraboloid. The rays passing through the ring travel at a distance from the axis within the range $[r', r' + \Delta r']$—see figure 5. After double reflection off the mirror, these rays intersect the plane A at points lying inside the ring of width $\Delta r''$ with outer radius r''. That is, the outer radius $r' + \Delta r'$ of the "incident" ring corresponds to the inner radius $r'' - \Delta r''$ of the "reflected" ring. According to equation (5), the product of these radii is $4f^2$ (we'll dispense with the absolute value signs and take r to be the numerical value of

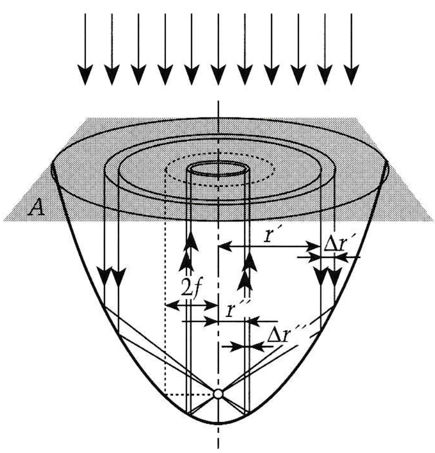


Figure 5

[2]Actually, in this case we don't have a "pure" inversion, but rather a combination of two transformations: an inversion and a symmetry transformation relative to the point O.

the distance). That is,

$$(r' + \Delta r')(r'' - \Delta r'') = r'r'' + r''\Delta r' - r'\Delta r'' - \Delta r'\Delta r'' = 4f^2.$$

Taking into account that $r'r'' = 4f^2$, we get

$$r''\Delta r' - r'\Delta r'' - \Delta r'\Delta r'' = 0.$$

Since $\Delta r' << r'$ and $\Delta r'' << r''$, the last member in the equation can be neglected and we can write

$$r''\Delta r' - r'\Delta r'' = 0.$$

It follows directly that

$$\frac{\Delta r''}{\Delta r'} = \frac{r''}{r'} = \frac{4f^2}{r'^2},$$

or

$$\Delta r'' = \Delta r' \frac{4f^2}{r'^2}. \quad (6)$$

Thus, the energy that in the incident beam passes through a ring of width $\Delta r'$ after double reflection of the rays off the mirror passes through a ring of width $\Delta r'' < \Delta r'$. This means that the amount of energy passing through a unit area per unit time (this characteristic value is known as the beam's intensity) is greater in the "reflected" ring than in the "incident" ring. We considered the case when $r' > 2f$. If $r' < 2f$, the reflected ring will be larger than the incident ring, which means that the intensity of the reflected light is less than that of the incident light.

So, the reflection of light from the paraboloid mirror results in a redistribution of energy. Let's find how the intensity changes with distance from the symmetry axis of the paraboloid. Let the intensity of the incident light be J_0. Then the amount of energy passing through a ring of width $\Delta r'$ per unit time (that is, the light flux through the ring) is the product of J_0 and the area of the ring—that is,

$$\Delta\Phi = J_0 2\pi r'\Delta r',$$

where r' is the inner radius of the ring.[3] After double reflection from the mirror all of this flux will come out through an area $2\pi r''\Delta r''$. So the intensity of light in the reflected beam will be

$$J = \frac{\Delta\Phi}{2\pi r''\Delta r''} = J_0 \frac{r'\Delta r'}{r''\Delta r''} = J_0 \frac{16f^4}{r''^4}. \quad (7)$$

Equation (7) describes the sharp increase in the intensity with a decrease in r''—that is, as the distance from the symmetry axis of the paraboloid decreases.

However, this result does not mean that the intensity becomes infinite at the axis itself. After all, the rays that emerge closest to the symmetry axis are those that were on the periphery of the incident beam. Clearly the maximum distance of the incident ray r'_{max} from the symmetry axis is equal to the radius of the mirror. This corresponds to the minimum distance r''_{min} in the exiting beam. The intensity of the reflected light increases according to equation (7) when r'' decreases to r''_{min}, but it becomes practically zero when $r'' < r''_{min}$.[4] This is why there's a trough in the intensity curve near the axis in figure 2. However, the maximum intensity, which occurs right near the trough, can be many times the intensity J_0 of the incident beam.

We extracted all this information from one simple formula—equation (5)!

Now let's consider the equations (4). They describe the passage of the rays "inside" the paraboloid between two reflections. These rays, converging at the focus and diverging then from it, form a kind of point source in the focus. This source differs, however, from a real one. It emits light non-isotropically—that is, the intensity of the emitted light changes with the beam's direction.

One can't help thinking that there is some mystical beauty in all these properties of the paraboloid!

I should point out that the conclusions we've drawn are correct only within the framework of idealized considerations that do not strictly correspond to reality. We spoke about the paraboloid as a mathematically ideal surface. A real mirror is only an approximation of such an ideal. Its surface is not strictly symmetrical; it has some spots that aren't quite as smooth as others; and so on.

Also, we used the concept of a beam of parallel rays. In reality any pencil of light propagates inside some solid angle that is not zero. We can speak only of a practically parallel beam—that is, a beam that diverges so little that we can neglect this divergence in a particular stretch of space. For example, the light pencil falling on the paraboloid from a star can be considered practically parallel. But even in this case the emerging beam will be distorted due to so-called diffractive divergence caused by the wave nature of light. And the narrower this beam (that is, the larger the mirror's diameter), the greater the divergence of the rays. Quantitatively the diffractive divergence is characterized by the angle of deviation of a ray from the direction of light propagation. The order of magnitude of this angle is

$$\alpha \cong \frac{\lambda}{r''_{min}},$$

where λ is the wavelength of the incident light. Thus, the thinner and more concentrated the light pencil is, the more it diverges. Let's evaluate the distance z where the increase of the beam's radius caused by the divergence becomes equal to the initial radius of the beam (up to this distance the beam can be considered practically parallel). We find that

$$z \cong \frac{r''_{min}}{\tan\alpha} \cong \frac{r''_{min}}{\alpha} \cong \frac{\left(r''_{min}\right)^2}{\lambda}.$$

According to this formula, the thinner and more intense the beam (all

CONTINUED ON PAGE 55

[3]The area of the ring delimited by the circles with radii r and $r + \Delta r$ is $S = \pi(r + \Delta r)^2 - \pi r^2 = 2\pi r\Delta r + \pi(\Delta r)^2$. If $\Delta r << r$, we can neglect the term with $(\Delta r)^2$ and consider that $S = 2\pi r\Delta r$.

[4]In reality the distribution of the intensity in the outgoing light is more complicated due to diffraction at the mirror's edge.

PROFILE

For the love of her subject

Marina Ratner's energetic path in mathematics

by Julia Angwin

LUCKILY FOR HER, THE YEAR that Marina Ratner applied for college in the Soviet Union coincided with Nikita Khrushchev's denouncement of Stalin. It was 1956, and the young Jewish girl was trying to get into Moscow State University—the Harvard of the USSR. The doors to the university, normally frozen shut to Jews, opened a crack that year during the political thaw . . . and Marina Ratner slipped in.

The training normally denied to people of her religion set Ratner on her course as a mathematician. She is now a professor at the University of California, Berkeley, and recently won a $25,000 award from the National Academy of Sciences.

The award recognizes her "striking proof" of a fundamentally important theorem that originated in number theory and is called the Raghunathan conjectures. Previous attempts to prove the conjectures for some particular cases were very intricate and provided little insight into what was going on. Ratner tackled the proof using a branch of mathematics called ergodic theory that originated in the study of thermodynamics. The roots of the name are the Greek words *erg* (energy) and *hodos* (path). This area of mathematics is also closely related to probability theory and statistics. Ratner's knowledge of ergodic theory helped her come up with the ideas needed to prove the conjectures. Her solution led to further developments in number theory and the theory of quadratic forms.

Ratner traces her interest in math partly to the tutelage of a particular high school teacher, although as the child of scientists she had always excelled in the subject.

"I really got a lot out of that one teacher," she recalled. "He was very tough and he was very difficult but interesting."

About her life as a woman in mathematics, Ratner said: "I don't

believe when they say that women have different brains, or that women aren't treated the same as men or boys. In my life I did not encounter any gender discrimination."

The most important thing for students is to love their subject, she said. That is how she won her teacher's love.

"He would sometimes even ask me to help in grading the test that he had given the class," she recalled.

He assigned the students difficult problems, teaching them to work in three dimensions as well as in plane geometry.

"Even students whom he gave a C, they did very well in college tests," Ratner recalled.

At Moscow State University, Ratner honed her mathematical skills. For four years she studied mostly math, peppering her curriculum with only physics classes and the required Marxism and Communist Party history courses.

After that, she took a four-year hiatus, working in a statistics research group. She also gave birth to her daughter, Anna.

When Anna was three years old, her mother went back to school to get her Ph.D. in mathematics. Russian students could stay in graduate school a maximum of three years. If their doctorate took any longer, they had to complete it on their own. So, by 1969, Ratner had her doctorate and was looking for work.

She taught for a while at a technical engineering school, but quickly decided to emigrate to Israel. She applied for a visa and was immediately fired from her job.

"It was considered unpatriotic and they called us traitors just because we wanted to emigrate," she recalled.

Fortunately her visa took only three months to arrive, and she quickly joined her relatives in Israel. After a few years' teaching at the Hebrew University of Jerusalem, she was hired by the University of California at Berkeley.

"I liked America from the very first day, despite the many things that are not good here," she said. "I think that not all Americans realize how great a country it is."

This country is undoubtedly the richer for the influx of such talented scholars as Marina Ratner. The award she received is merely a sign of what she had already given her adopted land.

Bulletin Board

Quantum and Quantumites win awards

The Professional and Scholarly Publishing Division of the Association of American Publishers (AAP) named *Quantum* the award winner for Excellence in Design and Production in the journals category for 1993. As part of its commitment to excellence, this division of AAP sponsors a prestigious annual awards program that acknowledges and promotes outstanding examples of professional and scholarly publishing. The National Science Teachers Association (NSTA) and Springer-Verlag New York, Inc., were joint recipients of the award in the AAP's Eighteenth Annual Awards Competition.

Several *Quantum* covers have received special notice over the past several years. Artist Leonid Tishkov's covers for the November/December 1991 and March/April 1992 issues won awards for excellence from *Print* magazine and were reprinted in its regional design annual, a showcase of the best in illustration and design.

Staff artist Sergey Ivanov was an award winner in the Magazine Cover category in the Creativity '93 competition. His cover for the July/August 1993 issue of *Quantum* will appear in the *Creativity '93 Annual.* Sergey also cited by the Educational Press Association of America (EdPress) for his work on *The Pillbug Project*, an NSTA publication. EdPress honored *Quantum* publisher Bill G. Aldridge for his Publisher's Page editorial "Photosynthesism" in the July/August 1992 issue of *Quantum.*

Our congratulations to the artists, to Bill Aldridge, and to our colleagues at Springer-Verlag.

New IMAX film

If your vacation plans include a visit to Washington, D.C., be sure and stop by the National Air and Space Museum and take in the new IMAX film, "Destiny in Space." This summer marks the 25th anniversary of the Apollo 11 moon landing, and as we look back at that pivotal event and take stock of what has been achieved since then, we also look for signs of what lies ahead in space exploration. "Destiny" provides insights into how our current space activities are preparing us for the future.

The 40-minute film is the third in a trilogy of movies shot by astronauts aboard space shuttles. It explains the exploratory background laid by life sciences research, robotic planetary missions to Venus and Mars, investigations involving the Hubble Space Telescope, and other studies. "Star Trek" actor Leonard Nimoy narrates the film.

Guide to scholarships

Students who are considering careers in math or science, or those who are already well on their way in these fields, would do well to consult *The Prentice Hall Guide to Scholarships and Fellowships for Math and Science Students* by Mark Kantrowitz and Joann P. DiGennaro. According to the publisher, this is the "first and only resource to focus on the more than 250 scholarships and fellowships available to math and science students at the high school, undergraduate, and graduate levels." It also provides the latest information on more than 80 contests and competitions, internships, summer employment offerings, and opportunities to study abroad.

The guide includes information on financial aid programs that span the range of careers open to students in science, math, and engineering, from acoustics, biotechnology, and computer science to meteorology, physics, and zoology. It describes programs directed toward female and minority students as well as programs of a more general nature that do not restrict the student's field of study.

The book provides guidance on such topics as how to choose an undergraduate school that suits your needs (evaluating everything from courses of study and instructors to social atmosphere and extracurricular activities); how to uncover all possible sources of financial aid; and how to improve your chances of being accepted by the graduate school of your choice.

The book is capped off with an annotated bibliography of additional sources of academic and career information of potential interest to science and math students. (325 pages, $19.95 paper—ISBN 0-13-045345-5, $29.95 cloth—ISBN 0-13-045337-4)

Talk back to us

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

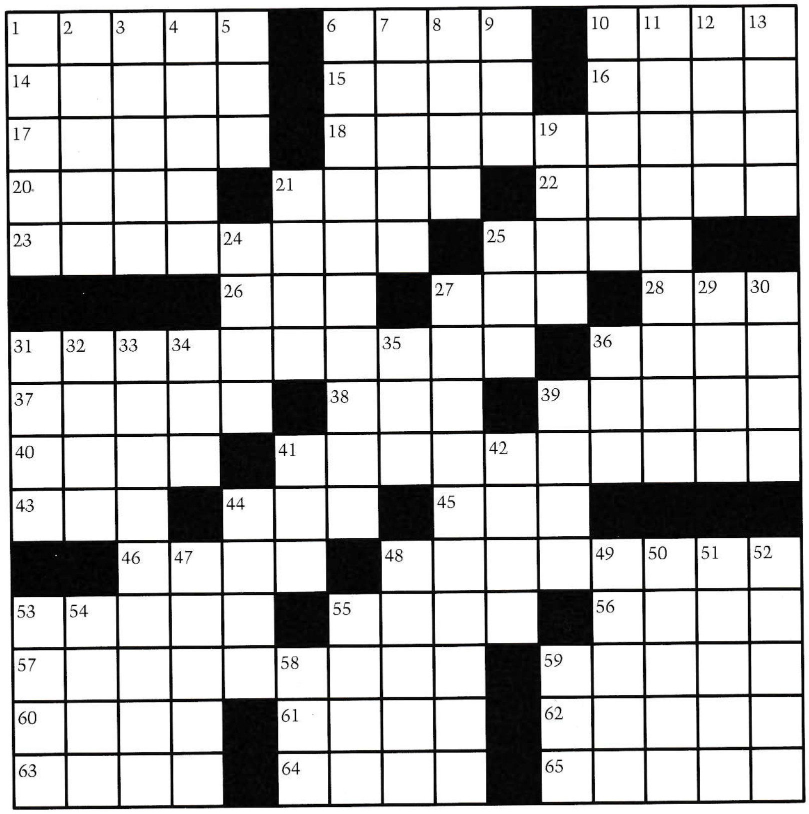
or

zip us an electron or two
at quantum@nsta.org

Volume 4 (1993–94)

Criss × cross science

by David R. Martin



Across

1 Behind
6 Experimental results
10 Large plasma ball
14 Milton ___
15 Mild oath
16 Albacore or bluefin
17 Plait
18 Oxide of cerium or erbium, e.g.
20 Discharge
21 Former Yugoslavian leader
22 Emotional emanations (slang)
23 Oscilloscope pattern
25 TV's Jay ___
26 Bank account
27 Atmosphere
28 Plastic ingredient: abbr.
31 Branch of mechanics
36 Leguminous plant
37 Seize illegally
38 Fish disease
39 ___ ray (high energy photon)
40 Coin
41 Not critically damped
43 Employ
44 Chemical suffix
45 Mixture of hydrocarbons
46 Speed
48 Heating ore mixtures
53 Large nail
55 Laser ___
56 Nonpareil
57 ___ line (on a tidal map)
59 French river
60 Birth-control advocate ___ Guttmacher
61 "The King ___"
62 Attend to
63 Courageous
64 Pear
65 Unit of length

Down

1 Abbots
2 Italian physicist Enrico ___
3 Comet tail
4 A type
5 ___ giant star
6 Instantaneous rate of change
7 Ornamental chalcedony
8 Poi source
9 Summer drink
10 Prepare a slide
11 Turbine driven machine
12 Poker stake
13 Arena noises
19 For ___ more
21 Scarlet's home
24 Weakling
25 ___ pendens
27 Like some lenses and prisms
29 Arrived
30 Nail
31 African antelope
32 Egyptian goddess
33 Not analytical
34 Poet's before
35 Dry ___
36 Group of whales
39 Type of bladder
41 Single
42 Per ___
44 Fusion project: abbr.
47 Insulation fiber
48 Dispatches
49 "___ a fine lady"
50 Like a noble gas
51 ___ cell (neuron)
52 Feminist Germaine ___
53 Wound cover
54 Rich man's sport
55 Seismologist ___ Gutenberg (1889-1960)
58 Mortar mixer
59 Belief

SOLUTION IN THE NEXT ISSUE

SOLUTION TO THE MAY/JUNE PUZZLE

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

ANSWERS, HINTS & SOLUTIONS

Math

M116

Let n be the number of sides of the given polygon (of course, $n \geq 4$). Then the number of its diagonals is $n(n-3)/2$ (because there are $n-3$ diagonals issuing from each of its n vertices, and this accounts twice for each diagonal). So we have to prove that

$$\frac{s}{n} < \frac{d}{n(n-3)/2},$$

where s and d are the sums of the lengths of all the polygon's sides and diagonals, respectively.

Consider two nonadjacent sides AB and CD, and diagonals AC and BD joining their endpoints and intersecting at O (fig. 1). Since $AB < AO + OB$, $CD < CO + OD$, we have $AB + CD < (AO + OC) + (BO + OD) = AC + BD$.

If we write out such inequalities for all pairs of nonadjacent sides and add them up, we'll get $(n-3)s$ on the left side (because each side enters $n-3$ pairs) and $2d$ on the right side (because each diagonal occurs in two inequalities—it joins two pairs of nonadjacent sides in the position of figure 1). So $(n-3)s < 2d$, which is equivalent to what we set out to prove.

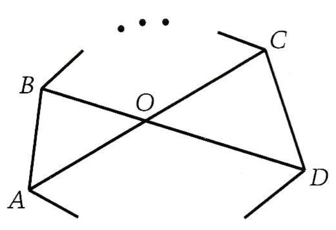


Figure 1

M117

The answer is 2^{n-1}. Consider any of the permutations in question. Suppose the number in the first place (from the left) is k. If $k > 1$, then the numbers $k-1, k-2, \ldots, 1$ will be met just in this (decreasing) order as we move from left to right. Indeed, the number 1 must have 2 on its left side, the number 2 must have 1 or 3—that is, 3—on the left, 3 must have 2 or 4—that is, 4—on the left, and so on. Similarly, for $k < n$ the numbers $k+1, k+2, \ldots, n$ must be arranged in increasing order, because $n-1$ must be to the left of n, then $n-2$ to the left of $n-1$, and so on. Therefore, any of the permutations considered in the problem is uniquely determined by the set of places occupied by the numbers $1, 2, \ldots, k-1$ (there may be no such places if $k = 1$—that is, for the identity permutation $1, 2, \ldots, n$): we have to arrange these numbers in these places in decreasing order and the remaining numbers in the remaining places in increasing order.

It's not hard to see that the number of such sets is simply the number of all subsets of the set of all the $n-1$ places except the first, and is equal to 2^{n-1}.

M118

We'll use induction over $m + n$ to prove a slightly generalized statement with *not less* than k greatest numbers underlined in each column, and *not less* than l in each row (this will be more convenient for inductive reasoning).

Obviously, for $m = n = k = l = 1$ the statement is true: $kl = 1$ number (the only one) will be underlined twice. Let's now show that the statement for an $m \times n$ array can be reduced to the case of an $(m-1) \times n$ or $m \times (n-1)$ array.

If all the underlined numbers in an $m \times n$ array are underlined twice, then there are no less than kl of them. Otherwise, let a be the greatest of the numbers underlined once. It's either one of the k greatest numbers in its column or one of the l greatest numbers in its row. Assume it was underlined "along a column." Then the l greatest numbers in its row are greater than a and so are underlined twice. Cancel out this row. We get an $(m-1) \times n$ array, in which at least l greatest numbers are underlined in each line and at least $k-1$ numbers in each column. The induction hypothesis implies that at least $(k-1)l$ numbers in this reduced array are underlined twice. The same numbers are underlined twice in the big array; together with at least l numbers underlined twice in the line that was deleted, this makes at least $(k-1)l + l = kl$ numbers, completing the proof.

M119

Let's first prove that $PA_1 = PA_2 = \ldots = PA_n$. Suppose this is not true. Choose the shortest and the longest of these edges, PA_s and PA_l, respectively. Now draw an angle BAC congruent to angles $PA_1A_2, \ldots, PA_nA_1$, and mark off the segment $AB = A_1A_2 = \ldots = A_nA_1$ on one of its sides, and segments $AC = A_{s-1}P$ and $AD = A_{l-1}P$ on the other side. (Figure 2 shows the

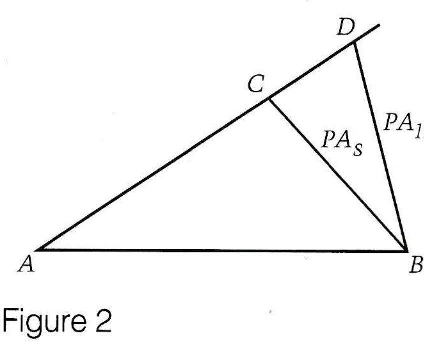


Figure 2

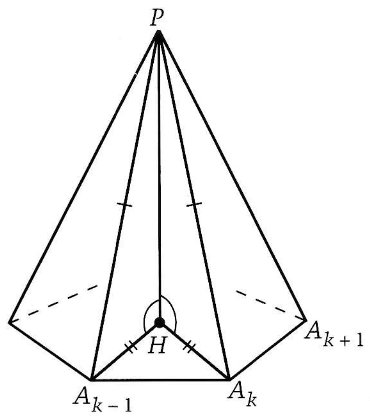


Figure 3

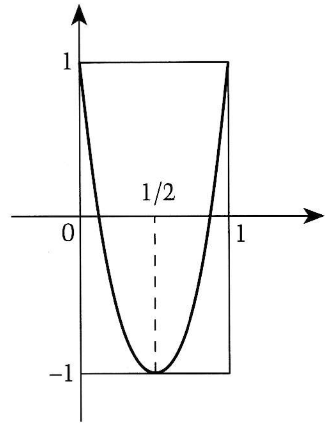


Figure 4

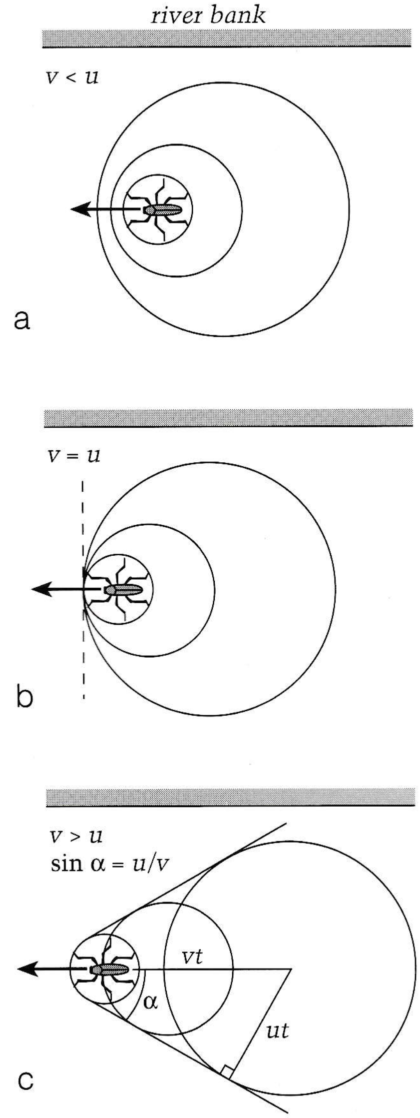


Figure 5

case where $AD < AC$. The same argument will hold if $AD > AC$. Of course, when $s = 1$ or l, A_{s-1} or A_{l-1} should be replaced by A_n.) Triangles ABC and ABD are congruent to the faces $A_{s-1}A_sP$ and $A_{l-1}A_lP$ of the pyramid, respectively, so $BC = PA_s$ and $BD = PA_l$. By the Triangle Inequality, $BD - BC < DC = |AD - AC|$—that is, $PA_l - PA_s < |PA_{l-1} - PA_{s-1}| \le PA_l - PA_s$, because PA_l is the longest and PA_s the shortest of the edges PA_k—and this is a contradiction that proves that all these edges are the same length. Now drop the height PH of the pyramid (fig. 3). By the Pythagorean Theorem, $HA_k^2 = PA_k^2 - PH^2$, so $HA_1 = HA_2 = \ldots = HA_n$. Therefore, the vertices $A_1, \ldots, A_n$ all lie on a circle with center H and so divide it into equal arcs (subtended by equal chords A_kA_{k+1}). This means that $A_1A_2 \ldots A_n$ is a regular n-gon, and the base H of the pyramid's height is its center, so we're done.

M120

The answer is $a = 8$. If $f(x) = Ax^2 + Bx + C$, then $f'(1) = 2A + B$. Now

$$f(0) = C,$$
$$f(1/2) = A/4 + B/2 + C,$$
$$f(1) = A + B + C,$$

so we have

$$|f'(1)| = |2A + B|$$
$$= |3f(1) - 4f(1/2) + f(0)|$$
$$\le 3|f(1)| + 4|f(1/2)| + |f(0)|$$
$$\le 8.$$

So $a \le 8$. On the other hand, the above inequalities turn into exact equalities if $f(1) = -f(1/2) = f(0) = 1$, that is, for $f_0(x) = 8x^2 - 8x + 1$ (see the graph in figure 4). The function $f_0(x)$ satisfies the inequality $|f_0(x)| \le 1$ for $0 \le x \le 1$ (its graph is a parabola with the vertex at $x = 1/2$), so the number 8 in our estimate for $|f'(1)|$ cannot be decreased, and therefore $a = 8$.

This solution was based on the fact that a polynomial of degree n is uniquely determined by its values at $n + 1$ points. In our case, $n = 2$; try to solve this problem for cubic polynomials. It's worth noting that for an arbitrary degree n the best estimate in this problem is given by the so-called *Chebyshev polynomials*, which we intend to discuss in detail in upcoming issues of *Quantum*.

Physics

P116

Three characteristic cases are possible: (a) the beetle's velocity v relative to the water is less than the velocity of the surface waves u—that is, $v < u$; (b) $v = u$; (c) $v > u$. Each case is illustrated in figure 5. Figure 5c looks like the shock wave produced by a supersonic aircraft, or like Cherenkov radiation, which arises when electrons travel faster than light in a particular medium (though not, of course, faster than light in a vacuum).

P117

Denote the normal force as **N** (the vector is directed toward the center of the dome), and the force of friction as f. To find the correct direction of this force, we need to see where the support point moves if the friction disappears. Let's write the equilibrium equation for the rod (fig. 6):

$$2N \cos \alpha - 2f \sin \alpha + mg = 0.$$

If the rod were initially pressed into place more strongly, N and f would increase, but mg would not change. As we are interested in the minimal

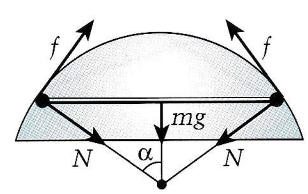


Figure 6

possible value of the coefficient of friction, we assume that the rod has been pressed very strongly and, hence, the force mg can be neglected. Then

$$N\cos\alpha = f\sin\alpha \le \mu N\sin\alpha,$$

or

$$\mu \ge \cot\alpha = \frac{\sqrt{R^2 - l^2/4}}{l/2} = \sqrt{\frac{4R^2}{l^2} - 1}.$$

We've obtained the answer, but some comments are in order. If $l > 2R$, we can't insert the rod into the dome. If, on the other hand, the rod is too small, then the corresponding value for μ will be too large and will make no sense. Finally, rubber is a substance with very complicated properties; its interaction with a solid surface is not adequately described by the laws of dry friction (but that's another story).

P118

The conduction of heat from the inner shell of the thermos (that is, from the tea) is performed by molecules of gas between the inner and outer shells. That's why the space between the shells was pumped out: to prevent the gas that remains from forming a continuous heat-conducting medium—that is, to make the molecules move between the shells without colliding with one another.

Let's assume that after a collision with a shell, a molecule on average has an energy proportional to the temperature of the shell. Thus, at the outer shell the kinetic energy of the translational movement of a molecule is $K_1 \sim \frac{3}{2}kT_r$, where T_r is room temperature; at the inner shell the corresponding value is $K_2 \sim \frac{3}{2}kT_t$, where T_t is the temperature of the tea. Although this temperature changes with time, the range of its variation is not large: $\Delta T = 363$ K $- 343$ K $= 20$ K. So for the purpose of our estimates we can assume that $K_1 \sim \frac{3}{2}kT_s$ where $T_s = 353$ K is the average temperature of the hot shell.

Consequently, after a collision with the hot shell, each molecule "takes" energy from the tea that is equal to

$$\Delta K \sim \tfrac{3}{2}k(T_s - T_r).$$

The number of molecules colliding with the hot shell per unit time is proportional to $\frac{1}{2}nvS$, where n is the density of the gas molecules and v is the average magnitude of the velocity projected perpendicular to the wall. The number of molecules that collide with the wall during time Δt is

$$N \sim \tfrac{1}{2}nvS\Delta t.$$

These molecules take from the tea the energy

$$\Delta E = N\Delta K \sim \tfrac{3}{4}nvSk(T_s - T_r)\Delta t.$$

This energy is equal to the change ΔU in the internal energy of the tea in the period Δt—that is, $\Delta E = \Delta U$. Insofar as

$$\Delta U = Mc\Delta T,$$

where ΔT is the change in the temperature of the tea, then

$$\tfrac{3}{4}nvSk(T_s - T_r)\Delta t \sim Mc\Delta T.$$

From this we get

$$\Delta t \sim \frac{4Mc\Delta T}{3nvSk(T_s - T_r)}.$$

Taking into account that $v \sim \sqrt{RT_r/\mu}$, where μ is the molar mass of the gas, and $n = P/kT_r$ (which follows from the Clapeyron–Clausius equation), we finally obtain

$$\Delta t \sim \frac{4Mc\Delta T\sqrt{\mu T_r}}{3P(T_s - T_r)S\sqrt{R}}.$$

Substituting numerical data ($M = 1$ kg, $\Delta T = 20$ K, $T_r = 293$ K, $P = 1$ Pa, $\mu = 29 \cdot 10^{-3}$ kg/mol, and so on) results in

$$\Delta t \sim 3 \cdot 10^4 \text{ s} \sim 8 \text{ h}.$$

P119

At an arbitrary point A at a distance R from the wire, the velocity of the particle is directed at a small angle α with the x-axis, so

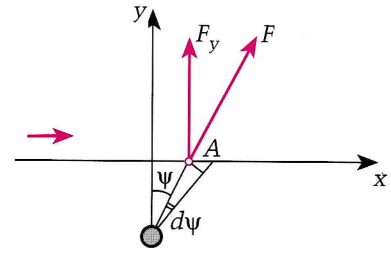


Figure 7

$$\alpha = \frac{v_y}{v_x},$$

where v_y is the vertical projection of the velocity and $v_x = \sqrt{2K/m}$ is the horizontal projection.

Let's write down Newton's second law for the y-component (fig. 7):

$$F_y dt = m dv_y,$$

where

$$F_y = eE\cos\psi = \frac{e\lambda\cos\psi}{2\pi\varepsilon_0 R}.$$

The small time period dt can be derived from the formula $v_x = dx/dt$:

$$dt = \frac{dx}{v_x} = \frac{R\,d\psi}{v_x\cos\psi}.$$

During this time the vertical projection changes by

$$dv_y = \frac{F_y}{m}dt = \frac{e\lambda}{2\pi\varepsilon_0 m v_x}d\psi.$$

The total projection of the velocity on the y-axis is composed of small increments:

$$v_y = \int_{-\pi/2}^{\pi/2} dv_y = \frac{e\lambda}{2\varepsilon_0 m v_x}.$$

Thus, the angle we're seeking is

$$\alpha = \frac{v_y}{v_x} = \frac{e\lambda}{2\varepsilon_0 m v_x^2} = \frac{e\lambda}{4\varepsilon_0 K}.$$

P120

Two cases are possible. (1) The lens is convergent. Drawing the

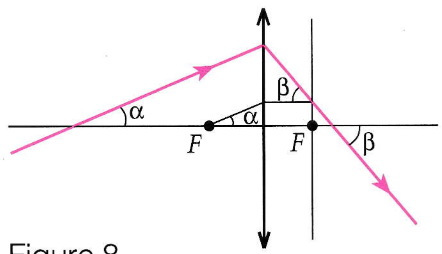


Figure 8

light path through the lens (fig. 8) yields the following relationship:

$$d \tan \alpha = F \tan \alpha + F \tan \beta.$$

From this it follows that

$$F = \frac{d}{1 + \tan\beta/\tan\alpha} = 4 \text{ cm}.$$

(2) The lens is divergent. From figure 9 we have

$$F \tan \beta = F \tan \alpha + d \tan \alpha,$$

from which we get

$$F = \frac{d}{\tan\beta/\tan\alpha - 1} = 12 \text{ cm}$$

—that is, the beam exits from the focus of the lens.

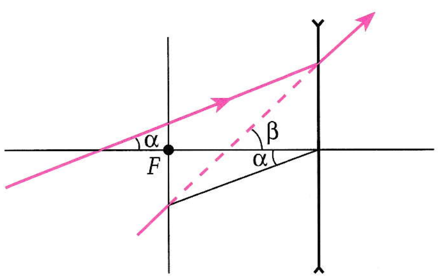


Figure 9

Brainteasers

B116

Suppose originally the band formed an $n \times n$ square, so that the number of musicians is n^2. By the statement of the problem, n^2 must be divisible by $n + 5$ (so that it's possible to form $n + 5$ rows). Since $n^2 = (n + 5)(n - 5) + 25$, this means that 25 is divisible by $n + 5$. The only divisor of 25 greater than 5 is 25 itself, so $n + 5 = 25$, $n = 20$, and the answer is $n^2 = 400$. (V. Dubrovsky)

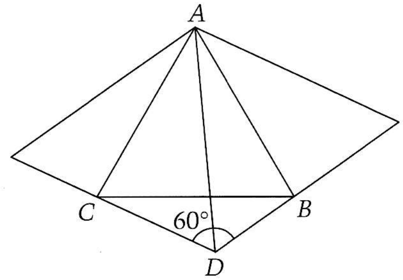


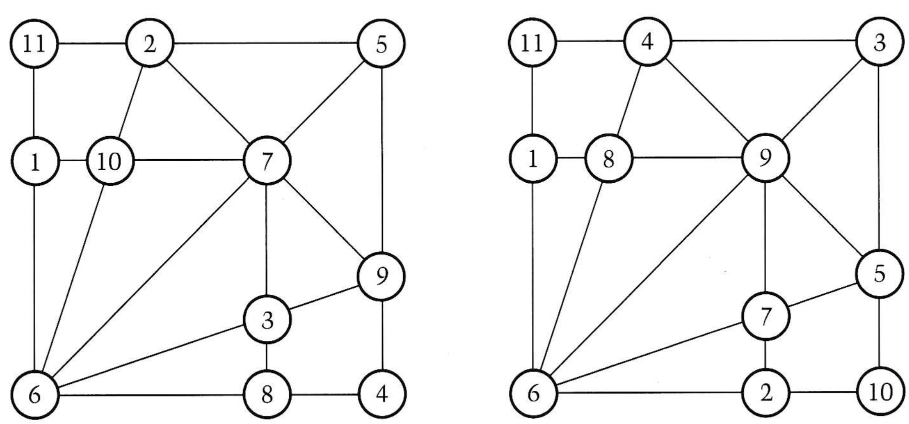
Figure 10

B117

From the properties of a rhombus it is clear that $\angle ACD = \angle BDA = 60°$ (fig. 10). Suppose angle ACB were 60°. Then quadrilateral $ACDB$ can be inscribed in a circle (since $\angle ACD = \angle ADB$), so angle CAB is supplementary to angle CDB, and is 60° as well. This makes triangle ABC equilateral. A similar proof holds if angle ABC is assumed to be 60°. Finally, if the 60° angle is supplementary to angle ADB, so again quadrilateral $ACBD$ is cyclic. Hence $\angle ACB = \angle ADB = 60°$, and triangle ABC is once more equilateral.

B118

Turn on one of the switches, wait a while, turn it off, then turn on a second switch and enter the room with the lamps. The lamp that's on is connected to the second switch, one of the other two lamps is hot and is connected to the first switch, and the remaining lamp is connected to the third switch.



a

b

Figure 11

B119

Let s be the sum of the three numbers along any of the lines and a the number in the left bottom corner. If we add together the five triple sums along the lines through a, we'll get $(1 + 2 + \ldots + 11) + 4a = 66 + 4a = 5s$. Now consider three horizontal lines and the line through the two numbers not on these lines. Adding the numbers along these four lines, we get $4s = 66 + a$. From these two equations we find $a = 6$, $s = 18$. One of the lines through $a = 6$ contains the numbers 1 and 11. Other lines that may pass through 1 can contain 8 and 9 or 7 and 10, while lines through 11 can contain 2, 5 or 3, 4. So a line through 1 can't intersect a line through 11. This leaves only the circles on the left or bottom sides of the square as possible locations of 1 and 11. Writing 11 in the top left corner, by trial and error we find the two solutions in figure 11. If we simultaneously exchange the pairs of numbers other than 6 on all lines through 6 (for instance, 1 and 11, 10 and 2, 7 and 5, and so on in figure 11a), we'll get two other solutions. (Using the fact that $s = 18$ and $a = 6$, we can make sure that this operation always transforms a solution into another solution.) And, finally, these four solutions can be reflected about the diagonal line to give four more solutions. So the total number of solutions is 8. (V. Dubrovsky)

B120

By the statement of the problem, the two odd numbers in question are different—otherwise, their product would be equal to the square of their half-sum. So the smallest possible product is $11 \cdot 13 = 143$; the next one is $11 \cdot 15 = 165$. Since there are no pyramids as high as 165 m, the numbers are 11 and 13, and the height of the pyramid is greater than 143 m but less than $[(11 + 13)/2]^2 = 144$ m. There are two pyramids in Egypt higher than 143 m: the pyramid of Khufu (Cheops) at 146.6 m and that of Khafre (Chephren) at 143.5 m. So the pyramid was built for Chephren.

Tilings

1. For figure 3a in the article the answer is $1/(mn + 1)$; for figure 3b it's $1/(mn - 1)$. In both cases the answer is obtained by extending the grid of the small parallelograms and counting nodes inside the big parallelogram.

2. It can be proved that the points A, B, C and the intersection points of AA_1, BB_1, and CC_1 can be considered nodes of a parallelogram grid (fig. 12). Then the area of ABC equals $3 + 3 \cdot 1/6 = 3\frac{1}{2}$ times the area of a grid parallelogram, while the area of the small (shaded) triangle is 1/2 of this unit. So the answer is $1/2 \div 3\frac{1}{2} = \frac{1}{7}$.

The existence of the parallelogram grid in the solution above can be derived, for example, from the fact that the segments AA_1, BB_1, CC_1 divide each other in the ratio 3 : 3 : 1. A proof is left to the reader.

3. Let's tile the plane with copies

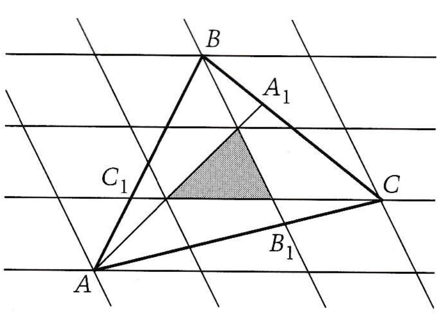


Figure 12

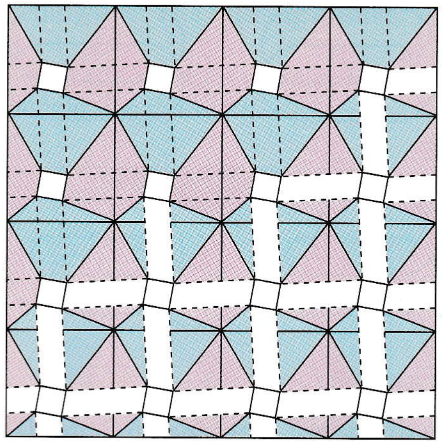


Figure 13

of the given big square with the small square inside it (fig. 13). This creates blue and red hexagons (we see them in the top part of the figure), and we have to prove they are of equal area. Join the small squares as shown. Then a "thick" grid consisting of small squares and parallelograms joining them emerges (see the lower part of the figure). Obviously this grid fits on itself when rotated 90° about the center of any small square. Therefore, the parallelograms are all congruent to one another and divide the plane into squares. Each of these squares consists of two blue and two pink triangles with a common vertex, and, as was mentioned in the article, the two blue triangles have the same total area as the two pink ones. Now, looking at the hexagons again, we see that a blue one comprises two blue triangles and a parallelogram, and a pink one consists of two pink triangles and a parallelogram. It follows that the hexagons have the same area, and we're done.

4. Let M and N be the midpoints of the sides AB and CD of a quadrilateral $ABCD$. Think of $ABCD$ as a tile in a quadrilateral tiling of the sort considered in the article. Let $CDEF$ be an adjacent tile (symmetric to $ABCD$ about N), and L be the midpoint of its side EF (fig. 14). Since M and L are symmetric about N, all these three points lie on one line, so $ML = MN + NL = 2MN$. On the other hand, $ML = AE$. Now the statement of the problem follows from the Triangle Inequality for triangle ADE

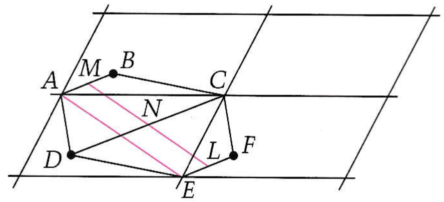


Figure 14

and the relation $DE = BC$.

5. We use the fact that the area of a parallelogram is less than half the product of its diagonals, and is equal to this product only if the diagonals are perpendicular. We can see this, for example, by looking at the four triangles formed by a parallelogram and its diagonals. We can then use the facts that the area of a triangle is half the product of two adjacent sides and the sine of the included angle, and that this sine is at most equal to 1 (when the diagonals are perpendicular).

Consider the given quadrilateral $ABCD$ as a sample tile, the corresponding tiling, and the parallelogram tiling associated with it. The area of any of the parallelograms is twice the area of $ABCD$, and it's no greater than the half-product of the diagonals of the parallelogram. As seen in figure 14, one of the diagonals, $AE \le AD + DE = AD + BC$. Similarly, the other diagonal is not greater than $AB + CD$. So the area of $ABCD$ does not exceed $\frac{1}{4}(AD + BC)(AB + CD)$.

To obtain equality, we must have $AD + BC = AE$, or $AD \parallel BC$ and $AB \parallel CD$; in addition, the diagonals of the parallelogram must be perpendicular. This means that $ABCD$ is a rectangle.

6. In figure 15, the area of $ABCD$ is twice the area of $APDQ$ (because

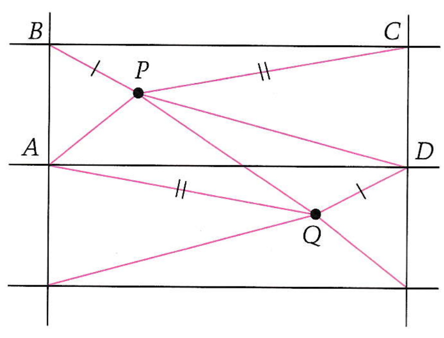


Figure 15

$\triangle ADQ$ is congruent to $\triangle CBP$). Now it's not hard to prove (see the solution to problem 5) that the area of a triangle is not more than 1/2 the product of two adjacent sides. Therefore, area $(APDQ)$ = area(APQ) + area$(DPQ) \le AP \cdot AQ/2 + DP \cdot DQ/2$ = $\frac{1}{2}(PA \cdot PC + PD \cdot PB)$, so area$(ABCD) \le PA \cdot PC + PB \cdot PD$.

7. The hexagon can be included in a tiling as shown in figure 5 in the article so that the vertices of the triangle in question are nodes of the corresponding grid. Then every hexagon covers three nodes, each of which belongs to three hexagons. So the hexagon's area is $3 \cdot 1/3 = 1$ "grid unit," while the area of the triangle is 1/2.

Tale of one city

(See the Happenings department in the May/June issue)

1. The regular networks are 3b, 3e, 3f, 3h, 3k, 3l, 3m, 3n, 3o, 3p, and 3q.

2. Consider a section of the network without any crossing over, as shown in figure 16. There are four branches, which we label A, B, C, and D. On either side, each of A and B will join up with one of C and D. Hence there are at most two components.

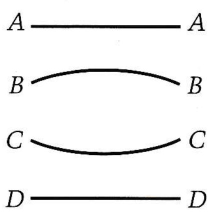


Figure 16

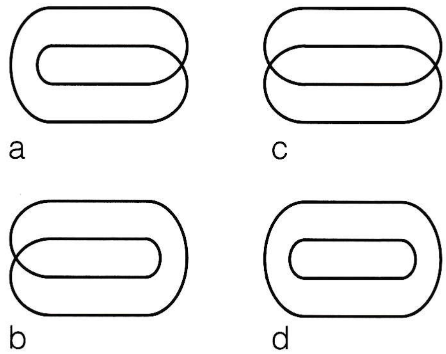


Figure 17

While the networks themselves can be very complicated, the solution to question 2 shows that there are only four basic patterns, as shown in figure 17. Two of them lead to regular networks and the other two lead to irregular networks.

3. By symmetry we need only consider the addition of integration points. Figure 18a shows the results when both points are added on the same side of the circle, while in figure 18b they are on opposite sides. Comparison with figure 16 shows that the former does not affect regularity. We can see that the same holds for the latter by going through the four cases in figure 17.

4. Suppose there is an integration point on one side of a circle. We can move it to the other side as follows.

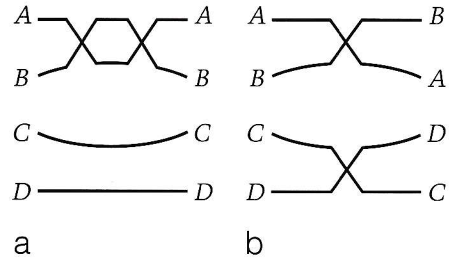


Figure 18

By question 3 we can first add two points on the other side and then remove one of them along with the original point.

5. The regular networks are (c), (d), (e), (h), (i), and (k).

6. The results of the three operations are shown in figure 19. Clearly regularity is not affected.

7. Observe first that the elimination of initial or terminal 0's does not affect the divisibility by 3 of the alternate sum, nor does any of the three operations in question 6. We can eliminate all 0's from a sequence as follows. First, we get rid of the outside 0's and the inside 00's. Any remaining 0 must be flanked by two 1's, both of which can be removed. This is because we can replace 101 with 0. Repeating, if necessary, we can either force the 0 to the outside or bring it next to another inside 0. In either case, the 0 can be eliminated. When only 1's remain in the sequence, we can get rid of 111's, so that the final

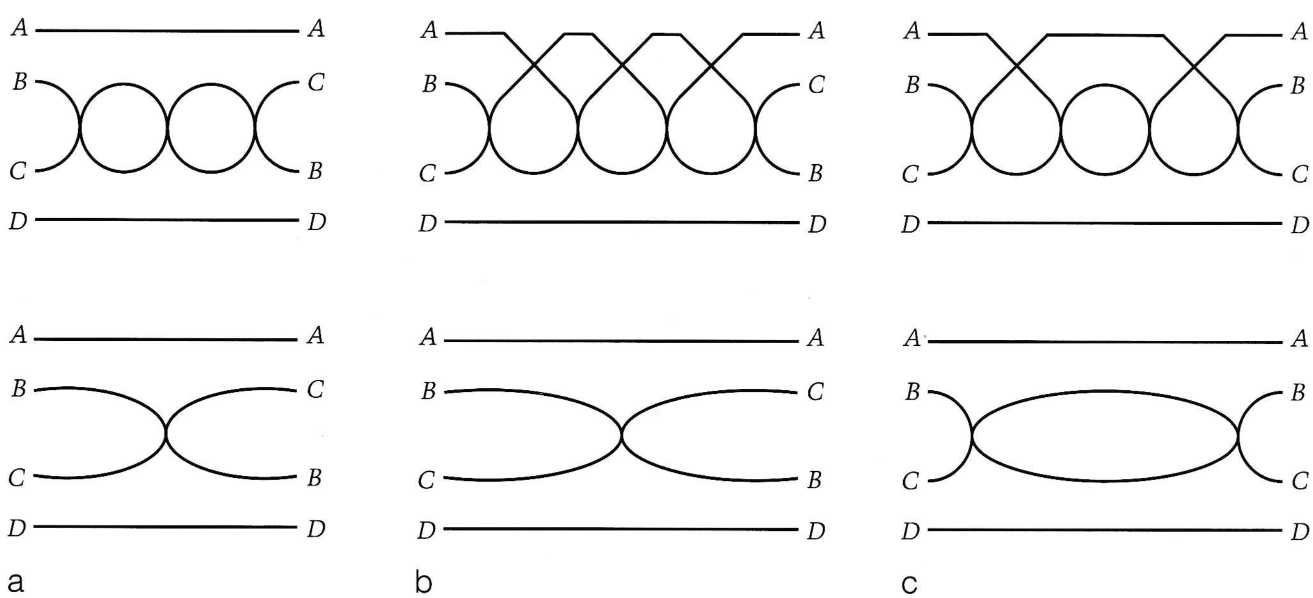


Figure 19

sequence is either empty or consists of one or two 1's. In the first case, the network is irregular, and the alternating sum 0 is divisible by 3. In the last two cases, the networks are regular, and the alternating sum, which is 1 or 2, is not divisible by 3.

Training problem. (a) We have

$$a^p - b^p = (a-b)(a^{p-1} + a^{p-2}b + \dots + ab^{p-2} + b^{p-1}).$$

Since $a \equiv b \pmod p$, $a - b \equiv 0 \pmod p$, and

$$a^{p-1} + a^{p-2}b + \dots + ab^{p-2} + b^{p-1} \equiv pa^{p-1} \equiv 0 \pmod p.$$

The desired result follows.

(b) (1) We use the fact, which the reader can prove, that if x is a central angle (see figure 20), a necessary and sufficient condition for point P to be on circle O is that $x + 2y = 2\pi$. In the figure accompanying the original problem, arc AB = arc DK (since AB

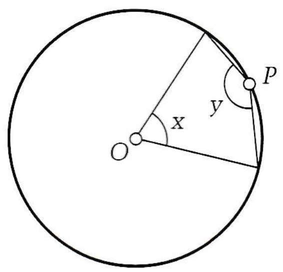


Figure 20

$\parallel BC$) and arc AK = arc AN (since $\angle B = \angle ADC$). Hence arc MN = arc NK (since AN is a diameter). This means that $NM = NK$, so we need only show that point C is on the circle with center N through M and K. Let arc AB = arc $DK = a$, arc AK = arc $AN = b$, and arc MN = arc $NK = c$. Then $\angle MNK = a + b$, $2\angle MCK = a + b + 2c$, and $\angle MNK + 2\angle MCK = 2a + 2b + 2c = 2\pi$. The result follows from the first statement in this solution. (2) Triangles ABC and KLC are similar, as are ADC and MLC. Hence $LC/LK = CB/AB = AD/CD = LM/LC$, so that $LC = \sqrt{ab}$.

Toy Store

The answer to Tim Rowett's puzzle: in the sequence 77, 49, 36, 18, each number except the first is the product of the digits of the preceding one. The missing number is $8 = 1 \cdot 8$.

For the solution to Peter Hajek's puzzle, see figure 21.

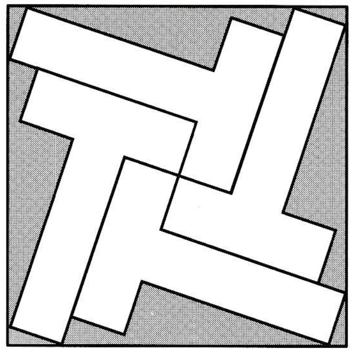


Figure 21

"AMAZING PARABOLOID" CONTINUED FROM PAGE 43

other conditions being equal), the shorter its "parallelism." As a consequence, not only is it practically impossible to construct a very deep paraboloid, it's also senseless from the theoretical point of view. The wave theory of light doesn't allow us to obtain a beam of light that is as powerful and as parallel as we like.

Thus, the beautiful and unexpected effects related to double reflection from the paraboloid of revolution are possible only within certain limits—when we can ignore the wave character of light. This occurs when the smallest of all the dimensions in the system—the radius r''_{min}—is many times greater than the wavelength. This condition imposes a limit on the degree to which light can be "amplified."

"MATH INVESTIGATIONS" CONTINUED FROM PAGE 30

the following corollary to a theorem of Gauss often useful: *It is impossible to construct with ruler and compasses a line whose length is a root or the negative of a root of a cubic equation with rational coefficients having no rational root, when the unit of length is given.* To use this theorem, he would assign carefully chosen rational values to the data, derive a cubic equation with roots the sides of the triangle (or other essential entities) and constructible coefficients, and then show that this cubic has no rational roots. In some other problems his data led to the constructibility of an angle of 20° or some other impossible situation. However, it should be noted that in all cases, even though it cannot be constructed by Euclidean tools, there should be a triangle satisfying the given data. In fact, all 186 problems may be viewed as reconstructions of a given triangle from the ingredients given.

Leroy Meyers's interest in these problems dates back to his high school days, when—like many other incipient mathematicians—he was intrigued by the variety of such problems and the clever problem-solving techniques needed for their solutions. Later his interest was rekindled by an article in the *Mathematics Magazine* written by William Wernick. Their subsequent collaboration will be the subject of my next column. Since I plan to feature there 20 other unsolved problems, in the meantime I challenge my readers to **resolve the 28 problems listed above**.

Index of Advertisers

The annual puzzle party

Everyone brings one present and takes away a bundle

by Anatoly Kalinin

ON AUGUST 20, 1994, THE city of Seattle, Washington, will host an event unusual in many respects. The most famous "puzzle people" from all over the world will gather there at the 14th International Puzzle Party. To give you an idea of what happens at these meetings, I'll tell you about the last one—the 13th Puzzle Party, which took place in the Netherlands last summer.

A year ago, on August 20, 1993, the halls of a fashionable restaurant outside Amsterdam were crowded from early morning on. More than a hundred people from the US, the Netherlands, Italy, France, the Czech Republic, Ireland, Malaysia, Portugal, Russia, Ukraine, and New Zealand had arrived there to hold their "congress." The international union of mechanical puzzle lovers arose many years ago, but it became really popular and well-organized in the early 1980s, when the famous Rubik's cube enthralled millions of people of all ages all over the world.

Today, however, the Hungarian Ernö Rubik is no longer the most famous personality in the puzzle community. Just as in the ancient view the world rested on three giant fishes, the modern world of clever toys is shouldered by three of its most famous puzzle-personalities: Jerry Slocum of the US, Edward Hordern of England, and Nob Yoshigahara of Japan. They possess the biggest collections (tens of thousands of puzzles in each), they've published many books about puzzles, and, of course, they've invented original puzzles of their own.

But let's return to Amsterdam. By tradition, everyone who comes to a Puzzle Party is obliged to bring a copy of an original puzzle (most often of his or her own devising) for every participant, who receives it free of charge. This multilateral trading results in full suitcases of new puzzles heading to new homes around the world. And for a puzzle fanatic, new puzzles are more valuable than Amsterdam's famous diamonds.

The conclave in and around Amsterdam continued for two days—a time filled with puzzle trading, attempts (often vain) to unlock the secret of a new puzzle right away, discussions of ideas for new puzzles, and excursions into the thousand-year history of puzzles.

The toy that was declared the best of those that were brought to the party was invented by one of the three contemporary giants—Edward Hordern. It's a brass six-sided prism pierced by a round rod. You have to draw the rod out of the prism. The surfaces of the prism and rod are absolutely smooth, without any projections or moving parts. All in all, the puzzle looks completely inaccessible and unsolvable—and no wonder. Only a month after the party each participant received a letter that revealed the secret: to disassemble the thing, you have to . . . kiss it! It turns out there's a little hole in the surface of the prism—so small you hardly notice it. You place your lips over this hole and draw the air in. The rod slips out of the prism all by itself.

Second place was won by another luminary of the puzzle world, Jerry Slocum. His puzzle was astonishing and paradoxical. It consisted only of four identical pieces resembling familiar jig-saw-puzzle shapes. All four pieces are linked to one another, and the task is to unhook them. It seemed as if witchcraft were involved: you could easily disconnect any two pieces, but to disassemble all four was absolutely impossible!

By way of illustration, I'd like to present a few puzzle gifts from the 13th Puzzle Party that can be solved in your head, or are fairly easy to make.

If there's the slightest chance that you can make your way to Seattle this August, take advantage of it and come to the 14th International Puzzle Party. You won't regret it! For details, contact Gary Foshee, 16006 266th Ave. S.E., Issaquah WA 98027, phone 206 392-2907.

Read the sentence directly below:

GREAT FUN FOR PEOPLE OF ALL AGES—TRY PUZZLETTS POCKET PUZZLES™—HAND MADE OF BRASS AND FINE WOOD.

Now, count the F's in the sentence above. Only once, though! Don't go back and count them again.

Solution: If you spotted four, you're a sharp one. If you got all five, you're a genius & much too good to be wasting your time on foolishness like this.

Figure 1

The puzzle from Mike Green (Kent, Washington).

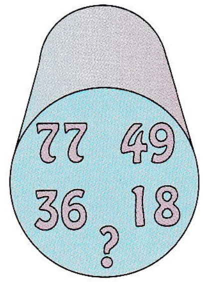


Figure 2

Tim Rowett, an English toy and puzzle designer, presented each party guest with a candy bar 1 inch in diameter and 8 inches long. The edible stuffing of the candy is permeated with nine figured (literally "figured") bands—also edible—such that as you bite off a piece you always see these mysterious figures on the end. So the process of solving the puzzle is combined with its consumption. To win this "biathalon" you not only have to find the relationship between the numbers and figure out the missing number, you have to eat the puzzle as fast as you can.

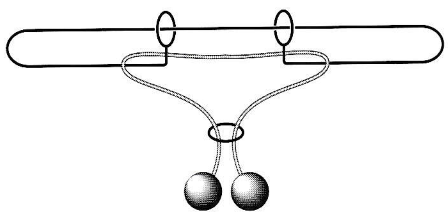


Figure 3

The puzzle from Howard R. Swift (Toledo, Ohio). The object is to remove the ring.

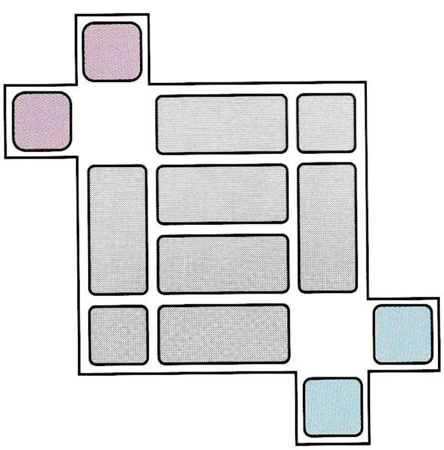


Figure 4

The present from Herman Witteveen (the Netherlands): exchange the two red blocks with the two blue blocks by sliding only.

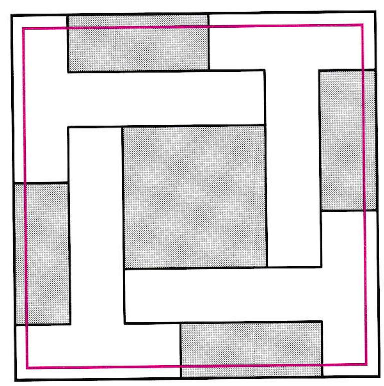


Figure 5

The present from Peter Hajek (UK): fit the four T's inside the red square. And a special additional question for Quantum *readers: after you solve this puzzle, try to figure out the smallest size of the red square, given the size of the big square, such that the solution is still possible. (We can vary the dimensions of the letters.)*

ANSWERS, HINTS & SOLUTIONS ON PAGE 55

New by SPRINGER

Ludwig Wittgenstein – Vienna Edition

edited by Michael Nedo

Parts already published:

Michael Nedo
Vienna Edition – Introduction

1993. 148 pages.
Section-sewn, paper bound DM 20,–,
US $ 24.50, £ 8.00
ISBN 3-211-82498-7

Vienna Edition Volume 1
Philosophische Bemerkungen

1994. XIX, 196 pages.
Cloth-bound DM 150,–, approx. US $ 100.00, £ 59.00
ISBN 3-211-82499-5

Vienna Edition Volume 2
Philosophische Betrachtungen
Philosophische Bemerkungen

1994. XIII, 333 pages.
Cloth-bound DM 210,–, approx. US $ 140.00, £ 83.00
ISBN 3-211-82502-9

Subscribers to the series are entitled to a 20% discount on the price of each individual volume.
The work will appear in the original German.
The introductory volume is free with subscription to the series.

Between two and five volumes will appear annually.
Each text volume will comprise 120 – 450 pages, of dimensions 8 1/2'' x 13'', section-sewn, cloth-bound.

Eleven volumes of edited text are initially planned, containing the writings from 1929 to 1933.
These will be thoroughly indexed by an accompanying series of concordance volumes.
An extension of the edition is intended.

Plan of the Edition

Introduction
English and German by Michael Nedo

Volume 1
„Philosophische Bemerkungen"

Volume 2
„Philosophische Betrachtungen"
„Philosophische Bemerkungen"

Volume 3
„Bemerkungen"
„Philosophische Bemerkungen"

Volume 4
„Bemerkungen zur Philosophie"
„Bemerkungen zur Philosophischen Grammatik"

Volume 5
„Philosophische Grammatik"

Volume 6
Synopses to manuscript volumes I–IV

Volume 7/1–2
Synopses to manuscript volumes V–X/1

Volume 8
„Philosophische Bemerkungen"
A reworking of the synopses
to manuscript volumes I–IV

Volume 9/1–2
A reworking of the synopses
to manuscript volumes I–X/1

Volume 10/1–2
The Big Typescript

Volume 11
Notebooks to manuscript volumes VI–IX

Prices are subject to change without notice.

Springer-Verlag Wien New York

Sachsenplatz 4–6, P.O.Box 89, A-1201 Wien · 175 Fifth Avenue, New York, NY 10010, USA
Heidelberger Platz 3, D-14197 Berlin · 37-3, Hongo 3-chome, Bunkyo-ku, Tokyo 113, Japan