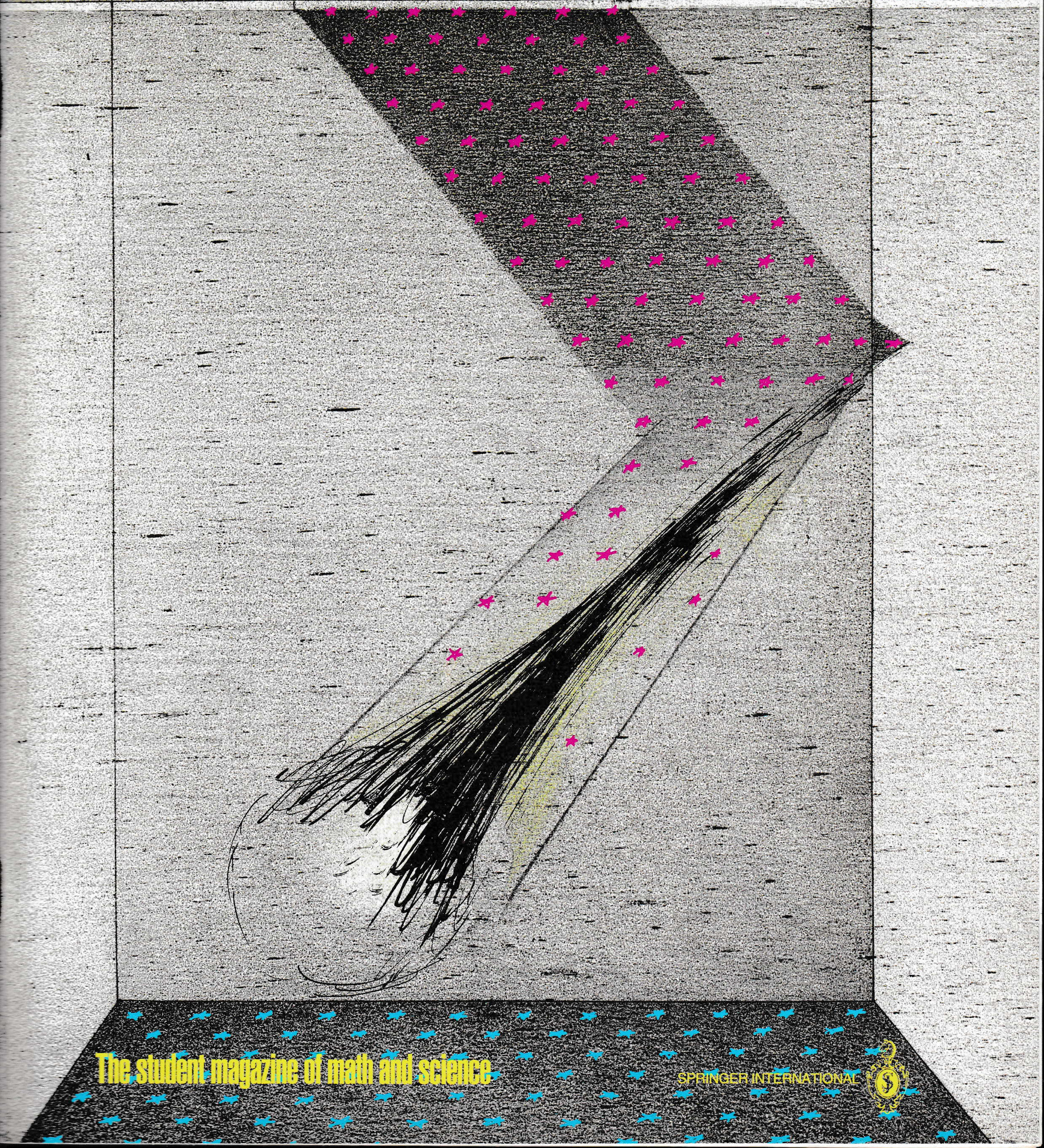
QUANTUM

NOVEMBER/DECEMBER 1992

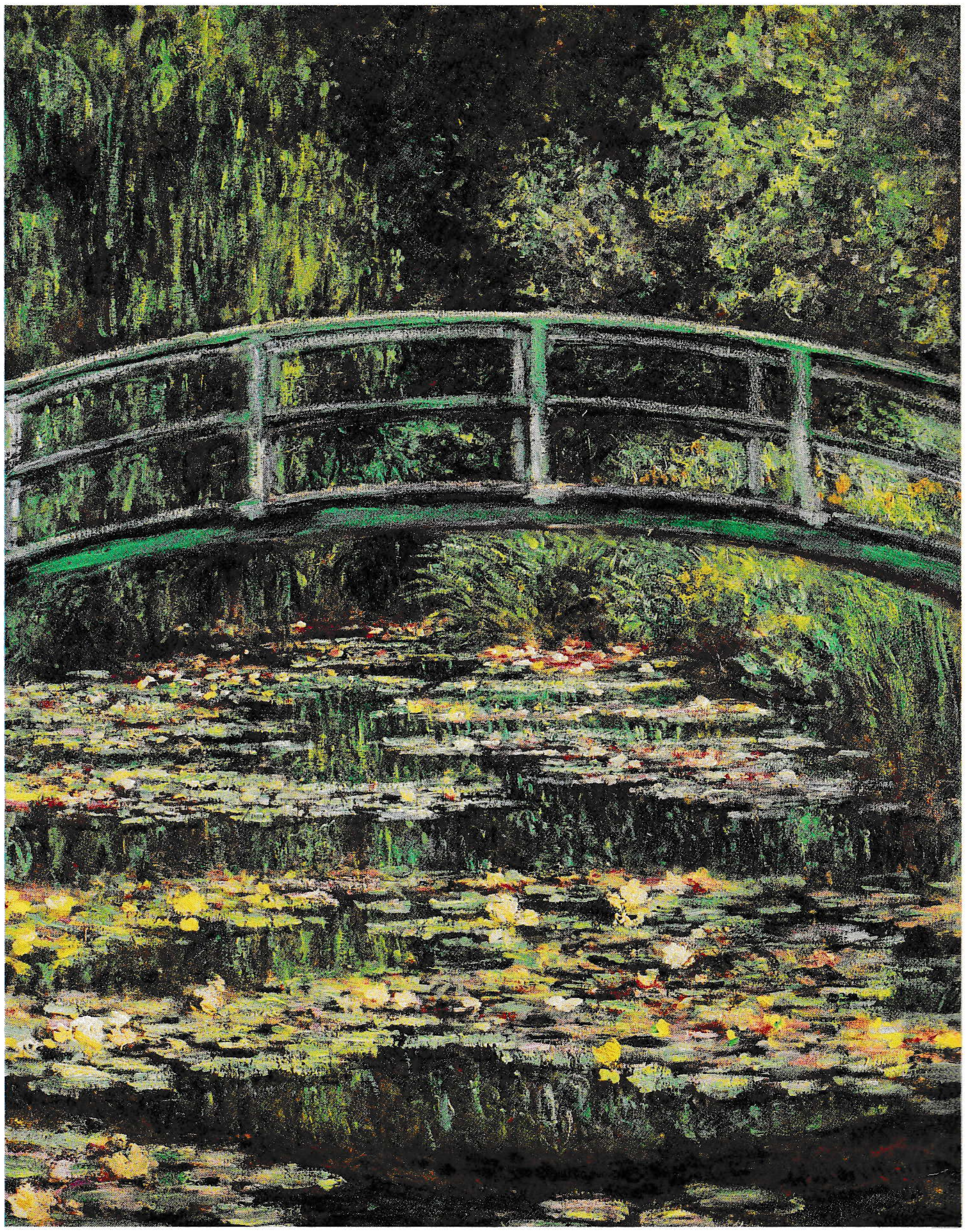
$5.00



The student magazine of math and science

SPRINGER INTERNATIONAL

GALLERY Q



Bridge over a Pool of Water Lilies (1899) by Claude Monet

At the age of 15, Claude Monet had already achieved a degree of artistic success. He sold some caricatures. They were well drawn, and they showed a keenly observant mind at work. But if the painter Eugène Boudin hadn't befriended him, Monet might have continued to sketch amusing portraits—might never have become the innovator in oils who helped change the course of modern art.

Boudin introduced Monet to the practice of painting out-of-doors, which was uncommon at the time. For the next 60 years, Monet turned his powers of observation on the play of light and color, water and sky, the subtle effects of atmosphere and weather. He would paint a single subject under various lighting and weather conditions—ordinary haystacks, for instance, or the Rouen Cathedral. The sets of paintings constitute a visual dissertation on the act of visual perception and the transformation of subjective impressions into pigmented canvas.

Late in life Monet created the famous garden that was to be his last refuge. This teeming piece of environmental art was tended by five gardeners. In the midst of the abundant vegetation was a pond with water lilies, over which a green Japanese bridge gently arched. The painting above is one of many studies Monet made of this tranquil scene.

You may not have noticed the two-dimensional frog on a lily pad directly below the bridge. When it finally jumps up toward the bridge, it will be a very long jump—in fact, an infinite one. It's a rare *Poincarian frog*, not visible to the naked eye. Turn to page 20 and you'll learn more about Poincaria (if not frogs).

QUANTUM

NOVEMBER/DECEMBER 1992 VOLUME 3, NUMBER 2

Cover art by Yury Vashchenko

One after another, they appeared in the sky, eliciting comment and superstitious dread as they had since time immemorial. But the three comets of 1618 occupy a special place in the history of science. They were immortalized in a series of polemical tracts on the nature of comets and, more generally, the proper approach to scientific questions.

Galileo was moved to respond to the opening salvo in the debate and, as it turned out, guessed wrong about comets. He hypothesized that they are terrestrial phenomena, not visitors from the far reaches of the galaxy. But in *The Assayer*, his major contribution to the controversy, Galileo succeeded in making a case for a new, open approach to science—one based on observational data and mathematical reasoning rather than armchair metaphysics and established authority.

The excerpt from *The Assayer* in the Anthology is Galileo's parable about the search for knowledge: the more we learn, the more careful we become in proclaiming The Truth.

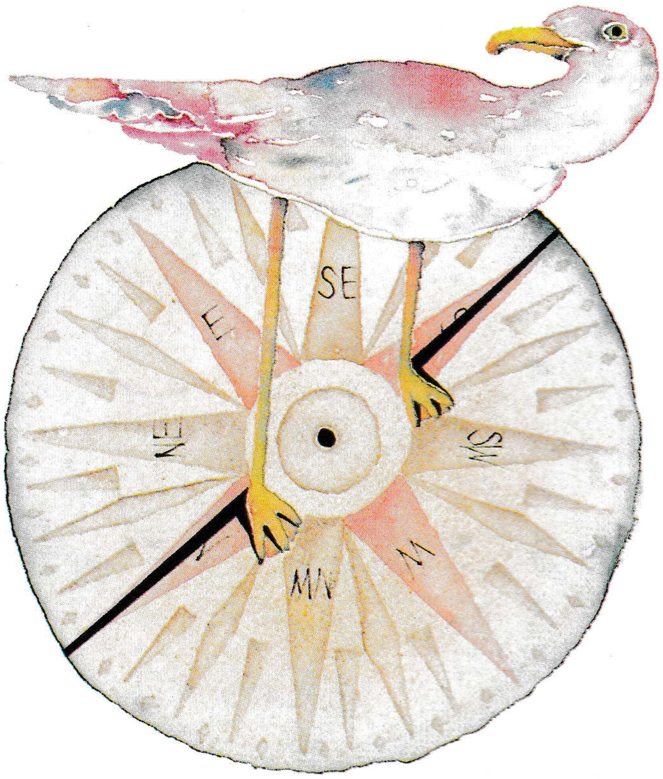
FEATURES

DEPARTMENTS

PUBLISHER'S PAGE

To err is human

To correct is divine



Art by Sergey Ivanov

IN THE SUMMER OF 1992 I conducted workshops for middle-level teachers in Houston, Texas. The workshops included lab work as well as discussion of basic concepts in physics, including astronomy. During one of these sessions I asserted something that, as it turned out, was not true. When I realized I had made a mistake and had found how to correct it, I sent an apologetic letter to the 77 teachers in the workshop. Here's a slightly revised excerpt:

You may remember that when I was doing my physics workshops there in June, I gave you a little quiz. One of the questions was: "When are shadows cast toward the south in Houston?" We told you, and I elaborated on the assertion that shadows would *never* be cast toward the south in Houston. We said that this was because the maximum angle of tilt of the Earth's axis with the pole of the ecliptic was 23.45°. I assured you that since Houston was at 29.77° north latitude, we were short some 6.32°.

When I arrived home in June, I went out on my boat. I was anchored for the day and was watching the sun set in the west. I was about 38° north latitude. I noticed that the Sun was much too far north, and shadows were quite clearly cast toward the south. This happened to be on June 21, the longest day of the year. I started the engine, reversed the prop, and strained the anchor line, steering so that the boat would take a heading toward the setting Sun. From the compass reading, taking into account local declination, the Sun was setting at about 304°. That would be 34° north of due west. From what I had told you, that angle should have been 6.32° south of west, or about 264°. The Sun was setting 40° further toward the north than I thought it should have been. My compass certainly does not create errors that large. I realized that what I had asserted at the workshop was wrong!

At first I tried to construct a model, using a ball. Then I tried to find some way to calculate the angles involved. The more I thought about this problem, the more I realized that I needed to use vectors and do geometry. After doing all of that work, I found the solution. . . .

Now, how did I arrive at my conclusion? Send us your solution to this problem. It should give times for sunrise and sunset, shadow lengths, and other such information as a function of latitude and time of day (based upon longitude). We'll publish the most interesting and clearly written solution to this problem. Try to take into account the fact that the Earth is not a sphere (although spherical-Earth solutions will be accepted as well).

Teachers make mistakes. After all, we're not the source of knowledge—nature is! What troubles me the most is that nobody challenged my unsupported, wrong assertion. You must always ask questions and challenge things that sound fishy, or look like cardboard props with nothing behind them. Admit it when you don't know the facts or don't understand a concept. Ignorance isn't bliss. But when we recognize it, we've taken the first step toward solid knowledge.

—Bill G. Aldridge

QUANTUM

THE STUDENT 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)

Quantum (ISSN 1048-8820) is published bimonthly by the National Science Teachers Association in cooperation with Springer-Verlag New York Inc. Volume 3 (6 issues) will be published in 1992–1993. *Quantum* contains authorized English-language translations from *Kvant*, a physics and mathematics magazine published by the Russian Academy of Sciences and the Russian Academy of Pedagogical Sciences, as well as original material in English. **Editorial offices:** NSTA, 3140 North Washington Boulevard, 2nd Floor, Arlington, VA 22201. **Production offices:** Springer-Verlag New York, Inc., 175 Fifth Avenue, New York, NY 10010. Production Editor: Madeline R. Kraner

Advertising:

Advertising Representatives: (Washington) Paul Kuntzler (202) 328-5800; (New York) Thomas Heitzman (212) 460-1675; and G. Probst, Springer-Verlag, Heidelberger Platz 3, 1000 Berlin 33, FRG, telephone (0) 30-82 07-1, fax 30-82 07-1.

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

Subscription Information: North America

Student rate: $14; Personal rate (nonstudent): $18; Institutional rate: $30; Single Issue Price: $5. Rates include postage and handling. (Canadian customers please add 7% GST to subscription price. Springer-Verlag GST registration number is 123394918.) Subscriptions begin with next published issue (backstarts may be requested). Bulk rates for students are available. Send orders to *Quantum*, Springer-Verlag New York, Inc., P.O. Box 2485, Secaucus, NJ 07096; or call 1-800-SPRINGER (in New York, call [201] 348-4033).

All Countries Outside North America

Subscription rates in U.S. currency as above (all rates calculated in DM at the exchange rate current at the time of purchase) plus postage and handling. SAL (Surface Airmail Listed) is mandatory for Japan, India, Australia, and New Zealand. Customers should ask for the appropriate price list. Air mail delivery to all other countries is available upon request. Orders may be placed through your bookseller or directly through Springer-Verlag, Heidelberger Platz 3, 1000 Berlin 33, FRG.

The dark power of conventional wisdom

How even a genius can fail to see—or share—a new truth

by A. D. Alexandrov

ANY PERSON WHO HAS worked with mathematics—solving problems, demonstrating propositions, developing new concepts—has certainly had occasion (and not just one) to be amazed by one's own dullness. You think the problem through, again and again, but can't solve it, and when you learn the answer you think: "What a dummy I am! Why didn't I think of that?" Or sometimes you think and think and finally solve it, and you're glad, but you still think: "What a numbskull I am! Why didn't I think of that sooner?"

Problems are one thing, but entirely new concepts are another. Suppose you're working on some problems, but it never occurs to you to look at them from another angle or from a more general point of view. And so you never formulate the general concepts that would clear up a whole family of problems. If—oh, joy!—you happen to figure them out, you're surprised: "Why didn't this occur to me earlier?" And if someone else already thought of it, then no matter how gratified you are by the march of science, you're angry with yourself: "How could I have missed that! What a blockhead I am!"

In short, anyone who has thought hard, pondered deeply, and searched thoroughly for an answer knows how thick and dull a person can be. Usually only people who have never cogitated, pondered, or worked at difficult problems are capable of admiring their own sagacity. Success comes easily to those who place no difficult problems or serious goals before themselves.

Now I'm going to tell you a story of human thickheadedness and genius, a story much more significant than the things I've just been talking about. I'll be talking about one of the greatest achievements of the human spirit, in which first-class talents and true geniuses took part. Our topic will be non-Euclidean geometry and its 2,000-year-plus history. But first there's something I'd like to explain.

What is Lobachevskian geometry?

You probably know by now that Lobachevskian geometry is the geometry obtained from Euclidean geometry by changing only one axiom—the parallel axiom. Lobachevsky took the following as an axiom: if a point is not on a given line, then there are at least two straight lines parallel to the given line that pass through the point. The statements, or theorems, derived from Euclidean foundations altered in this way constitute *Lobachevskian geometry*. It's all very simple, isn't it? It trips off the tongue easily and clearly. But the difficulty lies in the fact that Lobachevsky's axiom doesn't correspond to our visual notions. And so the conclusions drawn from it (the many theorems of Lobachevskian geometry) turn out to be unconventional and hard to imagine. The real meaning of this geometry is not at all clear from the simple formal definition given above.

Lobachevsky himself called his geometry *imaginary*. He considered it a theory that could be applied to real space. *Could* be . . . but there simply were no practical applications. So the logical consistency of his geometry remained unestablished. No matter how deeply Lobachevsky developed his theory, a contradiction still might show up later.

The real meaning and logical consistency of Lobachevskian geometry follow from a simple model of it devised by the German mathematician Felix Klein. Here's the model.

The interior of a circle (fig. 1) is taken to be a "plane," all the points of the interior stand for "points," and chords are "straight lines"—of course, their ends are excluded, since only the inside of the circle is under consideration. "Isometries" are taken to be the transformations of the circle that take it into itself and chords into chords. Correspondingly,

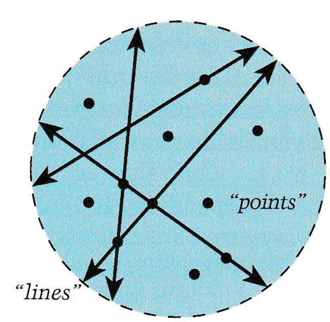


Figure 1

Opposite (l to r): János Bólyai, Nikolay Lobachevsky, and Carl Friedrich Gauss.

Art by Pavel Chernusky

geometrical figures mapped onto one another by such transformations are called "congruent."

Any theorem of Lobachevskian plane geometry becomes, in this model, a theorem of Euclidean geometry. This general statement is proved by verifying the validity of the axioms of Lobachevskian geometry in the model.[1] The fact that the parallel axiom fails in this model is immediately seen from figure 2: through the point C outside the straight line (that is, chord) AB there are an infinite number of "straight lines" (chords) that do not intersect AB.

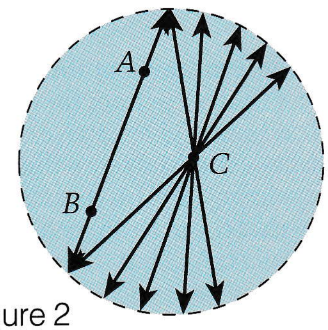


Figure 2

Therefore, if there is a contradiction in Lobachevskian geometry, the same contradiction (rendered into the "language of the circle") emerges in Euclidean geometry as well. So Lobachevskian geometry has real meaning—as much meaning as geometry applied to real objects can have.

Lobachevskian geometry is as consistent as Euclidean geometry, and has the same degree of real, experimentally established meaning.

From Euclid to Lobachevsky

Back in the 4th century B.C. Euclid himself stated the "parallel axiom" (his Fifth Postulate) as follows: *If a straight line crosses two straight lines and forms interior opposite angles whose sum is less than two right angles, then, when indefinitely produced, these lines will meet on the side where the angles are less than two right angles.* We've given Euclid's formulation of the statement only to show how complicated it is. Other postulates are much simpler and their formulations are shorter—for example, the First Postulate: *A straight line can be drawn through any two points.*

Naturally attempts were made to get rid of this rather complicated Fifth Postulate, to derive it from other fundamental premises of geometry. I think Euclid made such attempts himself, or, in any case, they were undertaken already in his lifetime. Some Arab authors mentioned Archimedes's treatise "On Parallels" (3rd century B.C.) where, presumably, the Fifth Postulate was deduced from some simpler premises.

Attempts to prove the Fifth Postulate continued from that point on, over the course of 2,000 years. A lot of scientists took up the challenge. Here's a partial list: the Greeks Ptolemy (2nd century A.D.—the very Ptolemy who created the Ptolemaic system in astronomy) and Proclus (5th century); the Arab al-Khaisam (10th century); the Persian (or Tadjik) Omar Khayyam (11th century or beginning of the 12th century—the same Omar Khayyam who is known as a great poet); the Azerbaijani at-Tousi (13th century); the German Clavius-Schlussel (1514—from here on the date of the treatise is given); the Italians Cataldi (1603), Borelli (1658), and Vitale (1680); the Englishman Wallis (1663); the Italian Saccheri (1733); the German Lambert (1766); the Frenchmen Bertrand (1778) and Legendre (1794, 1823); the Russian Guryev (1798). All their attempts amounted to the deduction of the Fifth Postulate from some other assumption. Many of them didn't notice this and thought they had succeeded in proving the postulate. Others, who were more penetrating and critical, formulated the statement from which the Fifth Postulate had been deduced—as, for instance, Omar Khayyam did.

In the 17th and 18th centuries, as mathematics developed rapidly, the search for a proof intensified. An Italian monk—a teacher of mathematics and grammar by the name of Girolamo Saccheri—made significant efforts. The treatise containing his attempt at a proof of the Fifth Postulate appeared in 1733—the year of his death. Its title was *Euclid Cleansed of All Blemishes, or A Geometrical Attempt to Establish the First Principles of All Geometry.* Relying on his predecessors' works, Saccherri tried to prove the Fifth Postulate by *reductio ad absurdum*: having drawn a conclusion equivalent to negating the Fifth Postulate, he deduced consequences from it, trying to come to a contradiction. But as the negation of the Fifth Postulate is Lobachevsky's axiom, his (Saccheri's) conclusions were nothing more or less than some theorems of Lobachevskian geometry. In other words, Saccheri was developing a new geometry without realizing what he was doing. He failed to come to a contradiction; he nevertheless concluded that he had proved the Fifth Postulate, although he apparently wasn't too sure of this himself.

The 18th century saw a rash of treatises on the theory of parallels (55 of them!). The most prominent among them was written by the German mathematician, physicist, and astronomer Johann Heinrich Lambert in 1766. Attempting a proof of the Fifth Postulate by *reductio ad absurdum*, he deduced many consequences from its negation. We might say it was Lambert who built the foundations of Lobachevskian geometry. There was no contradiction in his conclusions, and he didn't think he had found any (unlike almost all of his predecessors). Lambert even expressed the thought that he "almost must draw the conclusion" that the hypothesis he tried to refute is "valid on some imaginary sphere." Nevertheless, he was still certain that geometry based on the negation of the Fifth Postulate is impossible. His paper, however, didn't prove this conviction. So Lambert was apparently dissatisfied with his work and didn't publish it. It was published nine years after his death—20 years after it was written. Lambert came very close to discovering a new geometry, but didn't.

The German mathematicians Schweikart (1818) and Taurinus (1825) fully realized the possibility of non-Euclidean geometry, but they

[1] You can read more about models of axiomatic systems, and models of non-Euclidean geometry in particular, in the article by Vladimir Boltyansky in the last issue of *Quantum*. See also the article by Simon Gindikin, "The Wonderland of Poincaria," in this issue.

never expressly stated that the theory they posited would be just as logically rigorous as Euclidean geometry.

Carl Friedrich Gauss, according to his own testimony, began analyzing the theory of parallels in 1792. Judging from his correspondence, we can conclude that he was gradually approaching the idea that a proof of the Fifth Postulate is impossible. In 1817 he wrote to Olbers: "I am coming closer and closer to the conviction that the necessity of our geometry cannot be proven, at least by a human reasoner and for human reason." Since he wrote "I am coming closer and closer," he clearly hadn't reached that conviction yet. He wrote further: "Perhaps in another life we might come upon other views of the nature of space that are inaccessible now. Till then, geometry must be ranked not with arithmetic, which is purely a priori, but with mechanics . . ." At the same time he moved non-Euclidean geometry forward considerably. But it wasn't until 1824, in a letter to Taurinus, that he wrote quite definitely that non-Euclidean geometry, "in which the sum of angles in a triangle is less than 180°, is quite logical" and that he had "developed it for himself rather satisfactorily." Nevertheless, it wasn't until 1831 that he set down his deductions (albeit briefly); and he never published anything on non-Euclidean geometry. In 1829, in a letter to Bessel, Gauss wrote: "I fear the Boeotians will bellow if I express my views."[2] He was afraid of losing his authority in the scientific community.

While Gauss was busy writing all this, a man was found who not only developed a quite satisfactory geometry negating the Fifth Postulate, and not only came to the conclusion that this geometry was absolutely logical; he also reported his ideas to a scientific gathering without fear of anybody "bellowing." It was Nikolay Ivanovich Lobachevsky, who came to the conclusion as early as 1824 that non-Euclidean geometry was feasible. On February 11, 1826, he submitted his report to the physical and mathematical department of Kazan University. This report was enlarged in his paper "On Elements of Geometry," which was published in the *Proceedings of Kazan University* from February 1829 through August 1830.

In 1835–38 Lobachevsky published an elaborated statement of his theory, "New Elements of Geometry with a Complete Theory of Parallels." He wrote in the preface: "The fruitless efforts since Euclid's time, over the course of two thousand years, have led me to suspect that the notions themselves do not contain the truth that people have tried to prove and that, like other physical laws, can be verified only by experiment—for example, by Astronomical observations." Lobachevsky regarded the problem of verification of any geometry as a matter of experience.[3] He considered his geometry a possible theory of the properties of real space—that is, of the structure of corresponding relations between material bodies and phenomena.

Almost simultaneously with Lobachevsky, in 1825 the young Hungarian mathematician János Bólyai[4] came upon the same geometry. In 1832 Bólyai stated his conclusions and published them as an appendix to his father's textbook on geometry. His father, Farkas Bólyai, sent the textbook to Gauss. The latter approved János Bólyai's results but at the same time said that he had been aware of this for a long time. Realizing the significance of his discovery, János thought Gauss was taking credit for the discovery. From that point he stopped working on non-Euclidean geometry for a long time. But Lobachevsky continued to develop his geometry and to publish papers presenting his ideas right up until he died.

One shouldn't be surprised that the new geometry seemed impossible. Look at figure 3: clearly, the straight line CM is certain to cross the line AB when extended far enough. The reverse assumption seems absurd. There can be no such thing as non-Euclidean geometry! Which makes Lobachevsky's and Bólyai's idea all the more bold—they decided to allow such "absurdity." Nowadays, when the straightforward meaning of non-Euclidean geometry is common knowledge, you don't need to be bold at all—all you need is the merest capability of abstract thought.

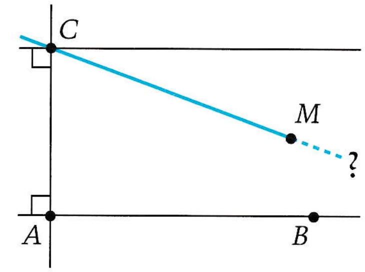


Figure 3

From belief to proof

And so Lobachevsky and Bólyai openly, and Gauss in his letters, expressed their conviction that non-Euclidean geometry is valid and developed it further. Nevertheless, until the real meaning of the new geometry was found, this great discovery was left hanging: Lobachevskian geometry was nothing more than an imaginary construct.

In 1839–40 two papers appeared, written by F. Minding, a professor at Derpt (Tartu) University. In them he examined certain special surfaces: surfaces of constant negative curvature. In essence, Minding came to the conclusion that the geometry on such surfaces is actually the geometry of Lobachevsky, though he didn't state this explicitly. It's interesting to note that two years before Minding's papers were published, one of Lobachevsky's works was published in the very same journal!

In 1854, upon taking the post of professor at Göttingen University, Bernhard Riemann delivered a trial lecture, as was the custom. It was entitled "Hypotheses Constituting the Principles of Geometry." The lecture was a treasure trove of fruitful ideas: from a general conceptualization of mathematical space to statements that anticipate the general theory of

[2]The inhabitants of Boeotia, a province of ancient Greece, were said to be incredibly stupid. "Boeotians" was a term of abuse in Gauss's time.

[3]The Russian word *opyt* can mean either experience or experiment.—*Ed.*

[4]Pronounced YAH-nosh BO-yai.—*Ed.*

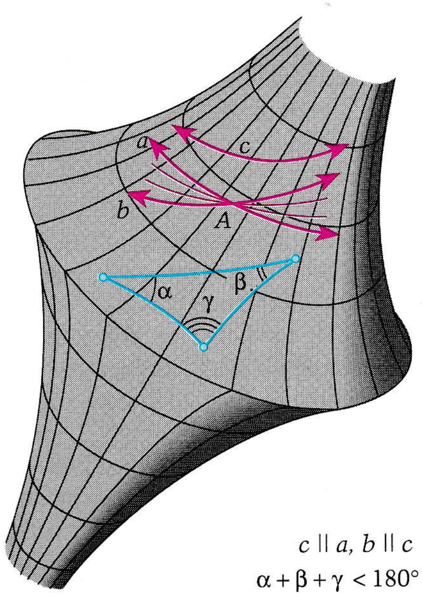
relativity. In addition, Riemann introduced the general theory of a certain kind of space (now called *Riemann space*), which includes, as the simplest instances, Euclidean and Lobachevskian space and so-called spherical space. Riemann gave a purely analytical definition of these kinds of space; in essence, this meant that Lobachevskian geometry is as consistent as mathematical analysis.

But nobody understood Riemann's lecture, or noticed it. Only one man listened to it with great interest and left the hall deep in thought. It was the 77-year-old Gauss. Riemann's lecture wasn't published until 1868, two years after his death. But then it immediately made a tremendous impression and inspired the rapid development of the theory introduced there.

In 1857 Codazzi developed Minding's ideas, but he didn't compare his results with non-Euclidean geometry. Maybe he didn't know about it, even though some of Lobachevsky's works had been published in French and German, and Bólyai's paper was issued in Latin in 1832.

It wasn't until 1868 that Beltrami, relying on Minding and Codazzi, made the necessary comparisons and produced a detailed proof that Lobachevskian geometry holds for surfaces of constant negative curvature (like the one in figure 4).

Beltrami's results were analytical, however, and far from elementary geometry, far from Euclid. Three years later Klein noticed his model on a circle, which was described at the beginning of this article. Later Poincaré found another interesting model that involved complex numbers.[5]



$c \parallel a,\ b \parallel c$

$\alpha + \beta + \gamma < 180°$

Figure 4

So, 40 years after the publication of the first papers by Lobachevsky and Bólyai, their conviction was proved and their geometry was acknowledged throughout the world.

Genius

The idea that some geometry other than that of Euclid was even thinkable dawned on Lambert as early as 1766; in the first half of the 19th century the same idea was proposed by Schweikart, Taurinus, Gauss, Lobachevsky, and Bólyai.

To formulate an idea refuting a universally accepted idea is a great gift in and of itself. But it isn't science yet—only an idea. Science requires the transformation of an idea into theory, just as engineering requires the transformation of an idea into an invention.

Genius is not merely a flight of fancy but persistence—steady work supported by inspiration, and inspiration fortified by hard work. Copernicus expressed the notion that the Sun, not the Earth, is at the center of the solar system (an idea expressed by Aristarchus of Samos way back in the 3rd century B.C.), but he went on to construct the Copernican system and give an exact description of a planet's motion around the Sun that agreed with observations. The same is true with Lobachevsky. Not only did he express his belief in the possible existence of non-Euclidean geometry, he also constructed that geometry. And just as Copernicus opened a new path in astronomy, leading to the modern concept of a universe with many "worlds," planetary systems, galaxies, and so on, Lobachevsky blazed a trail in geometry, leading to the creation of many different "geometries," the most varied geometrical theories of "imaginary" space—topological, Riemann, Finsler . . . there's no counting them.

[5]See "The Wonderland of Poincaria," page 20.

In the 1860s and 1870s, when mathematicians began to reconstruct geometry in earnest, Lobachevsky was called "the Copernicus of geometry," and justifiably so. We mustn't forget, of course, that a new geometry was developed and published by Bólyai as well, but the credit goes to Lobachevsky because he did it earlier, and thereafter he continued his investigations and continued to publish his results.

The appearance of non-Euclidean geometry marked the beginning of a revolutionary transformation of geometry. As happens with every revolution, reactionary elements arose along with the forces bearing the revolution forward. Just when the new geometry was about to be discovered, the philosophy of Immanuel Kant appeared on the scene. In 1781, in his *Critique of Pure Reason*, Kant held that geometry is a priori (independent of experience) and deduced from this the apriority of space itself, as an a priori form of contemplation. Any geometry other than the one inherent in this form of contemplation seemed unthinkable.

Lobachevsky openly opposed these views. Like all great scientists, he was a philosopher: science can't make progress without philosophy. The appearance of the new geometry resulted in unknown and as yet inconceivable ways of proceeding in science. The revolution took hold. Genius is revolution, and revolution is genius in action.

Dullness

The history of the Fifth Postulate and non-Euclidean geometry demonstrates human genius, but it also shows the mind's clumsiness (if not outright dullness).

To begin with, most attempts to prove the Fifth Postulate were based on errors. To the authors it appeared that they had found a proof. This persisted even to the beginning of the 19th century. Only a few realized that they were relying on additional assumptions equivalent to the Fifth Postulate and explicitly formulated them.

The most characteristic example is Saccheri: along with his deep and refined conclusions concerning non-Euclidean geometry, he still ends up concluding that he had succeeded in taking the hypothesis that denies the Fifth Postulate and "pulling it out by the roots," cleansing Euclid of his "blemishes" in the process.

When non-Euclidean geometry was discovered and promulgated, and questions about its real meaning arose, dimwittedness showed itself in full force.

In 1827 Gauss developed the principles of the general theory of geometry on surfaces in which the shortest curves play the part of straight segments. In particular, he proved that on a surface of negative curvature, the sum of the angles of a triangle is less than 180°. He knew also that it was the same in non-Euclidean geometry. But he didn't compare those conclusions, nor did he surmise that non-Euclidean geometry must be realized on certain surfaces. If he had grasped the idea, the proof would not have been difficult for him, a mathematician of extraordinary intellectual power.

Such a phenomenon is rather common. It doesn't occur to people to compare things that seem quite different but, upon scrupulous examination, prove to be closely connected or even coincident.

It was the same with non-Euclidean geometry and geometry on surfaces with constant negative curvature. Neither Minding and Codazzi, nor even Lobachevsky, made such a comparison! It was Beltrami who did it 40 years after Gauss's work.

Meanwhile, in 1859 Cayley created a theory of distance containing a model of Lobachevskian geometry, but he didn't realize it because he hadn't juxtaposed his theory and Lobachevskian geometry. And later, in 1861, he published a paper on Lobachevskian geometry!

In wasn't until 1871 that, after making such a comparison, Klein came upon a simple model in a circle. This elementary model settles the question of the unprovability of the Fifth Postulate. That's where it all ended—there lies the final resting place of the 2,000-year problem that tortured the world's greatest mathematical minds!

Nowadays there are people who are still trying to "prove" the Fifth Postulate and who besiege mathematicians with their "results." But since the problem of the Fifth Postulate is solved by the model in a circle and it's not difficult for anyone to understand this solution, all the so-called "proofs" and "results" can't be relegated to mental clumsiness but to plain old stupidity, without ruling out the possibility that the authors are bona fide medical cases. Stupidity is not the same as dullness. Mental clumsiness can be characteristic even of geniuses: this is all too readily apparent from the history of the Fifth Postulate and non-Euclidean geometry.

Character

Gauss, Bólyai, and Lobachevsky are the three mathematicians who discovered non-Euclidean geometry. Three persons—three personalities.

Friedrich Gauss was a mathematician of extraordinary ability. He called "the great Gauss," *princeps mathematicorum* ("first among mathematicians"), the "prince of mathematicians."

But for all his mathematical power, Gauss was intellectually cautious and indecisive by nature. He spent more than 30 years investigating the theory of parallels before he decided to express, even to himself, his firm conviction that non-Euclidean geometry is valid. Then there was another kind of caution—cowardice, which kept him from publishing his results for fear of rousing the Boeotians against him.

János Bólyai contrasts sharply with Gauss. He was only 23 years old when he came up with non-Euclidean geometry (Gauss did it at the age of 47, Lobachevsky at 31). Lobachevsky was 32 when he began publishing, Bólyai was 30, and Gauss never published on this topic during his lifetime. Bólyai's paper on non-Euclidean geometry was brilliantly written—if anything, it was rather too concise. The brilliance of his talent corresponded to other features of his passionate nature. He was an army officer—one of the proud Hungarian hussars—and a duelist. Once he managed to get into a sword fight with several opponents. He fought them, one after the other. He had reserved the right to play the violin between duels to restore flexibility to his hand. He ended up pinning them all (not to death, though).

Pride ruined Bólyai—not because he was killed in a duel, but because his pride extended to his mathematics. As was mentioned above, after hearing Gauss's response to his work on non-Euclidean geometry, Bólyai concluded that Gauss had simply appropriated his discovery. When one of Lobachevsky's books was published in German, he thought it was Gauss who had taken the pseudonym "Lobachevsky" and had stolen his (Bólyai's) results. Besides the discovery of non-Euclidean geometry, Bólyai wrote another mathematical paper with ideas that were ahead of their time, but the paper wasn't done properly. During his last year of life, his mind became clouded. He died in 1860, at the age of 58.

Lobachevsky differed greatly from both Gauss and Bólyai: he combined boldness with persistence and thoroughness, the power of theoretical thought with a strong will. His discovery was not accepted, and he was thought to be a bit crazy, as Chernyshevsky characterized him.[6] Recognition from Gauss came later. But Lobachevsky wasn't abashed and continued his "crazy" research on "crazy" geometry and continued publishing his papers after his first fundamental work appeared in 1829–30. He went blind in his old age and had to dictate his last book, *Pangeometry*.

Lobachevsky was known for more than his scientific activity. For 18 years he was rector of Kazan University, showing outstanding energy, administrative dexterity, and an understanding of the challenges involved in educating the younger generation. During the difficult times of a cholera epidemic in 1835, his energetic and competent behavior seemed

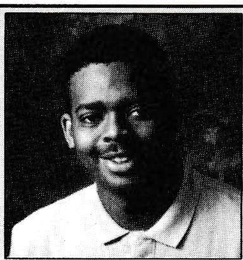
[6]See the beginning of "The Wonderland of Poincaria."

unlikely for a man who dealt with imaginary geometry, one of the most abstract fields in the most abstract of sciences—mathematics. But maybe we shouldn't be so surprised. A strong will, necessary for resolute action under hard conditions, is just as necessary for developing and defending one's scientific convictions and the truth in the face of bellowing Boeotians.

Talent and genius are not merely specific capabilities, but a matter of character as well. Magellan and Nansen needed determination to sail to unknown lands; theorists need intellectual determination to think the "improbable" and develop it in spite of established views and traditions, and sometimes even despite their own doubts. But to be convinced of one's own ideas isn't enough; the ideas must be communicated to others. And this step also requires determination, because people often misunderstand, or toss away, or even ridicule, new ideas and results. And one's colleagues are often the first to do it—scholars who are convinced of the incontrovertibility of their views, sure of their intellectual infallibility. They are the academic Babbitts occupying prestigious chairs at universities—the Boeotians whom Gauss didn't dare stand up to.

Even though Gauss had nothing to fear except the uncomplimentary opinions of his colleagues, he nevertheless concealed his scientific convictions—he hid the truth. He behaved wisely from the point of view of philistines—past or present—who pursue occupations in science (or in any other field, for that matter).

Real science requires boldness of thought and boldness in openly professing a bold thought. The history of the Fifth Postulate and non-Euclidean geometry shows how many difficulties must be overcome even by bold-thinking people to reach truths that seem so simple once they have been discovered. It shows how the most brilliant minds are capable of dull thought. So great scientists combine boldness of thought with modesty in assessing their own achievements. Along with laws and axioms, we would do well to learn that lesson from the history of science.



ALISTON THOMAS

B.S. in Chemistry; Ph.D. Graduate Fellow and Teaching Assistantship in Chemistry at the University of Illinois, Urbana-Champaign; summer internship Pralex Corporation, Christiansted, St. Croix, USVI.

PROFESSIONAL & GRADUATE OPPORTUNITIES

✓ After freshman year, students can enhance academic programs each summer with professional paid work experiences for credit in companies and research institutions throughout the United States and in Europe.

Recent opportunities have included
Proctor and Gamble, Cincinnati, OH;
Monsanto Agriculture New Products Division, St. Louis, MO;
Owens Corning Fiberglass, Granville, OH;
Michael Reese Medical Research Center, Chicago, IL;
Brookhaven National Laboratory, Upton, NY;
Argonne National Laboratory, Argonne, IL;
Meharry Medical College, Nashville, TN.

✓ 95% of Clarke graduates are employed or in graduate school within six months of graduation.

Celebrating 150 years as a four-year liberal arts college.

Join the SUPERACHIEVERS in SCIENCE and MATH at CLARKE COLLEGE

DUBUQUE, IOWA

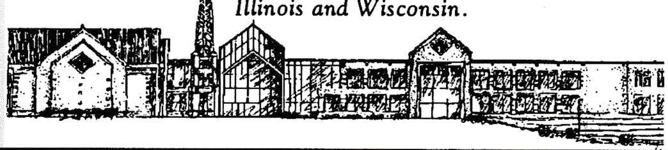
STATE OF THE ART EQUIPMENT AND INSTRUCTIONAL MATERIALS

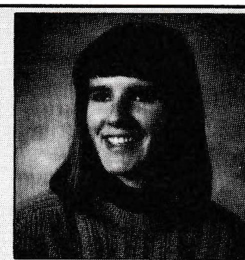
✓ Selected for National Science Foundation Grant for **Multimedia Electronic Classroom Laboratory** for Math, Biology and Computer Science.
✓ Awarded National Science Foundation Grant for the acquisition of **Nuclear Magnetic Resonance** (NMR) instrument to be used in chemical structural analysis.
✓ Computers incorporated into hands-on investigative approach to all science and math courses.
✓ ***Mathematica***, a graphics/symbolic computation program, used in calculus and other math classes
✓ Designated as an **IBM test site** for evaluating prototype hardware and software. Faculty and students involved in research which integrates new computer technology into chemistry analysis.

QUALITY ACADEMIC ENVIRONMENT

✓ Clarke's **12:1 teacher ratio** guarantees attention from faculty who will challenge you to achieve.
✓ Academic excellence rewarded with renewable non-need **scholarships**.
✓ Supports student chapters of the American Chemical Society (**ACS**); Mathematics Association of America (**MAA**); and Association for Computing Machinery (**ACM**) where students compete in national computing programming contest.
✓ Offers a chapter of National Computer Honor Society, **Upsilon Pi Epsilon.**
✓ **Strong undergraduate research programs** in biology developed for on-site study.

Located in scenic northeast Iowa on the bluffs of the Mississippi River bordering Illinois and Wisconsin.





DEANNA ANGST

B.S. in Chemistry and Mathematics; summer internship for ICI Americas, Brussels, Belgium; Teaching Assistantship & Emily Berger Fellowship for Ph.D. in Chemistry, University of Kansas, Lawrence.

UNDERGRADUATE SCIENCE MAJORS & PROGRAMS:

Biology/Molecular Biology
Chemistry
Computer Information Science/Management Information Science
Computer Sciences
Education
Health:
- Medical Technology
- Pre-Chiropractic
- Pre-Dentistry
- Pre-Medicine
- Pre-Pharmacy
- Pre-Veterinary

Mathematics
Nursing
Physical Therapy
3-2 Engineering

✓ Recent graduates offered opportunities for **totally funded** teaching fellowships and stipends for graduate studies at universities throughout the country including
University of Arizona,
University of Iowa,
University of Minnesota,
University of Notre Dame,
University of Oklahoma,
University of Osteopathic Medicine and Health Science,
Virginia Polytechnic Institute and State University,
Washington State University

For more information, call 319-588-6316 OR toll-free 1-800-383-2345 (1-800-255-2255 ext. 6316 U.S. Virgin Islands & Puerto Rico)

Circle No. 7 on Reader Service Card

The talking wave of the future

Words hanging by a thread—of light

by Yury Nosov

FIBER-OPTIC COMMUNICATION is a fundamentally new type of information exchange with virtually inexhaustible possibilities. It's an instructive outgrowth of the scientific revolution, embodying the hopes of humankind for a worldwide communication system. Today there's not a single developed country in the world that isn't working on a fiber-optic system, and every educated person has heard of this marvel of the twentieth century. But it's one thing to know *about* something, another thing to *know* something. So, listen up! I'm going to take you into the what, where, when, how, and why of fiber-optic communication.

It's much easier to learn about science from our own experiences. So imagine diving deep into a river or an ocean. After the water calms down, look back at the world you just left. If you look directly overhead, you'll see the brightness of the summer sky, the sun high above, and, if you're lucky, a sea gull skimming over the water. But if you look up at an angle, you'll see nothing but the surface of the water, which from the inside seems like a murky, fluttering mirror. Your eye can't escape from this "watery prison." And so we've managed to combine the pleasant and the useful: we've had a nice swim and we've gotten acquainted with the physical phenomenon of *total internal reflection*.

One of the most basic of the "optical ABCs" is the fact that a light ray (ray 1 in figure 1) incident at the interface between two different media is split into two rays: a reflected ray and a refracted ray. If the original ray propagates in a medium with a greater refractive index n, the refracted ray is bent toward the interface more than the reflected ray is. Also, rays incident at the interface at very oblique angles do not even form refracted rays—all the energy is reflected back into the medium it comes from (rays 2 and 3 in figure 1). The interface for such rays works exactly like an ideal mirror. This is what physicists call total internal reflection. The medium with the greater refractive index n is called optically more dense. This term is justified by the fact that light propagates in this medium at a slower speed v (we recall that $v = c/n$, where $c = 3 \cdot 10^8$ m/s is the speed of light in a vacuum).

Only one step separates these simple general considerations from the subject at hand. The "classical" optical fiber is a thin, two-layer glass thread. It has a core region with an index of refraction n slightly higher than that of its surrounding sheath (figure 2). Because of this higher refractive index, rays that enter the end of the fiber at a shallow angle to the central axis are reflected back into the core when they strike the interface between the core and the sheath. We can see from geometric considerations that if a ray is reflected back into the core at its first encounter with the interface, it will continue to be confined indefinitely (ray 1 in figure 2). Naturally the trajectory of each

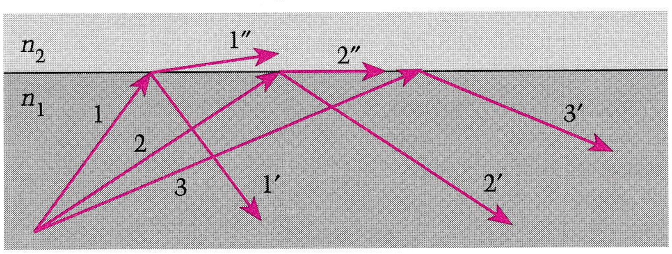


Figure 1
Light rays striking an interface. The two media have refractive indices n_1 and n_2 ($n_1 > n_2$). Ray 1 is split into the reflected ray 1′ and the refracted ray 1″; for ray 2, the refracted ray 2″ is in the plane of the interface between the media; all the energy of ray 3 is transformed into the energy of the reflected ray 3′. All the rays that strike the interface at angles shallower than that of ray 2 meet the condition for total internal reflection.

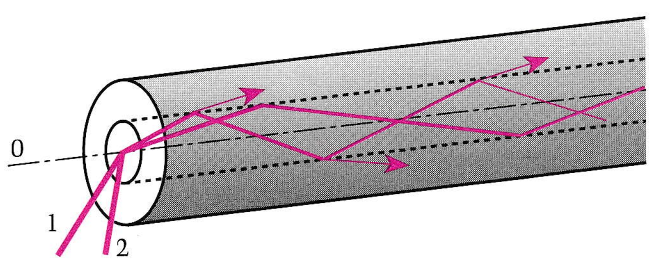


Figure 2
Light traveling along a two-layer fiber optic. Ray 1, owing to total internal reflection, caroms from side to side and is confined to the core of the fiber optic. The out-of-aperture ray 2 gradually loses its energy, passing it to the rays that leak into the sheath.

Art by Dmitry Krymov

ray undergoing total internal reflection is a single broken line, but the energy of all these rays propagates along the fiber's axis.

In this way a flux of light that diverges only slightly and enters one end of the fiber passes through the core without diminution and exits the other end of the fiber. The fiber functions here as a light guide, called a *fiber optic*. One of the popularizers of fiber optics figuratively compared the light ray passing along the optical fiber to a bullet racing through a metal pipe, ricocheting off its walls.

Now we can give some mathematical polish to this physical picture. Consider the geometry shown in figure 3, in which a light ray is incident on the end of the fiber at an angle ϕ' to the axis of the fiber. In your textbooks you usually encounter Snell's law in the form

$$n_1 \sin\theta_1 = n_2 \sin\theta_2,$$

where the angles θ_1 and θ_2 are measured relative to the normal to the interface. But Snell's law can also be stated in terms of the angles that the rays make with the interface itself. Applying Snell's law in this form to the rays at point A, we have the general expression

$$n_1 \cos\phi' = n_2 \cos\phi''.$$

The critical angle ϕ_c for total internal reflection is found by setting $\phi'' = 0$. Therefore,

$$\cos\phi_c = \frac{n_2}{n_1}.$$

Applying Snell's law in the usual form to the rays at point B yields

$$\sin\phi = n_1 \sin\phi'.$$

For the critical case, the angle of incidence $\phi = \phi_0$ is given by

$$\begin{aligned}\sin\phi_0 &= n_1 \sin\phi_c \\ &= \sqrt{n_1^2\left(1-\cos^2\phi_c\right)} \\ &= \sqrt{n_1^2 - n_2^2}.\end{aligned}$$

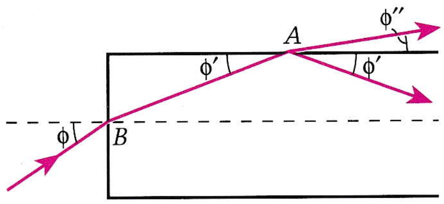


Figure 3
The geometry for calculating the numerical aperture for a fiber optic.

Since the values of n_1 and n_2 are usually close to each other, we can obtain an approximate formula

$$\begin{aligned}\sin\phi_0 &= \sqrt{(n_1+n_2)(n_1-n_2)} \\ &\cong \sqrt{2n_1\Delta n},\end{aligned}$$

where $\Delta n = n_1 - n_2$. The term $\sin\phi_0$ on the left side of this equation is called the numerical aperture. The greater this quantity, the larger the portion of the light flux from a source that is "captured" by a fiber optic. Rays that deviate from the fiber's axis at angles smaller than ϕ_0 are called *aperture rays* (also "channelized" or "directed"). Out-of-aperture rays (for example, ray 2 in figure 2), on the other hand, leak out. At each encounter with the sheath, part of the energy of such a ray is refracted into the sheath until all its energy is lost. If a two-layer fiber optic is bent slightly, this scenario remains valid except for a small decrease in ϕ_0.

This unusual curved propagation of light was first demonstrated as far back as 1870 by the Englishman John Tyndall. You can reproduce his experiment if you like (see figure 4). After what was said above, it's clear that the stream of water is the "core" and the surrounding air is the "sheath" of a light guide, and that it shines because the air–water interface shakes slightly, which leads to disturbances here and there that disrupt total internal reflection so that light can escape from the water.

"That's all very fine," you may be saying. "But arrows in a figure are one thing, actual rays of light are another. Tyndall's experiment merely demonstrates the principle of total internal reflection, not a practical method of transmitting light along curvilinear paths." And you're right: it took scientists a long time to understand that this principle is the best one and to develop the technology to make use of it. Before this they wasted a lot of effort trying to manufacture tubes with polished inner walls or with countless mirrors built in (sort of like a periscope, but more complicated).

The first fiber optics were made in the early 1950s. A cylindrical glass rod was slipped inside a glass tube; the compositions of the glasses in the rod and pipe were selected so as to provide a higher refractive index n for the rod. Then this assembly was heated while being rotated in the flame of a gas burner to form a solid glass rod, which was stretched into a thread as the glass softened. When the rod is stretched, the thickness of the sheath (the tube) decreases as much as the diameter of the core (rod) does. When you first encounter such fibers, you are struck by how they seem to belie their "glassy genealogy," for they are strong as steel and flexible enough to be wound onto a pencil, and they can reliably transmit a flux of light. Of course, a hair-thin fiber can't transmit a great deal of light, so individual fibers are combined into braids, cemented and polished at their ends. The fibers can be assembled into braids either chaotically, without any order, or regularly, according to some rule. If chaotically, the cables obtained can be used only to transport light fluxes (for example, for lighting places with difficult access). Cables with orderly arranged fibers can be

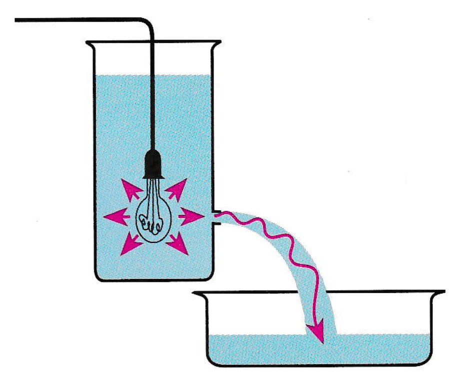


Figure 4
Tyndall's experiment. (If you try to reproduce this experiment, in which light propagates inside a stream of water, make sure that your current leads are protected against contact with the water.)

used to transmit images. And by arranging the fibers in a cable in certain ways, it's possible to process images (superposition, decomposition, transformation, and so on). Some of the applications of glass fibers are well known. No doubt you've seen fiber lamps, or maybe you've been forced to become acquainted (not so pleasantly) with a medical endoscope for examining the inside of the stomach. Fiber optics have also found their way into industry (instrumentation, computers) and into our everyday life. They have become an indispensable part of human civilization in the second half of the twentieth century.

Despite the attractive properties of these cables, they are only "short-range weapons" because a mere 1 to 2 meters of such a cable are enough to attenuate the flux of light by a half. One can easily calculate that only a millionth of the energy put into a 100-meter-long cable reaches the other end—there's no point thinking of a communication system with such losses. However, dreamers aren't deterred by the obvious, and scientists persisted in searching for ways to solve this problem. An added impulse was provided by the arrival of lasers in 1960. With their powerful, sharp light beams, lasers seemed specially made for optical communication. And that's how it began: a laser was placed on the roof of one skyscraper and a photoreceptor on the roof of another. Voilà—an optical telephone! The problem was, it worked only when there was nothing in the way: fog, rain, snow, buildings, hills . . . Besides, such a telephone could not transmit signals over large distances. No one played with this toy for long—it was too costly.

But in 1966 an Englishmen by the name of Kao combined (in his head) a laser and a fiber optic, and he predicted the possibility of a new type of communication if only . . . if only the fibers were a thousand times more transparent.

Let's make a detour to get acquainted with some numerical estimates. In radio, acoustics, and communication science a dimensionless unit—the decibel (dB)—is used to compare two values of the same physical characteristic. A decibel is defined as 10 times the logarithm of the ratio of two power levels. Applied to fiber optics, it may be defined as

$$B = 10\log\frac{P_{\text{in}}}{P_{\text{out}}},$$

where P_{in} and P_{out} stand for the power of the light at the entrance and exit, respectively, of the fiber optic. It's convenient to use this kind of characteristic of signal attenuation (that is, a logarithmic one) when designing complex systems, since it replaces the operation of multiplication with the simpler operation of addition. Here's an example. Let a transmission line consist of three sections. The first section attenuates the signal by a factor of 10; the second section reduces the power level of the signal by a factor of 1,000, and the third section by a factor of 2. The whole line attenuates the signal by a factor of $10 \times 1{,}000 \times 2$, or by $10 + 30 + 3 = 43$ dB. An additional advantage is that the new characteristic helps scientists avoid large numbers, which they generally don't like. The index of fiber transparency is the attenuation of a light signal per unit length of fiber $b = B/l$, measured in decibels per kilometer.

So in effect, Kao was asking for a transparency index $b \le 20$ dB/km. Typical fiber-optic cables had an index $b = 1{,}500$ dB/km, and normal window glass has an index $b = 100{,}000$ dB/km. Fortunately, we live in technological times, which means that any problem correctly posed gets solved, sooner or later.

In 1970 Corning Glass, an American company, announced that it had created a fiber with a transparency index $b \cong 16$ dB/km. Talk about a sensation! So the year 1970 marks the beginning of the era of fiber-optic communication.

The company had given up on the idea of using multicomponent glass for this purpose, since it showed no prospects, and concentrated its efforts on quartz (silicon dioxide, or SiO_2). Corning scientists developed a technique whereby a mixture of tetrachlorosilane ($SiCl_4$) and oxygen is fed through a heated quartz tube. In the high-temperature zone (where $T = 1{,}400$°C) this mixture undergoes a reaction to form extremely pure quartz, which uniformly precipitates onto the walls of the tube (fig. 5). By adding other reagents containing boron, phosphorous, germanium, and fluorine to the mixture, you can "dope" the quartz to precisely control its refractive index. After a quartz layer of the required thickness and composition has precipitated on the walls, the tube is reheated to collapse it into a solid rod, or "blank." This blank is then softened by heating and drawn through a system of dies until it forms a fiber several kilometers long. The fiber is typically about 0.125 mm in diameter, and the diameter of its core is about 0.05 mm. In the final stage, the fiber is drawn through an extruder with a polymer

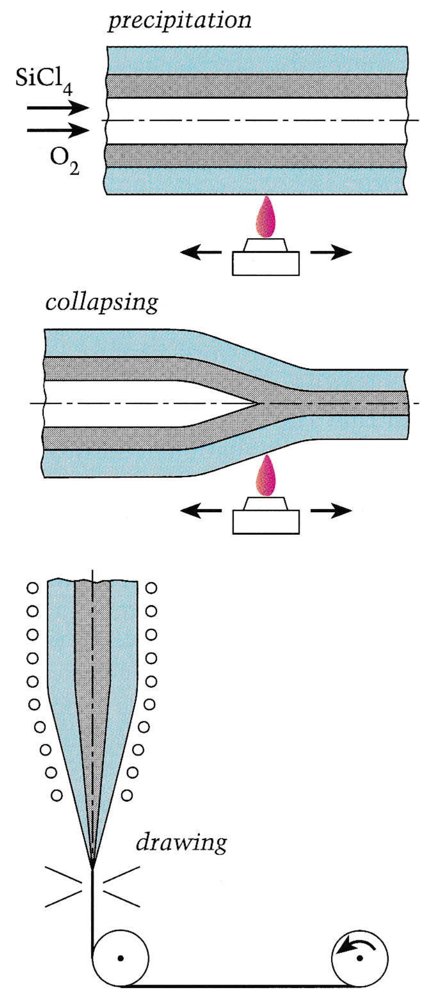


Figure 5
Stages of manufacture of two-layer quartz fibers.

melt and leaves the extruder coated with a protective jacket. Finally, the fibers—a few or hundreds—are twisted together with strengthening fibers and coated with several layers of polymer, and the fiber-optic cable is ready for action.

On the face of it, the process is quite logical and simple beyond words. But in fact, the machinery needed to manufacture the blanks is a highly complex, computerized set of equipment including a gas purification system that ensures "semiconductor" purity and sterility[1] as well as dozens of instruments for controlling and regulating the process. You'd be even more impressed if you saw the three-story building where the blanks are drawn out to form the fibers. When you watch the operators in protective blue overalls ascending to the upper level in a vibrating telescopic elevator, it makes you think of astronauts about to get into their spacecraft.

The success of this quartz technology triggered a "steeplechase" among scientists, engineers, and company managers. It seemed that every three months a new fantastic record was set, only to be forgotten after another three months. Finally, the transparency index stopped dead at 0.14 dB/km. Two kilometers of this fiber is more transparent than a freshly washed window! A lot of work went into this achievement. Early on it was discovered that the redder the light, the smaller the attenuation in the quartz. The situation is even better in the invisible infrared region (fig. 6). The absorption peaks can be explained by the inevitable presence of trace water in the quartz. They're not crucial, though, since one can always tune the wavelength to a valley between the absorption peaks. It was in one of these valleys, at $\lambda \cong 1.55$ µm, that the record transparency of 0.14 dB/km was achieved, although for almost a decade all research was done in the region $\lambda \cong 0.85$ µm, and scientists contented themselves with a transparency

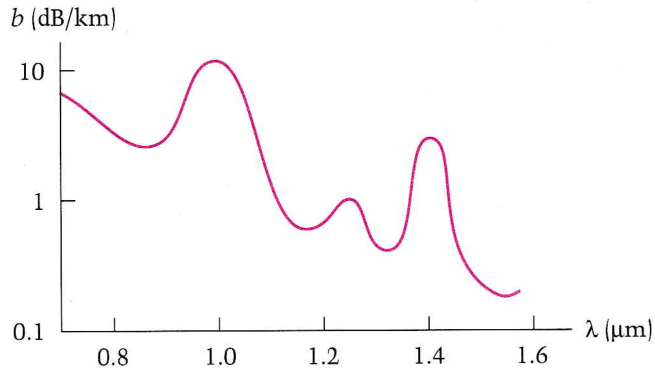


Figure 6
Absorption spectrum of quartz fiber. Such a spectrum is usually plotted on a semilogarithmic scale because of the pronounced and distinctive dependence of its transparency on the wavelength of the optical radiation.

of $b \cong 2$ dB/km. Why? Here we come upon problems that lie outside the fiber but are closely linked with it.

A fiber optic, in and of itself, is not a communication line. You also need a generator and a detector of light signals—in other words, a transmitter and a receiver. When superpure fibers appeared in the early seventies, semiconductor technology dominated the electronics industry. When "partners" were needed for the fiber, it was only natural to choose a laser and a photodiode—both semiconductors, naturally.

The semiconductor lasers of the early seventies—even the best of them (heterojunction lasers)—weren't as developed as the transistors and integrated circuits of the time. They were not efficient, not coherent,[2] and not durable. The only purpose they served was to validate certain general principles of quantum electronics; they were absolutely useless for technical applications. The development of fiber-optic communication logically required a technological revolution in the world of lasers. And that revolution in fact occurred. Step by step, all the "nots" in the list above were eliminated. It happened first with heterojunction lasers radiating at a wavelength $\lambda \cong 0.85$ µm. These lasers already existed in the "prefiber" period. Then heterojunction lasers operating at $\lambda = 1.3$ µm to 1.55 µm were developed specifically for communication purposes. In parallel with (sometimes, even ahead of) improvements in laser technology, photodetectors were also significantly improved.

In addition to these main elements, a communication system needed a host of subsidiary components. The laser had to be turned on and off quickly, so special excitation circuits were invented; the signal received from the photodiode is generally too weak to be processed any further and needs to be amplified and transformed, so a whole series of highly sensitive integrated circuits was developed; the fibers, transmitters, and receivers had to be combined into a system, so optical connectors, splitters, and switches were devised.

Taken together, all these elements formed the basis of a new branch of technology, and fiber-optic communication sprang to life. Already in 1972 a ten-kilometer fiber-optic line went into operation in England. Several years later a 1,000-kilometer system was inaugurated on the east coast of the US. The newspapers bubbled over with reports on advances in fiber-optic communication systems.

It soon became evident, however, that achieving a small attenuation of the optical signal, which makes long lines possible, is only half the battle. The rallying cry of the "fiber steeplechase"—more transparent, still more transparent, still more!—didn't tell the whole story.

Information is transmitted along the fiber in short flashes of light, or pulses. The higher the number of pulses transmitted over a line per unit time, the greater the capacity of the line, which is measured in bits per second. Let's turn back to figure 2. If the fiber is transparent enough, ray 1

[1]That is, the kind of clean atmosphere required for the manufacture of semiconductors and other computer components.—*Ed.*

[2]A light source is called coherent if it produces light waves whose phases correlate in time and/or space.

(which undergoes multiple total internal reflections and caroms from side to side) and ray 0 (which propagates in the fiber along its axis and normally travels a shorter distance) will both reach the other end of the fiber. No energy was wasted, and that's great. The only problem is, these two rays will not reach the other end simultaneously, because their paths are different. This will cause the pulse to spread out: the rectilinear light pulse, strictly defined in time, blurs and turns into a bell-shaped pulse. The adjacent pulses will overlap and partially or even completely merge with one another, which leads to distortion of the information transmitted over the fiber.

There is yet another mechanism that causes the pulses to spread. Light waves of different wavelengths travel in quartz (as in most other substances) at slightly different speeds. Although a laser is considered a source of monochromatic light, it nevertheless emits light in a narrow range of wavelengths. All these phenomena—the nonuniformity of light propagation, which is caused by the fiber geometry, and the physical properties of quartz—are combined in the term "dispersion of the fiber optic." Dispersion is very undesirable in communication systems, and its significance increases with the length of the fiber-optic line. Worst of all, it appeared to be impossible in principle to eliminate dispersion in fiber optics. "Don't ride your horses so hard!" came the jeers from the sidelines—from those who preferred to stick with traditional means of communication. But the riders in the fiber-optic steeplechase didn't rein in, they went to the whip: they were determined to unravel the intricacies of dispersion. They developed a unimodal fiber with an extremely thin core. In this fiber the rays travel almost strictly parallel to the fiber's axis, and so they got rid of the first component of dispersion. Then they jumped over the ravine with λ = 1.55 μm (fig. 6) and got rid of the second source of dispersion, since quartz is all but indifferent to wavelength in this region.

The switch from rather inefficient first-generation lines (λ = 0.85 μm) to second-generation lines (λ = 1.55 μm) was a turning point in the history of fiber-optic communication. It occurred in the early eighties and marked the beginning of the triumphant procession of fiber optics through the boundless fields of the communications world. For example, here are the characteristics of one of the fiber-optic lines of that time: the amount of information it could transmit per second in coded digital form is equivalent to the text of all the volumes of the *Encyclopædia Britannica*. Signal regenerators were 160 km apart; today this distance is closer to 300 km, and the 1,000-kilometer mark will probably be reached in the near future! Also, the line capacity mentioned above is far from being the limit—fibers have already been created in laboratories that can transmit several million phone calls simultaneously. It's hard to believe that all this is being done by a glass fiber one-tenth the thickness of a human hair.

It didn't take long for trunk lines of fiber optic cables to be laid. The trans-Atlantic line TAT-8, connecting the US and Great Britain, has been operating for several years now; a fiber-optic semicircle connects ten countries of the Mediterranean region; and fiber optics have brought Japan, the US, Australia, and New Zealand closer together. And these are just a few examples of the longest fiber-optic trunk lines.

Fiber optics are convenient for short-distance communication lines as well. They can be used in internal communication networks in factories and offices and on board airplanes and ships. Using glass fibers, you can transmit information between computers. The replacement of metallic telephone cables with fiber-optic cables is paving the way for cable television, including high-definition TV.

Still another remarkable feature becomes apparent in short fiber-optic lines: they are "electrically hermetic." This simply means that fiber-optic lines are completely immune to electromagnetic interference from lightning, sparks from electrical devices, radio waves, "cross-talk" from nearby metallic cables, and so on. Sometimes it is this one advantage of fiber over metal cables that causes engineers to lay a fiber-optic line. If the line is to be in an electromagnetically dirty environment, they really have no choice.

On the other hand, fiber optics do not emit any radiation, so you can't eavesdrop on a fiber-optic line. Nowadays, when information is the most valuable commodity around, its proprietors are prepared to pay a lot for protection against unauthorized access.

There's still more. Fiber-optic lines can be made extremely light and compact—nearly one-hundredth the weight of metallic lines. This allows fundamentally novel solutions to many aeronautical problems—for example, the communication line to a spacecraft in the initial stage of its flight, when radio communication is hampered by powerful launch interference. After fulfilling its function, the fiber line is simply torn away.

Another application that would be easy to implement is to "tie down" a meteorological balloon that rises to an altitude of 2–3 km. Why not, if a kilometer of a fiber-optic cable weighs less than a kilogram and can withstand shear loads of 30–40 kg?

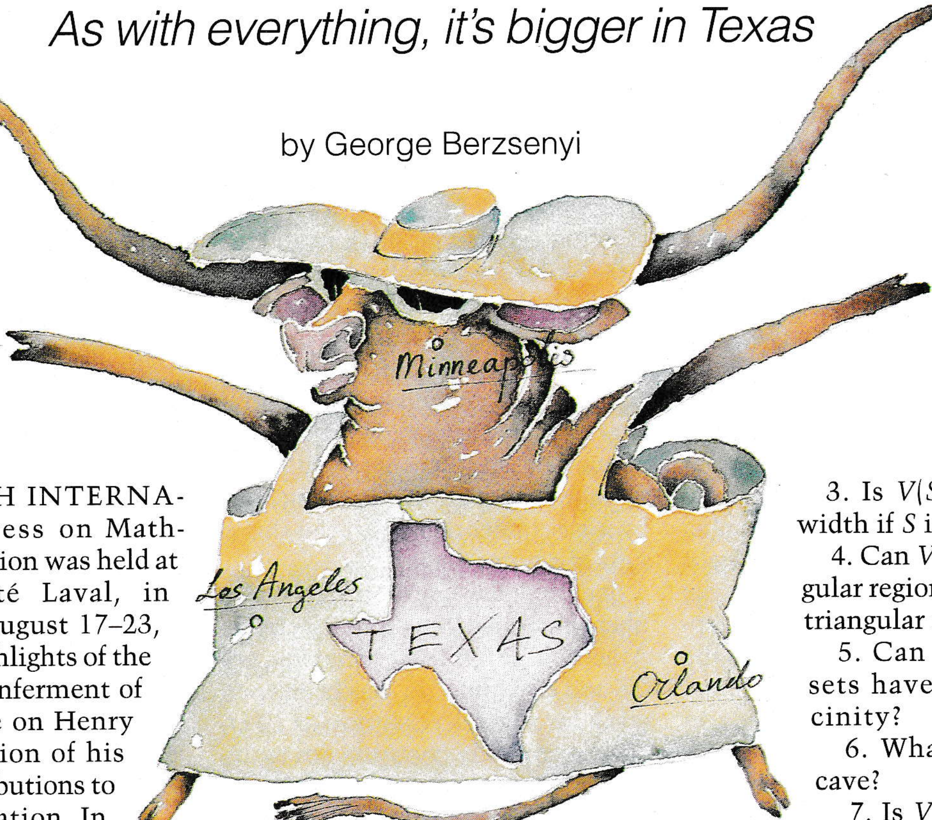
Ultimately, quartz fibers should be quite cheap, since the raw material for this product is ordinary sand. Right now, it's true, optical fiber is much more expensive than copper wire, but this is to be expected when a technologically complex new product is first put into production. So we can only speak of potential low cost. But then, metal cables made of copper and lead will never be exactly cheap.

Fiber-optic communication systems, with all their remarkable properties, initiated a technological boom comparable in its breadth and duration to the explosions caused by the personal computer, VCR, and television. One of the proponents of this new type of communication places it on the same level as the invention of the steam engine, the electric light bulb, the transistor . . .

The notion of vicinity

As with everything, it's bigger in Texas

by George Berzsenyi



THE SEVENTH INTERNATIONAL Congress on Mathematics Education was held at the Université Laval, in Québec, Canada, August 17–23, 1992. One of the highlights of the program was the conferment of an honorary degree on Henry Pollak, in recognition of his distinguished contributions to mathematics education. In this column, I'll share with my readers a really nice problem area he described to me some years ago. When I asked him about it in Québec, he assured me that as far as he knows, this problem area is still awaiting serious mathematical investigations.

In 1987, as the chairman of the Texas section of the Mathematical Association of America, I had invited Henry to present a minicourse to the members of the section. Following his presentation, we talked about the size of Texas, and he remarked that Minneapolis is in the vicinity of Texas in the sense that there are two cities in Texas (notably, Texline and Brownsville) that are further from one another than Minneapolis is from one of them.

More mathematically, one may say that a point P is in the vicinity of a set S if there are two points X and Y in S such that the distance $d(X, Y)$ between X and Y is greater than or equal to $d(P, X)$. We will denote by $V(S)$ the set of all points in the vicinity of S.

Even if we restrict our attention to planar sets, there are many interesting questions which readily come to mind. Some of these are listed below as possible appetizers.

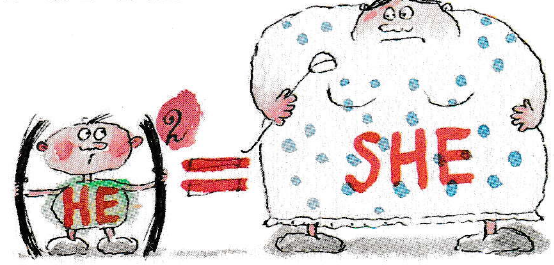
1. What is $V(S)$ if S is a segment? What if S is a circular, square, or triangular region? What if S is semicircular or elliptic?
2. What is $V(V(S))$ for each of the regions S named above?
3. Is $V(S)$ of constant width if S is?
4. Can $V(S)$ be a rectangular region for some S? A triangular region?
5. Can two different sets have the same vicinity?
6. What if S is concave?
7. Is $V(S)$ a subset of $V(T)$ whenever S is a subset of T?

One can also explore the notion of vicinity in three dimensions, in discrete sets, with respect to different measures of distance (for example, in "taxicab geometry"[1]), as well as in more general settings. The possibilities are endless, and I look forward to learning more about them from my readers.

[1]If you're not acquainted with this simple non-Euclidean geometry, you might look for *Taxicab Geometry* by Eugene F. Krause (New York: Dover Publications, 1986) in your library or bookstore.

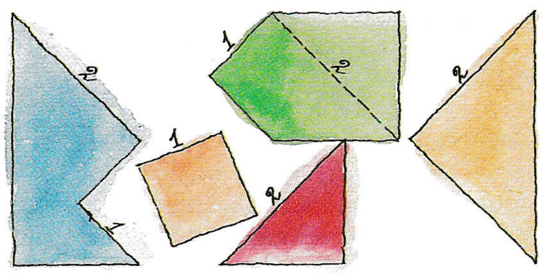
Art by Sergey Ivanov

Just for the fun of it!



B66

A feminist equation. Solve the number rebus $SHE = (HE)^2$, in which the same letters designate the same digits, different letters denote different digits. (A. Savin)



B67

Making squares. Take a piece of cardboard and cut out the polygons at left. Then try to fit the pieces together to make a square (1) using each piece except the small square once, (b) using each of all five pieces once, (c) using each of the five pieces twice. (The numbers 1 and 2 in the figure denote the relative dimensions of the pieces.) (V. Dubrovsky)

B68

Musical thermos. When we fill a thermos with water, we hear a sound. How will its tone change while the thermos is being filled? (A. Buzdin, S. Krotov)

B69

In search of special pairs. Find the smallest positive integer such that the sum of its digits and that of the subsequent integer are both divisible by 17. (G. Galperin)

Art by Pavel Chernusky

B70

Coins on a checkerboard. A number of coins are placed on each square of a checkerboard such that the sums on every two squares having a common side differ by one cent. Given that the sum on one of the squares is 3 cents, and on another one 17 cents, find the total amount of money on both diagonals of the checkerboard. (V. Proizvolov)

ANSWERS, HINTS & SOLUTIONS ON PAGE 57

А
я
в

The wonderland of Poincaria

"I have described an imaginary world whose inhabitants would inevitably have come to create Lobachevskian geometry."
—Henri Poincaré

by Simon Gindikin

WHEN THE STORY OF Lobachevskian (or hyperbolic[1]) geometry is recounted nowadays, one might get the impression that non-Euclidean geometry would have been favorably received if only its creators had proven its consistency. But what disconcerted the critics in the first place was not the absence of such a proof. People were used to thinking that geometry deals with the real space around us and that this space is correctly described by Euclidean geometry. It is significant that Gauss set geometry apart from other branches of mathematics, considering it an experimental science like mechanics. But in so doing, Gauss, as well as Lobachevsky and Bólyai, was aware that, first, there's room for logically coherent geometric constructions that do not reflect physical reality—"imaginary" geometries; and, second, it's not so unquestionable that Euclidean geometry is the one that applies in our world on the astronomical scale.

But this insight, perceived by only a few mathematicians, lay entirely beyond the grasp of the public at large. They measured the assertions of Lobachevskian geometry by the yardstick of their geometric intuition—and gained an inexhaustible source for exercising their wit. Nikolay Chernyshevsky (a writer and literary critic, one of the leaders of revolutionary movement of 1860s in Russia) wrote to his sons from exile that the whole city of Kazan was making fun of Lobachevsky: "What is this 'curvature of a ray' or 'curved space'? What is this geometry without the axiom of parallel lines?" He compares it to "squaring boots" and "extracting the roots of the boot tops," and says this is as absurd as "writing in Russian without verbs" (this is a gibe at the great Russian poet Afanasy Fet, with his famous poem written without a single verb: "Whispers, timid breathing, the trills of nightingales . . .," which also apparently "kept them in stitches").

A new stage in the development of non-Euclidean geometry came with the construction of its first models.[2] Today, we take these as a means to prove the consistency of hyperbolic geometry; but they were notable not only for this. Even considered favorably, Lobachevskian geometry seemed too sophisticated, divorced from the rest of mathematics, while the Caley–Klein model (see the article "The Dark Power of Conventional Wisdom" in this issue) showed that it quite naturally emerges from projective geometry, which was very popular at the time! On the other hand, an examination of a model whose basic concepts were built from the notions of our habitual Euclidean geometry made it possible to replace a formal axiomatic presentation of non-Euclidean geometry with a more visual one.

Another model was devised by Henri Poincaré while he was exploring some purely analytical problems of the theory of functions of a complex variable. Unexpectedly, he discovered that transformations appearing in his work could be interpreted as "displacements" (to be more accurate, *isometries*) in the hyperbolic plane. This discovery made such a strong impression on him that he recalled long afterward how it had come into his mind—"without, it seems, being preceded by any thoughts on the topic," as he mounted the footboard of an omnibus during an excursion to Coutances. Ten years later Poincaré made a remarkable addition to his model by providing its "physical" background. And that is the subject of this article.

Excursion into physics

Our geometric concepts have physical prerequisites. For example,

[1]The internationally accepted name for this kind of non-Euclidean geometry is "hyperbolic"; Russian authors, for obvious reasons, call it "Lobachevskian," while Hungarians prefer to attribute it to their compatriot Bólyai.—*Ed.*

[2]You can find more on models in the article "Turning the Incredible into the Obvious" in the last issue of *Quantum*.—*Ed.*

Art by Dmitry Krymov

we perceive light beams as straight lines. A light beam striking your eye seems to be a straight line even if it has been refracted on its way (say, coming from the air into the water). To dispel this illusion one must set up an experiment, or look at what's happening from a different perspective.

Suppose the upper half-plane ($y > 0$) is filled with an optically heterogeneous medium in which the speed of light $c(x, y)$ at point (x, y) varies according to the law $c(x, y) = y$ (no matter what the beam's direction). By Fermat's principle, the path of a light ray between two points is the one that takes the least time. It can be shown that in our medium, where $c(x, y) = y$, the light will propagate between two points along the curves L (fig. 1) such that

$$\frac{\sin\alpha(y)}{y} = k, \qquad (1)$$

where $\alpha(y)$ is the angle between the tangent to L at the ordinate y and the vertical, and k is the same for all the points of L.

Formula (1) must ring a bell for readers familiar with Snell's law of refraction at the interface of two media: $\sin\alpha_1/\sin\alpha_2 = c_1/c_2$, where c_1 and α_1 are the speed of the incident light ray and the angle it makes with the perpendicular to the interface, respectively; and c_2 and α_2 are similar values for the refracted ray. Poincaré's medium may be visualized as comprising an infinite number of infinitely thin horizontal layers with a speed of light y in the layer at a height of y. Snell's law applied to the interface of two adjacent layers yields condition (1). Condition (1) is valid for any circle whose center lies on the x-axis (that is, perpendicular to this axis). For any such circle, $k = 1/r$, where r is its radius. For $k = 0$ we get vertical lines. One can prove that no other curve satisfies condition (1); this fact has a physical explanation, too—for instance: the light propagates from a given point in a given direction along a unique path. (Snell's law is used in "The Talking Wave of the Future" and the Physics Contest solution in this issue and in "A Snail That Moves Like Light" in the September/October 1991 issue.)

The circles perpendicular to the x-axis and vertical lines (more exactly, their halves in the upper half-plane)

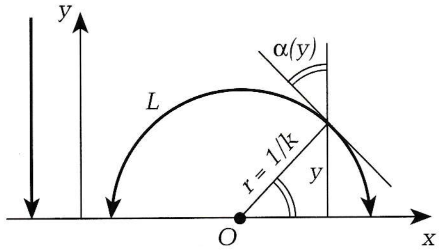


Figure 1

will play the leading roles in our story.

"Poincaria" and its geometry

Poincaré's world (I'll name it *Poincaria*) is the upper half-plane $\{(x, y) : y > 0\}$ without the border $y = 0$ (this is important!).[3] *Poincarians*, the creatures inhabiting Poincaria, perceive upper semicircles with centers on the x-axis and vertical rays as "straight lines" (fig. 2). We'll designate them as *p-lines*. *P*-lines seem infinite to Poincarians (it takes an infinite amount of time for light to travel along an entire p-line); the endpoints of p-lines—and the entire x-axis as well—are invisible. So Poincarians think their land extends to infinity in every direction. We'll call the invisible points of a p-line its *infinitely distant points*. For a p-line represented by a vertical ray, one of these is ∞—a special "point at infinity" added to the plane and common to all such p-lines. A p-line is uniquely determined by its pair of infinitely distant points (why?); this enables us to denote it by $L(a, b)$, where real numbers a and b are the coordinates of two infinitely distant points on the x-axis (one of them may be ∞).

Together with the Poincarians, let's try to construct the geometry of their space. Just as with us, who have lived in Euclidean space, some statements seem obvious to Poincarians—they accept them without proof, as axioms, and deduce more involved statements (theorems) from them. Looking at Poincaria from the outside, we'll see all these statements differently (for instance,

[3]One could consider the three-dimensional world as well, but it's easier to draw pictures in the plane, so for this reason we'll deal exclusively with "flat" creatures.

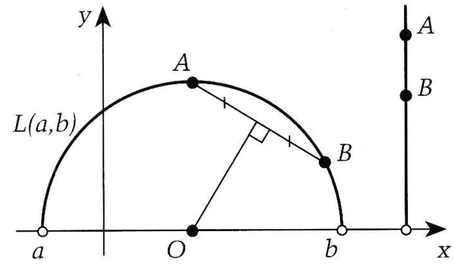


Figure 2

their p-lines are our semicircles or rays), so we'll translate Poincarian formulations into our "prosaic" Euclidean language and prove them in our own way.

For example, Poincarians know there's a unique p-line through any two distinct points. For us this means that through any two points of the half-plane there is a unique semicircle perpendicular to its border or a vertical ray (prove it using figure 2). Notice that the physical explanation of this statement, based on the fact that light travels between two points along a single path, is the same for Poincarians and for us. (This explanation is not a proof, though, from the geometric point of view.) It's easy to check that all the axioms of Euclidean geometry regarding the relative position of points and lines and the order of points on a line are valid in Poincaria, too.

To get used to Poincaria, consider two *p-half-planes* into which a p-line divides Poincaria; verify that the endpoints of a *p-segment* lie in two different p-half-planes sharing a border p-line if and only if the p-segment crosses the bordering p-line; draw p-triangles and p-polygons; think about p-convexity, if you're familiar with "normal" convexity. (Use figures 3a and 3b.)

The distinction between the geometry of Poincaria and Euclidean geometry becomes manifest when we look at the relative position of two p-lines. We already know that two different p-lines can intersect in no more than one point. If they do not intersect, they may either have a common infinite (invisible) point or no common points at all, even on the invisible border. In the first case we call them *parallel*; in the second case—*superparallel*. Given a p-line $L(a, b)$, there are only two p-lines (corre-

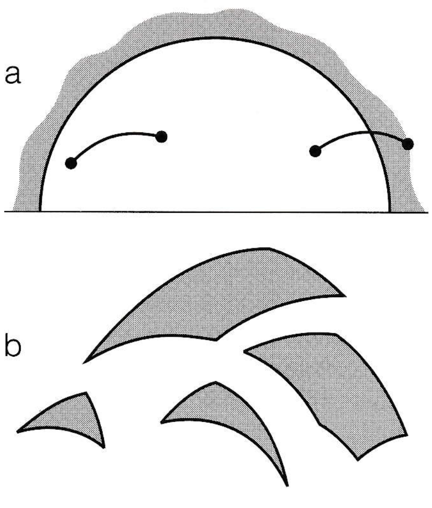


Figure 3

sponding to infinitely distant points a and b) through a point outside it that are parallel to $L(a, b)$ (fig. 4) and an infinite number of superparallels lying "between" the two parallels. Thus, the axiom of parallels does not hold in Poincaria. (For us, the observers, this comes as no surprise—we know what Poincarians don't: that their "lines" aren't "genuine"!) This allows us to hope that the geometry of Poincaria will turn out to be hyperbolic.

The main thing we have to do now is to define what "distance" and "isometry" mean in Poincaria.

Distances and isometries

From the optical point of view the most natural measure of distance between two points A and B in Poincaria is the time it takes light to travel from A to B: a p-line will then indeed be the shortest line between its points. It follows from physical considerations that the distance $\rho(A, B)$ thus defined has the following important properties, which it shares with the usual Euclidean distance:

(1) $\rho(A, B) = \rho(B, A)$;
(2) If B lies on a p-segment AC, then $\rho(A, B) + \rho(B, C) = \rho(A, C)$ (the light goes from A to C along a p-line and passes through point B on its way);
(3) For any points A, B, C, the *triangle inequality* $\rho(A, B) + \rho(B, C) \geq \rho(A, C)$ holds, and the equality here is valid only when B lies on p-segment AC (because the light's path from A to C by way of B cannot take less time than the shortest "straight" path AC).

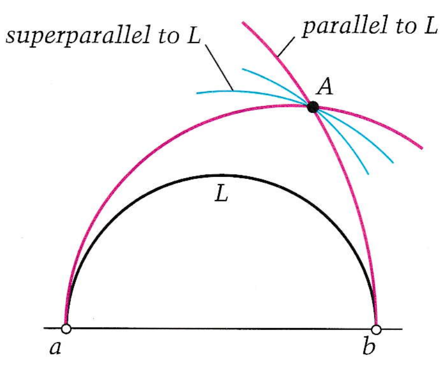


Figure 4

This distance ρ is the primary one for Poincarians (note that light propagates at unit speed relative to this distance), so they have no reason to express ρ in terms of something else. But *for us* it's natural to express ρ in terms of our Euclidean distance. This isn't very easy: we have to deal with nonuniform propagation of light, so to compute the travel time we need to calculate integrals. I'll give just the final expression:

$$\rho(A, B) = \ln\frac{r' + r}{r' - r}, \qquad (2)$$

where $r = AB$ (Euclidean distance) and $r' = AB'$, B' being the reflection of B in the x-axis; the logarithm is taken over the base e (a different base yields a value of ρ that differs by a constant factor). Euclidean distance is remarkable for the numerous transformations that preserve it—the term for such a transformation is *isometry*. Let's see how isometries look in Poincaria. These *p-isometries* are transformations that preserve the p-distance ρ and, consequently, take p-lines into p-lines.

Let's begin with isometries that don't leave any points in place. First of all, these are ordinary *translations* along the x-axis: $T_a(x, y) = (x + a, y)$. They preserve both Euclidean distance and the speed of light $c(x, y) = y$. Thus the travel time of light between any two points A and B—that is, the p-distance $\rho(A, B)$—remains unchanged. Translations take p-lines into p-lines, of course. On the other hand, the *dilation* $D_b(x, y) = (bx, by)$, $b > 0$, with center $(0, 0)$, which multiplies both Euclidean distances and the speed of light $c(x, y) = y$ by the constant b. It is not surprising, then, that it also retains the travel time of light, or p-distance. Formula (2) and a little algebra will verify this. So what looks a dilation (with its center on the x-axis) to us appears to be an isometry to Poincarians. Using these two kinds of p-isometries, one can move any point into any place. For instance, point (x_0, y_0) is taken into $(0, 1)$ under the p-isometry $((x - x_0)/y_0, y/y_0)$.

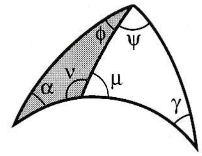
The set of all p-lines falls into two classes with respect to the p-isometries introduced—we'll call them *p-shifts*. The p-shift of a semicircle is always a semicircle, and the p-shift of a ray is always a ray (why?). But we cannot, by means of p-shifts, transform p-lines of one sort (semicircles) into those of the other sort (rays). So let's add another kind of p-isometry—*p-reflection* in a p-line. For a p-line that is a ray, this is a regular line reflection; for a semicircle, it's an inversion.

For instance, the p-reflection in p-line $L(-1, 1)$ is the inversion in the circle with center $O = (0, 0)$ and radius 1: by definition, it takes any point $A \neq O$ into point A' on ray OA such that $OA' \cdot OA = 1$. A detailed discussion of this transformation and its properties can be found in the article "Inversion" in the last issue of *Quantum*. These properties will be used in what follows without special comments and references, for the most part. In that same article you can find another version of the Poincaré model. Problem 1 below explains the connections between the two versions.

Inversion takes circles and lines into circles or lines and preserves the angles between them. In the language of Poincaria this implies that the p-reflection of p-line $L(a, b)$ in p-line $L(-1, 1)$ is p-line $L(1/a, 1/b)$. In particular, p-line $L(a, 0)$ (a semicircle for $a \neq \infty$) is taken into p-line $L(1/a, \infty)$ (a ray). So p-reflection maps Poincaria onto itself and maps p-lines into other p-lines. Also, it can be verified that p-reflections do not change the p-distance ρ. (In fact, it can be shown that any transformation of Poincaria that takes p-lines into p-lines preserves ρ—there are no dilations; and this is one of the principal distinctions between hyperbolic and Euclidean geometry.)

Combining p-shifts with p-reflections, we can create new types of p-isometries that then suffice to take any p-line into any p-line, and even to fit a p-ray given on the first p-line onto a p-ray on the second p-line (prove it!). It can be shown that *any* p-isometry can be represented as a combination of successive p-shifts and p-reflections.

Using p-isometries, Poincarians can "lay" a p-segment AB over p-segment A_1B_1 of equal p-length (when p-ray AB is fitted onto p-ray A_1B_1, point B automatically hits B_1 because $\rho(A, B) = \rho(A_1, B_1)$). Similarly, p-triangle ABC can be laid over p-triangle $A_1B_1C_1$ if they satisfy the side–angle–side (SAS) or angle–side–angle (ASA) condition of congruence. (Notice that the Poincarian measure of angles is the same as their Euclidean measure, because the latter is preserved under p-isometries. You can get an idea of p-congruence from figure 5.) Thus, Poincarians can prove the SAS and ASA properties of triangles in much the same way we do. As for the side–side–side (SSS) test, *our* usual proof involves the compass-and-ruler construction based on the fact that two circles intersect in no more than two points. Fortunately, p-circles turn out to be just the plain Euclidean circles (lying entirely in the upper half-plane) although their p-centers do not coincide with their Euclidean centers. (This is not entirely simple—see problem 2 below.) So the SSS test for congruence of triangles in Poincaria presents no problem either. But there's one more test—the angle–angle–angle (AAA) test in Poincaria: *two p-triangles are congruent if their corresponding angles are congruent*! (See problem 7.) In particular, the area of a triangle in Poincaria is determined by its angles α, β, and γ. In hyperbolic geometry the sum of a triangle's angles is *less* than π (problem 6). The difference $\pi - (\alpha + \beta + \gamma)$ is called the *angular defect of the triangle*. It behaves like an area—that is, if a triangle is cut into two triangles by a line through its vertex (fig. 6), then both its area and its angular defect will be equal to the sums of the areas or the angular defects, respectively, of the two pieces obtained. It can be deduced from this fact that the area of a triangle in hyperbolic geometry is proportional to its defect.



$$\pi - (\alpha + \beta + \gamma) = [\pi - (\alpha + \phi + \nu)] + [\pi - (\mu + \psi + \gamma)]$$

Figure 6

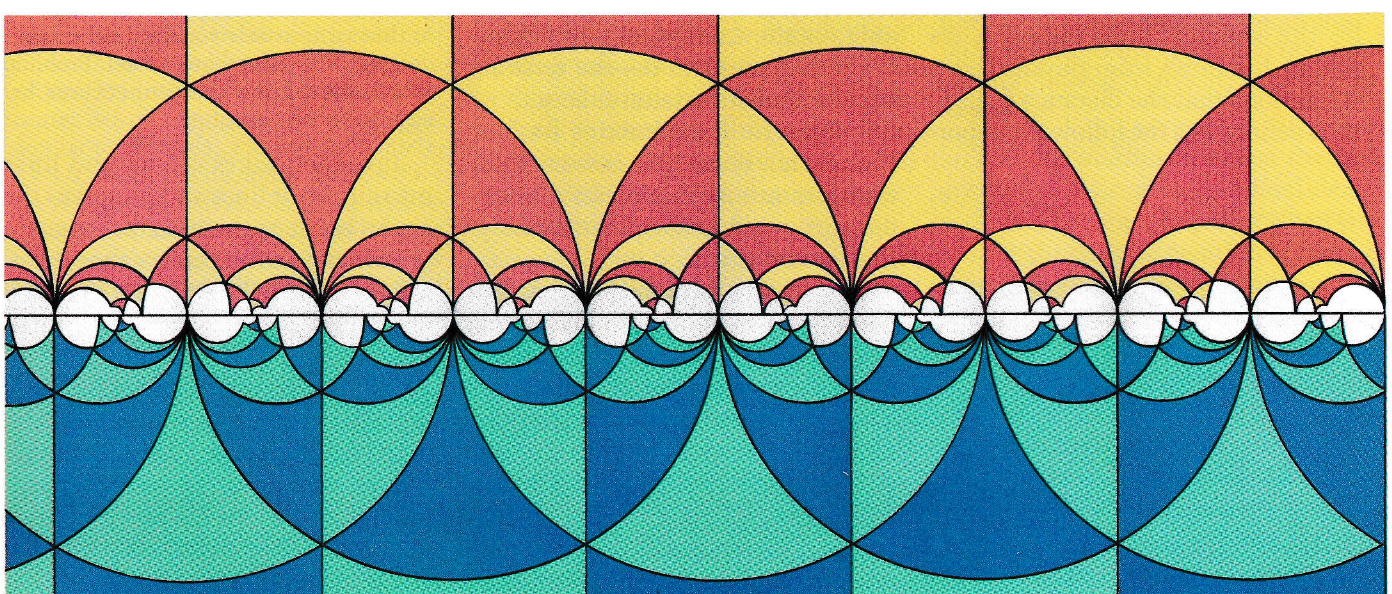
So you see that in some respects Euclidean and non-Euclidean (hyperbolic) geometries are quite similar to one another, while in other respects they are totally different or even opposite. I suggest that you continue your exploration of Poincaria by taking a crack at the problems in the final section.

Solid objects in Poincaria

So far we have been guided by optical prerequisites in all our geometrical considerations. It should be emphasized that the reason Poincarian geometry proved to be non-Euclidean is not that the optical laws differ from ours: Poincaria is constructed ("simulated") in our own world, and we do not change the laws of physics! The optical illusions of Poincarians are due to the optical heterogeneity of their world.

Figure 5

The so-called "modular figure" discovered by Gauss that plays an important role in the theory of functions of a complex variable and number theory. From the Poincarian point of view this is a tiling of two copies of a hyperbolic plane with isosceles p-triangles whose legs are parallel and whose base angles are 60°. (Adjacent triangles are p-symmetric about their common side.) As a matter of course, there are no such triangles in the Euclidean plane. Approaching the border between the half-planes, the p-triangles keep getting smaller and smaller in their Euclidean (but not Poincarian!) size. Try to figure out how to draw this figure yourself!

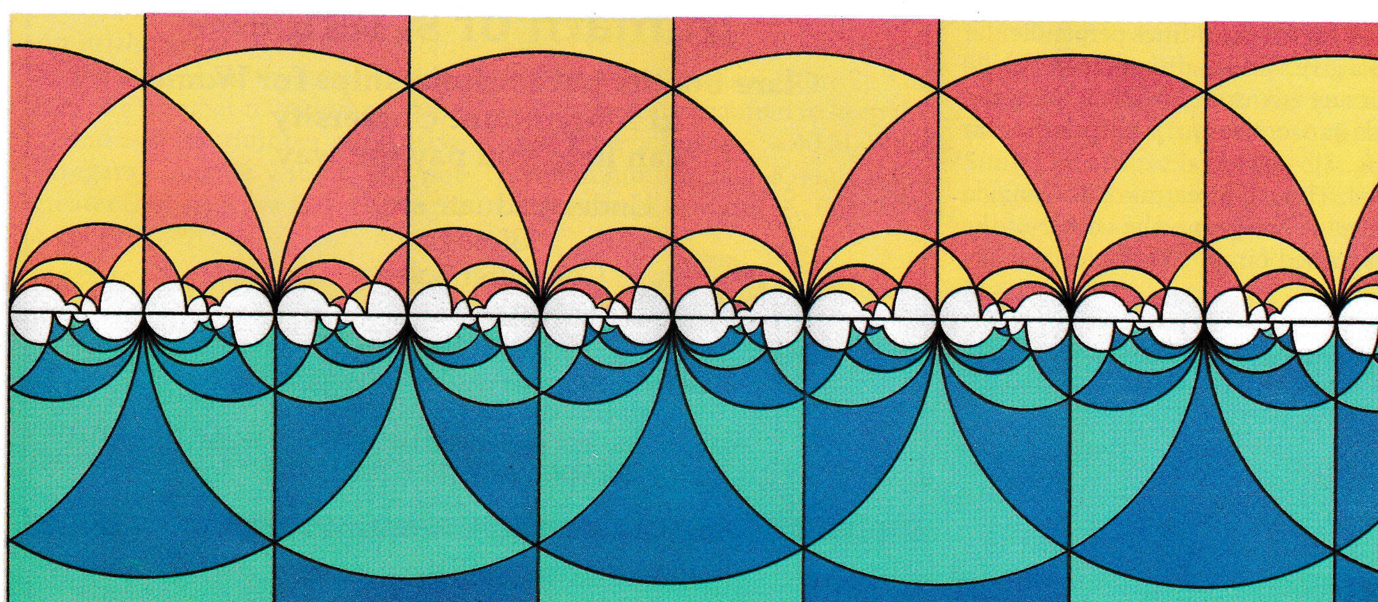


Although the most vivid representation of a straight line is undoubtedly a light beam, we nevertheless do not measure distances by the travel time of light—we use a ruler. It would probably be worthwhile for Poincarians to acquire a ruler, too. Of course, they will make their ruler "*p*-straight"; that is, it will have the shape of a segment of a *p*-line. But if Poincarian objects move in the same way ours do, then when such a ruler is moved to another place, it will no longer have the shape of nearby *p*-lines and so will not be "*p*-straight." To a Poincarian, a solid object will seem to change its shape when it moves. So how should a Poincarian react to this? The concept of a solid object has to be reconciled somehow with the geometry of Poincaria; otherwise, Poincarians will have to believe in supernatural powers. Henri Poincaré thought of an ingenious way out of this seemingly hopeless situation: he made use of *thermal expansion*.

Let all solid objects in Poincaria have the same coefficient of thermal expansion and zero thermal conductivity, and let their dimensions be proportional to their absolute temperature T. (Notice that under these circumstances Poincarians will be unable to measure temperature with an ordinary thermometer, because such a measurement presupposes comparing the expansion of objects with different coefficients of thermal expansion.) The distinguishing characteristic of a solid object is that when it moves in a medium with constant temperature, the (Euclidean) distance AB between any two of its points A and B remains the same. When an object moves from a region at temperature T_1 to a region at temperature T_2, the distance between its points changes by a factor of T_2/T_1; in other words, the ratio AB/T is preserved. But what happens when an object moves in a medium whose temperature varies from point to point? What value is preserved under these conditions? Suppose, for instance, that on one side of a certain line m the temperature of the medium is T_1, on the other side T_2. Let A be a point of a sufficiently large object located at temperature T_1 and B a point at temperature T_2. Consider a broken line ACB with node C on line m, and a value $r_1/T_1 + r_2/T_2$, where $r_1 = AC$, $r_2 = CB$ are the (Euclidean) lengths of the segments of our broken line. It turns out that when the object moves in such a medium, what is preserved is *the minimum value of* $r_1/T_1 + r_2/T_2$ *taken over all broken lines* ACB *with node* C *on line* m. (Try to figure out why!) This situation can be compared with the propagation of light in a two-layer medium: if the speed of light in one layer is numerically equal to T_1 and in the other to T_2, then $r_1/T_1 + r_2/T_2$ is just the time it takes the light to travel from A to C and then to B, and the minimum of this expression, by Fermat's principle, is the time it takes the light to travel from A "straight" to B—but, of course, refracting on line m according to Snell's law.

Extending this analogy to a medium with continuously varying temperature, assume that every point (x, y) in Poincaria is held at constant absolute temperature $T(x, y) = y$. Then, owing to these temperature conditions, the Poincarian travel time of light between any two fixed points A and B of a solid (in our usual sense) object will always be the same, no matter where and how this object is located, while the Euclidean distance AB will not. But this time is exactly the *p*-distance! So for Poincarians (who can't feel differences in temperature!) the dimensions of an object moving in such a medium are preserved, which means that it is "*p*-solid." What remains is to ensure that all objects have low heat capacities and move so slowly as to stay in thermal equilibrium, and that the variation in temperature is imperceptible for Poincarians. As a result, Poincarians will be unable not only to see the border of their world but to ever reach it: approaching the border, the temperature tends to absolute zero; therefore, the dimensions of objects will also approach zero, while their proportions will be retained.

Henri Poincaré tried to rule out every possibility for Poincarians to discover that their non-Euclidean world is nothing but a construction in our



Euclidean one. But did he make allowances for everything? If you discover any possibilities he failed to take into account, let us know.

Your turn

In solving the following problems you may find it helpful to draw the configuration in a position more convenient from the Euclidean point of view (although all positions are indistinguishable for Poincarians). For instance, we can always assume that one of the p-lines under consideration is a vertical ray, and so on. Also, sometimes it's better to use the "circular" version of the Poincaré model, which is discussed in the article "Inversion" in the September/October issue. The first problem explains how to pass from one version to the other.

1. Verify that inversion in a circle ω lying in the lower half-plane turns the Poincaré model in the upper half-plane, which is considered here, into the Poincaré model in some circle α (the model described in "Inversion")—that is, it maps the upper half-plane onto the interior of a circle α, p-lines into circular arcs orthogonal to α (or into diameters of α); and the p-distance $\rho(A, B)$ is equal to the p-distance $d(A', B')$ between the inverses A' and B' of points A and B, measured according to the formula given in "Inversion."

2. Using the model in a circle, prove that Poincarian circles are Euclidean circles lying in the upper half-plane.

Prove statements 3 through 10.

3. (a) All the p-lines perpendicular to a given p-line are superparallel; (b) for any two superparallels there is a unique common p-perpendicular (fig. 7).

4. Three p-bisectors of a p-triangle meet at one point, the center of the inscribed circle. What can you say about the circumcircle of a p-triangle? Does it always exist?

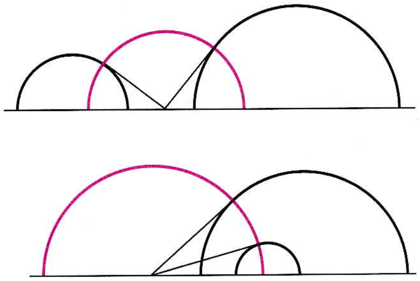


Figure 7

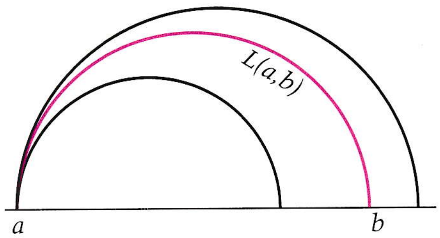


Figure 8

5. The base angles of an isosceles triangle are equal, and its bisector drawn from the vertex opposite to the base is also a median and an altitude.

6. The sum of the angles of a p-triangle is less than π.

7. If triangles ABC and AB_1C_1 have a common angle at vertex A (B_1 lying on p-ray AB, C_1 on AC) and equal angles at vertices B and B_1, C and C_1, respectively, then they coincide ($B = B_1$, $C = C_1$). Derive the AAA test for congruence of p-triangles. (Hint: if triangles ABC and AB_1C_1 do not coincide, consider quadrilateral BB_1C_1C and prove that statement 6 above is violated.)

8. Let $L(a, b)$, $L(a, b_1)$, $L(a, b_2)$ be three parallel p-lines (fig. 8). Then there exists a p-isometry taking $L(a, b)$ into itself, and $L(a, b_1)$ into $L(a, b_2)$. (This means that one cannot correctly define the "distance" between parallels in hyperbolic geometry.)

9. P-distance $\rho(A, B)$ is preserved under p-reflections and satisfies the equality $\rho(A, B) + \rho(B, C) = \rho(A, C)$ for any B on p-segment AC.

10. The perpendicular projection of a p-line L_1 onto a p-line L is a p-half-line if the lines are parallel, and a finite interval if they are not. ◘

ANSWERS, HINTS & SOLUTIONS ON PAGE 58

HOW DO YOU FIGURE?

Challenges in physics and math

Math

M66

Looking for order. In a one-round volleyball tournament each of the eight participating teams played each other. Prove that one can choose four teams A, B, C, D out of the eight such that A has beaten B, C, and D; B has beaten C and D; and C has beaten D.

M67

Intersecting parabolas. Two parabolas on the plane have perpendicular lines of symmetry and four common points. Prove that these four points lie on one circle. (L. Kuptsov)

M68

Suggestive coefficients. For any positive a and b prove the inequality $2\sqrt{a} + 3\sqrt[3]{b} \ge 5\sqrt[5]{ab}$. (N. Vasilyev)

M69

An arithmetic inequality. Let $\sigma(n)$ be the sum of all the positive divisors of a number n (including 1 and n) and $\phi(n)$ the number of positive integers coprime with n and not greater than n. Prove that $\sigma(n) + \phi(n) \ge 2n$. (V. Lev)

M70

Trapezoidal springboard. In trapezoid $ABCD$, diagonal AC is equal to leg BC and H is the midpoint of base AB. A variable line through H intersects line AD at P and line BD at Q. Show that angles ACP and QCB are either equal or supplementary. (I. Sharygin)

Physics

P66

Wedges and a washer. The inclined surfaces of two movable wedges of the same mass M meet the horizontal plane smoothly (fig. 1). A washer of mass m slides down the left-hand wedge from a height h. To what maximum height will the washer rise along the right-hand wedge? (Ignore the influence of friction.)

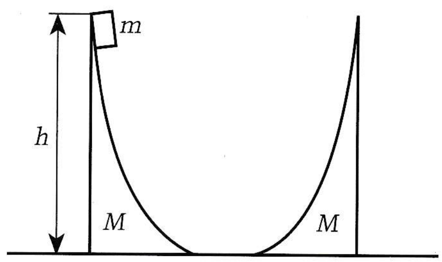


Figure 1

P67

To lift a load. A rope is thrown over a fixed, horizontal beam and fastened to a mass m = 6 kg (see figure 2). The minimum force required to keep the mass from falling is F_1 = 40 N. What is the minimum force F_2 required to pull the mass upward? (A. Buzdin)

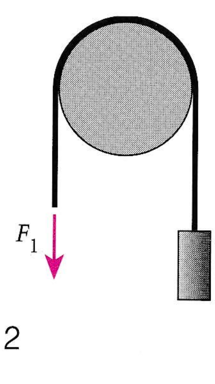


Figure 2

P68

Second skin. The surface of a space station is a blackened sphere in which a temperature T = 500 K is maintained due to the operation of equipment in the station. According to Stefan's law, the rate at which heat is given off from a unit surface area is proportional to the fourth power of the thermodynamic temperature. Determine the temperature T_x of the surface if the station is enveloped by a thin spherical black shell of nearly the same radius as the radius of the station. (A. Buzdin)

P69

Folded sheet. Imagine you take a very thin, square dielectric sheet that has a uniform charge on its surface and fold it twice so that you get a square whose sides are half as long as the original square. How much does its energy increase (or decrease)? (S. Krotov)

P70

Moon shadow. What is the velocity of the Moon's shadow on the Earth's surface during a total eclipse of the Sun? The eclipse is observed at the equator. For simplicity, the Earth's axis is taken to be perpendicular to the orbits of the Earth and the Moon.

ANSWERS, HINTS & SOLUTIONS ON PAGE 54

PHYSICS CONTEST

A topless roller coaster

"And in their motions harmony divine"
—*John Milton,* Paradise Lost

by Arthur Eisenkraft and Larry D. Kirkpatrick

WHY IS IT THAT AMUSEMENT park rides are such fun for some and a horror for others? Why will some of us rush to a park to try the newest, most frightening ride while others will only begrudgingly accept a ride on a ferris wheel? One of us (LDK) enjoys roller coasters so much that he spent 1.5 hours navigating on the New York subway system to take a 1.8-minute ride on the Cyclone at Coney Island. And this was an old-fashioned, wooden roller coaster!

The amusement park rides can take on an added dimension of interest and fun when you try to understand the physics inherent in the designs. Many amusement parks now have activities created by members of the American Association of Physics Teachers and the National Science Teachers Association that serve as guides for field trips for physics classes across the county. In these activities, people build accelerometers, predict and measure speeds, and record heart beat changes as you spin around, flip upside down, or watch the floor drop from under you. The rides provide many opportunities for testing the ideas of physics in "real-world" situations.

Next summer, the International Physics Olympiad will be held in Williamsburg, Virginia. During a respite from the grueling five-hour exams, the best high school students from approximately 40 countries will spend a day at Busch Gardens enjoying themselves as they ride and try to explain the physics in the park.

Designing an amusement park ride is quite a challenge. The ride must be entertaining *and* safe. Modern roller coasters have added a new dimension to roller coaster riding by adding such things as loops and corkscrews. These require readers of *Quantum*, as future ride designers, to apply knowledge of circular motion and centripetal acceleration in addition to the conservation of mechanical energy in analyzing the rides. Of course, in order to do this in a simple way we can make a number of assumptions such as (1) there are no frictional forces (including air resistance), (2) the kinetic energy of the wheels can be neglected, (3) the train of roller coaster cars stays on the track without the safety rail, and (4) the train is a point mass. The last assumption allows us to neglect such things as the rotational kinetic energy, the angular momentum, and the orientation of the train.

Up to now, all of the roller coasters of the world use a continuous track. But that does not restrict our imagination. Let's imagine that we remove the top portion of the track in a vertical loop, creating the so-called topless roller coaster. This allows us to combine the physics of circular motion in a gravitational field with that of projectile motion.

Assume that the vertical loop is a circle with a radius R and that the portion that is missing has an angle 2α centered about the top of the loop as shown in figure 1.

A. From what height H must a car be released so that it will leave the loop at one side of the gap and still arrive at the other side of the gap to continue the trip? Check your answer by setting $\alpha = 0$ to see if you get the expected result for a complete roller coaster.

B. It is useful to analyze the problem using the dimensionless ratio H/R because this sets the scale for

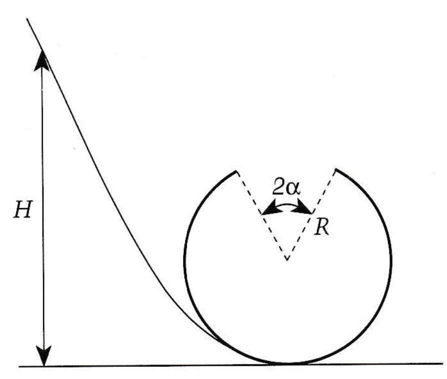


Figure 1

Art by Tomas Bunk


SOOORY, OUT OF ORDER!
Tom Bunk

the roller coaster. Draw a graph of H/R versus α.

C. For what range of angle α is this possible if the height is restricted to $H < 3R$?

D. Discuss the motion for the case $H = 5R/2$.

E. Discuss the motion for the case when H is a minimum.

Please send your solutions to *Quantum*, 3140 North Washington Boulevard, Arlington, VA 22201 within a month after receipt of this issue. The best solutions will be noted in this space and their authors will receive special certificates from *Quantum*.

Shake, rattle, and roll

In the May/June Contest Problem, we asked you to investigate earthquake behavior.

A. In part A, you were asked to calculate the time it would take for P or S waves emanating from the earthquake location E to reach an observation point X. From figure 2 we can see that

$$EX = 2R\sin\theta.$$

Therefore,

$$t = \frac{2R\sin\theta}{v},$$

where $v = v_P$ for P waves and $v = v_S$ for S waves. This is valid provided that X is at an angular separation less than or equal to X', defined by the tangential ray to the liquid core. From figure 2, X' has an angular separation given by

$$2\phi = 2\cos^{-1}\left(\frac{R_c}{R}\right).$$

B. Given the delay time between P and S waves, you were next asked to deduce the angular separation of E and X. Using the result from part A,

$$t = \frac{2R\sin\theta}{v},$$

we can express the time delay as

$$\Delta t = 2R\sin\theta\left(\frac{1}{v_S} - \frac{1}{v_P}\right). \qquad (1)$$

Substituting the data given, we get

$$131 = 2(6370)\left(\frac{1}{6.31} - \frac{1}{10.85}\right)\sin\theta.$$

Therefore, the angular separation of E and X is

$$2\theta = 17.84°.$$

This result is less than

$$2\cos^{-1}\left(\frac{R_c}{R}\right) = 2\cos^{-1}\left(\frac{3470}{6370}\right) = 114°,$$

and consequently the seismic wave is not refracted through the core.

C. If a second set of P and S waves had a longer delay, readers first had to hypothesize an explanation for the second set of delayed waves and then see if the result is consistent with the time delay given in part B.

The observations are most likely due to reflections from the mantle–core interface. Using the symbols in figure 3, we can express the time delay $\Delta t'$ as

$$\Delta t' = (ED + DX)\left(\frac{1}{v_S} - \frac{1}{v_P}\right)$$
$$= 2ED\left(\frac{1}{v_S} - \frac{1}{v_P}\right).$$

In triangle EYD,

$$(ED)^2 = (R\sin\theta)^2 + (R\cos\theta - R_c)^2$$
$$= R^2 + R_c^2 - 2RR_c\cos\theta,$$

since $\sin^2\theta + \cos^2\theta = 1$. Therefore,

$$\Delta t' = 2\sqrt{R^2 + R_c^2 - 2RR_c\cos\theta}\left(\frac{1}{v_S} - \frac{1}{v_P}\right).$$

Using equation (1), we get

$$\Delta t' = \frac{\Delta t\sqrt{R^2 + R_c^2 - 2RR_c\cos\theta}}{R\sin\theta}$$
$$= 396.7 \text{ s} = 6 \text{ min } 37 \text{ s}.$$

Thus, the subsequent interval produced by reflection of seismic waves at the mantle–core interface is consistent with an angular separation of 17.84°.

D. Since a P wave is able to travel through the core, you were asked to draw the path of the refracted P waves and derive the relation between the angle of incidence and the angular separation of E and X.

From figure 4, we get

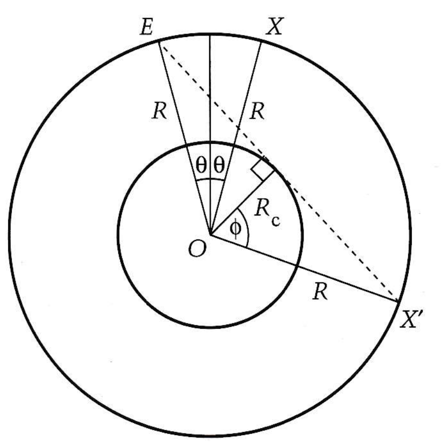


Figure 2

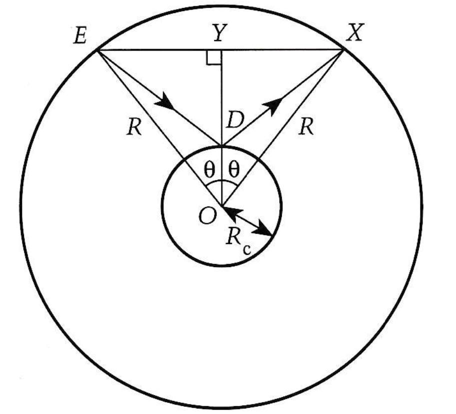


Figure 3

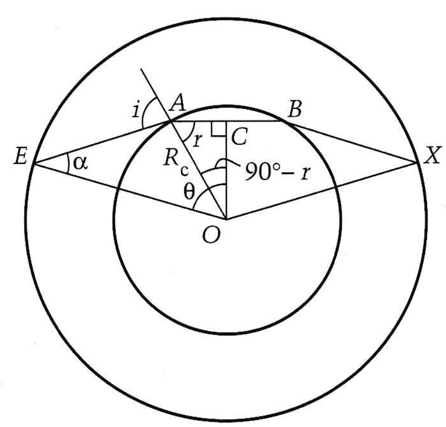


Figure 4

$$\begin{aligned}\theta &= \angle AOC + \angle EOA \\ &= (90 - r) + (i - \alpha).\end{aligned} \quad (2)$$

The law of refraction (Snell's law) gives

$$\frac{\sin i}{\sin r} = \frac{v_P}{v_{cP}}. \quad (3)$$

From triangle EAO and the law of sines, we get

$$\frac{R_c}{\sin\alpha} = \frac{R}{\sin i}. \quad (4)$$

Substituting equations (3) and (4) into (2) yields

$$\begin{aligned}\theta = 90 &- \sin^{-1}\left(\frac{v_{cP}}{v_P}\right)\sin i \\ &+ i - \sin^{-1}\left(\frac{R_c}{R}\right)\sin i.\end{aligned} \quad (5)$$

E. Our most talented readers were

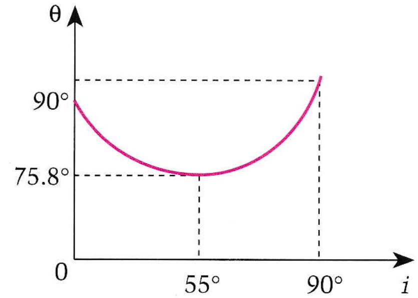


Figure 5

then asked to draw a graph of the relationship expressed in part D and to comment on the physical consequences.

Substituting $i = 0°$ into equation (5) gives $\theta = 90°$; $i = 90°$ gives $\theta = 90.8°$. Substituting numerical values for $i = 0°$ to $i = 90°$, one finds a minimum value at 55° and the corresponding minimum value of θ: $\theta_{min} = 75.8°$ (fig. 5). As θ has a minimum value of 75.8°, observers at positions for which $2\theta < 151.6°$ will not observe the earthquake as seismic waves. For $2\theta < 114°$, however, the

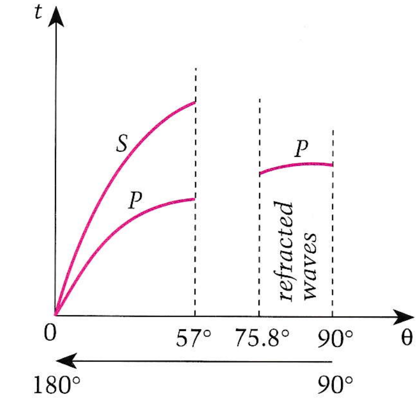


Figure 6

direct, nonrefracted waves will reach the observer.

F. Finally, readers were asked to sketch a comparison of the travel times for P and S waves for all angles. In this sketch (fig. 6) we can get a better sense of the "shadow" region where no earthquake waves will be observed.

KALEIDOSCOPE

Hit or miss

Take aim at these tricky problems from a master teaser

THIS MOTLEY COLLECTION of problems is borrowed from two books by the great Yakov Perelman, *For Young Mathematicians: The First Hundred Puzzles* (1924) and *Do You Know Physics?* (1934–35).[1]

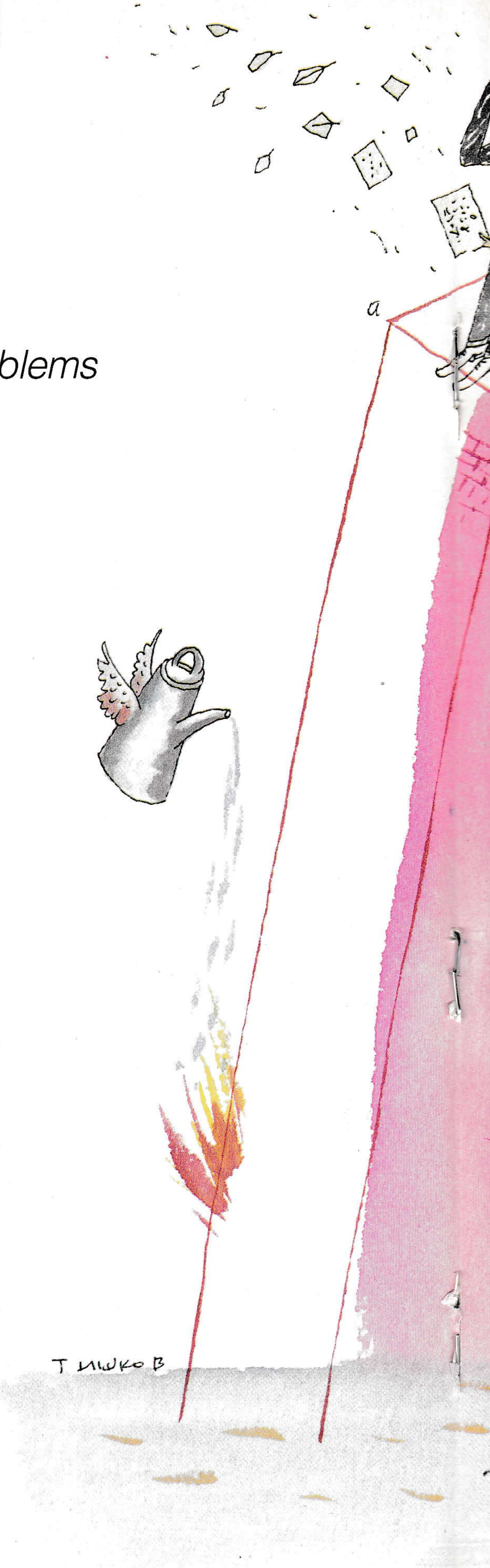
All of them seem pretty innocent, but they're not as easy as they look. When you come up with an answer, give it a second thought: did I hit the bull's eye, or did I miss the target entirely?

1. *Weight of a log.* A round log weighs 30 kilograms. How much would it weigh if it were twice as thick but half as long?

2. *Balance underwater.* One hundred kilograms of iron nails are weighed on a decimal balance (a balance with one arm ten times the length of the other) with iron weights in the other dish. The scale is immersed in water. Will the scale remain balanced?

3. *Hands together.* At twelve o'clock one hand covers the other one. But as you may have noticed, this isn't the only time the hands coincide: they run after one another all day long and line up several times a day. Can you name all the times when this occurs?

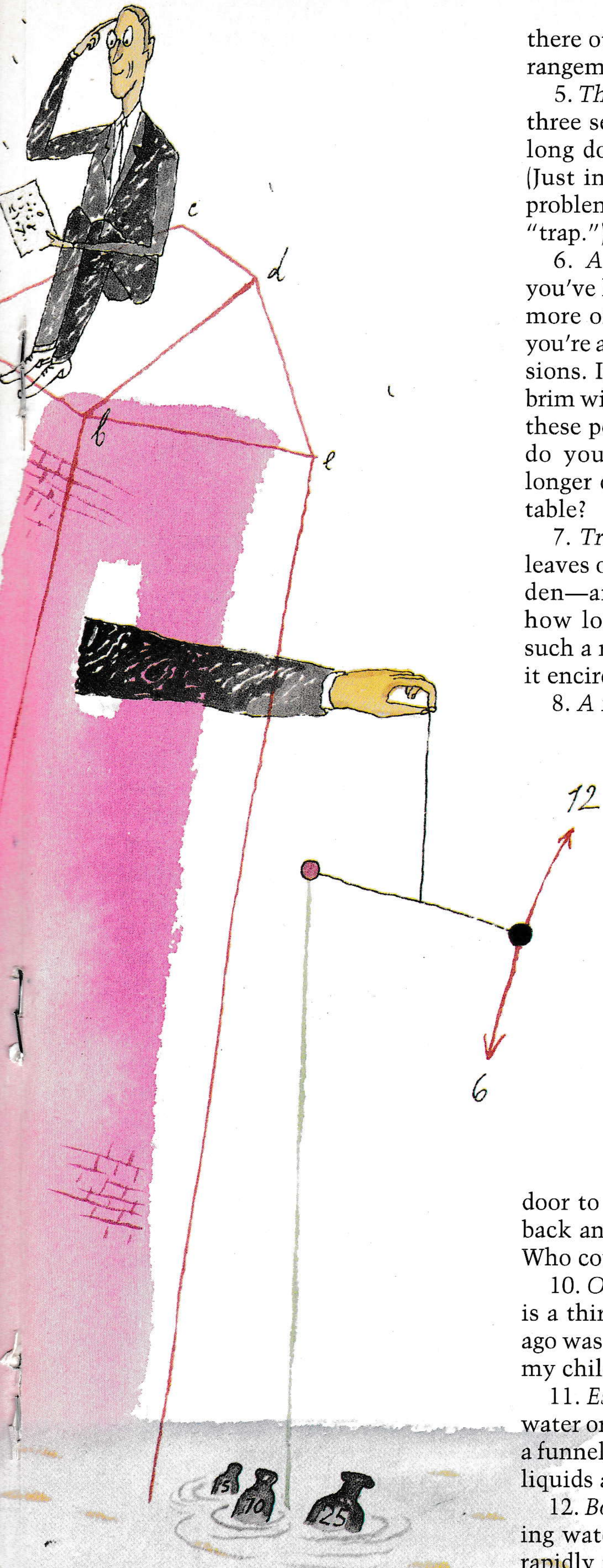
4. *Hands apart.* On the other hand, at six o'clock the hands point in opposite directions. Does this happen only at 6 o'clock, or are



---

[1]Books by this prolific popularizer of math and science are continually reprinted in English as well as Russian—for example, *Mathematics Can Be Fun* (Moscow: Mir Publishers, 1985) and its companion volume *Physics Can Be Fun*. The latter can be obtained from Victor Kamkin Bookstore, 4956 Boiling Brook Parkway, Rockville, MD 20852, phone 301 881-5973. (The former is currently out of stock.)—*Ed.*

Art by Leonid Tishkov



there other times when such an arrangement of the hands is possible?

5. *Three and seven.* A clock takes three seconds to strike three. How long does it take to strike seven? (Just in case, I warn you that this problem isn't a joke and there is no "trap.")

6. *A glass of peas.* No doubt you've held a pea in your hand, and more often you've held a glass, so you're acquainted with their dimensions. Imagine a glass filled to the brim with peas, and then imagine all these peas lined up in a row. What do you think—will this row be longer or shorter than your dinner table?

7. *Tree leaves.* If we tear all the leaves off some old tree—say, a linden—and line them up in a row, how long, approximately, would such a row be? For example, would it encircle a house?

8. *A million steps.* You know, of course, what a million is, and you know the length of your own stride. And because you know the one and the other, it won't be difficult for you to answer this question: how far will a million steps take you—more than 10 kilometers, or less?

9. *Who counts more?* In one hour, two men counted all the people who passed by. One of them stood near the door to his house, the other paced back and forth along the sidewalk. Who counted more passers-by?

10. *Old block and chip.* My child is a third as old as I, but five years ago was a fourth my age. How old is my child now?

11. *Escape speeds.* Which liquid, water or mercury, will empty out of a funnel sooner, if the heights of the liquids are the same?

12. *Boiling fire extinguisher.* Boiling water extinguishes a fire more rapidly than cold water, because it immediately removes the heat of evaporation from the flame and surrounds the fire with a layer of steam, which prevents the influx of fresh air. With this in mind, maybe the fire department should come with barrels of boiling water and pump that onto the burning buildings—what do you think?

13. *Bottom water.* When is the water at the bottom of a deep river warmer—in the summer or winter?

14. *Heating steel.* Why do steel structures collapse in a fire, even though steel doesn't burn or melt in the flames?

15. *Steamy color.* What color is steam?

16. *Match power.* What is the power of a burning matchstick?

17. *Salting water.* Can you dissolve more salt in water at 40°C or water at 70°C?

18. *Noisy shell.* Why do we hear the "sea" when we hold a cup or shell to our ears?

19. *Snow and black velvet.* What is lighter, black velvet on a sunny day or pure snow on a moonlit night?

20. *Color change.* When does gold have the color of silver?

ANSWERS, HINTS & SOLUTIONS IN THE NEXT ISSUE

Mathematics in living organisms

Does your kitty calculate vectors faster than you?

by M. Berkinblit and E. Glagoleva

THE NATURAL WORLD HAS contrived a great many "inventions" that people could understand and repeat only with a rather sophisticated level of development in science and technology. For instance, the echo location principle is used by dolphins and bats, but in technology it appeared only in the twentieth century. Many species of snakes catch their prey by infrared radiation, while devices for night vision are a recent development. And the list goes on. Not long ago the belief was prevalent that nature did not invent the wheel; that technology had been taking its own course. But then it was discovered that the flagella of bacteria rotate in special "bearings," which means that the wheel was also invented by nature at the earliest stages of evolution. There is a special branch of science called bionics that studies "nature's patents," so to speak. It turns out that some of these patents find practical application in "human" technology as well.

We can find phenomena in living organisms that would lead us to believe that nature made the first patent applications for electronic computers. These natural "devices" perform operations very much like the mathematical operations we tend to think of as achievements of human science. They in fact *are* electronic, since their operation is based on electrical phenomena in the organism. It looks like a repeat of the story of the invention of the wheel.

In this article we'll tell you about some operations of this kind: how nerve cells can count; how the eye takes logarithms and why; and how the brain of a cat and a monkey (and a human as well) makes use of vectors and trigonometric functions. Maybe some readers will conclude that it's not necessary to study these things, since they were given to living organisms by nature. But maybe others (and we hope there are more of these) would like to know about the mathematical and biological aspects of the matter.

How neurons count

Our first acquaintance with mathematics is counting:

One, two, buckle my shoe,
Three, four, shut the door . . .

And the simplest numbers seem to be the natural numbers. Even negative numbers made a slow entry into mathematics. They were known to the mathematicians of India in the early Middle Ages, but the negative numbers penetrated into European science only in the thirteenth and fourteenth centuries, and they weren't exactly warmly received when they got there. They were called "erroneous" or "absurd" numbers. But little by little the negative numbers affirmed their right to exist. They became commonplace, and not just for experts. The "cutting edge" of medieval science draws no blood in elementary schools today.

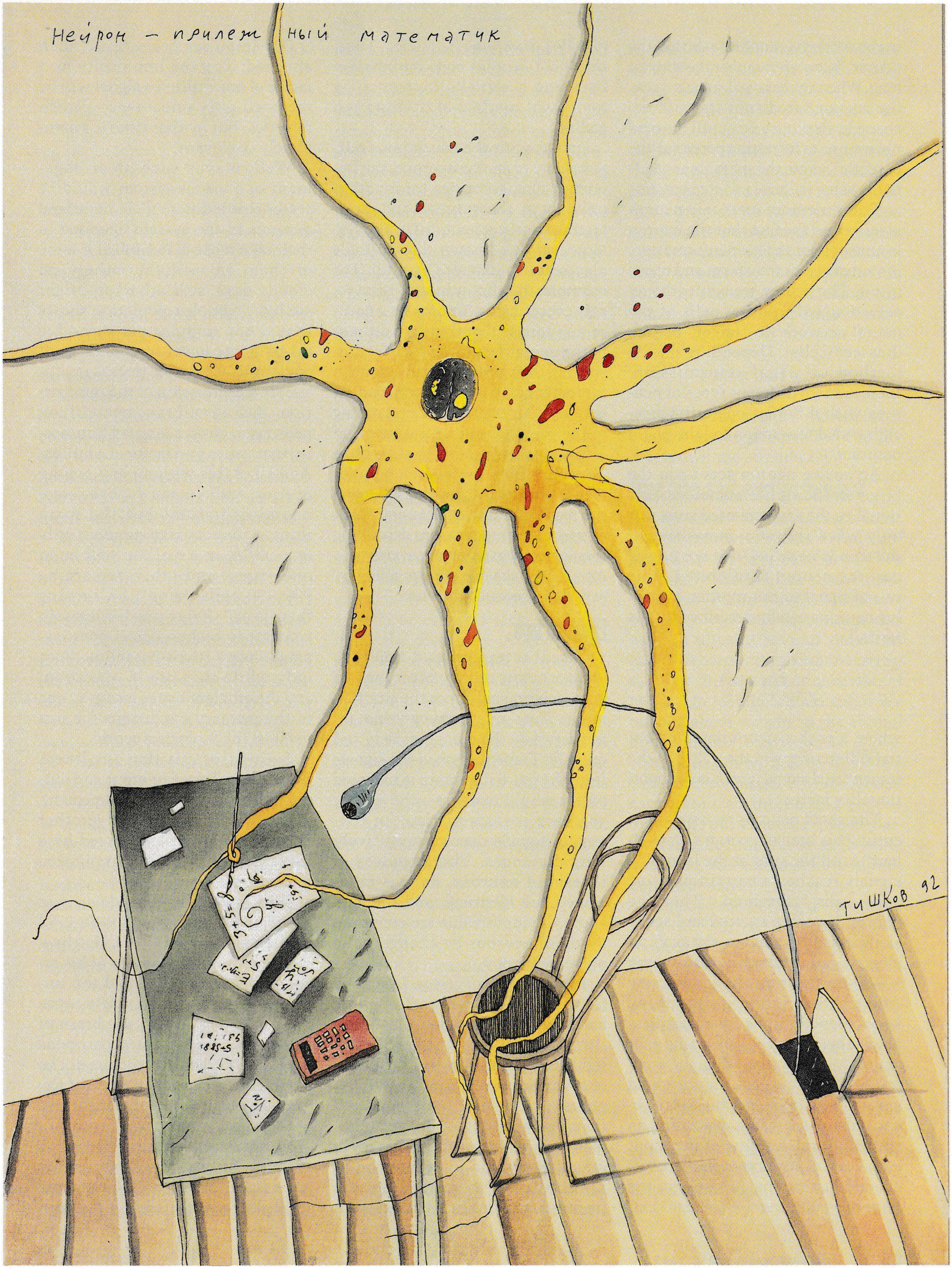
But in living organisms, everything is the other way around: it is natural and simple for the nerve cell (neuron) to carry out operations with positive and negative real "numbers." But in order to count even to two, a system consisting of several neurons—a primitive brain—is required.

How does a neuron operate? Like any cell, the neuron is separated from the external intercellular medium by a special envelope, or membrane. There is a difference in potential between the inner contents of the cell and the external medium. If the cell is not being activated, the difference in potential across its membrane does not change. This steady-state potential difference is called the "resting potential." It's natural to take this resting potential as zero (just as zero on the Celsius temperature scale was taken to be the point at which ice melts).

Other nerve cells—stimulating and inhibiting cells—can influence a neuron. The signals received from these cells cause changes in the difference in potential across the membrane in opposite directions. When different signals arrive at the neuron simultaneously, they are combined and the sign is naturally taken into account—that is, the neuron sums the positive and negative signals it receives. This sum can be either positive or negative.

An interesting feature of a neuron—as opposed to artificial summing devices from the ancient abacus

Art by Leonid Tishkov


Нейрон — прилежный математик
ТИШКОВ 92

to the modern computer—is that the neuron doesn't remember the sum for long. When the external signals cease, the accumulated sum starts to decrease in absolute value until the neuron returns to its resting potential (the potential across the membrane tends to the value we have taken to be zero).

This "memory loss" in a neuron is related to the fact that the neuron is not supposed to keep information, it's supposed to transform and transmit it. The neuron transmits the received signal to other cells of the neural network (to "target" or "addressee" cells). Depending on the method of signal transmission, there are two different types of neurons with different operating principles: analog neurons and threshold neurons.

An analog neuron acts upon the target cells with a signal proportional to the accumulated sum, but only when this sum is positive. If the sum is negative, it is not passed on—the neuron is inhibited. This rule of signal transformation by analog neurons can be described by the formula

$$y = \frac{k(x + |x|)}{2},$$

where x is the accumulated potential difference, y is the value of the transmitted signal, and k is a proportionality factor.

Threshold neurons operate differently. This kind of neuron "keeps silent" until the sum of the incoming signals reaches a certain positive value—the "threshold." Then the neuron becomes excited and sends along its exit appendage, the axon, an electric pulse (always of the same value) that acts on the target cells. After excitation, the neuron rests and keeps silent for some time regardless of whether other cells are acting on it or not. And after that, when the accumulated sum exceeds the threshold value, the neuron sends the next electric pulse. As a result, depending on the value and duration of the input signal, and depending on the neuron's characteristics, the output is in the form of a pulse train of constant amplitude but varying frequency. Thus, threshold neurons code information by using a signal's frequency—a noutrivial method of information coding.

Just as with the continuous output signal of analog neurons, the frequency change carries information only about the value of the input signal that varies continuously. However, it's known that animals also react to discrete stimuli. For example, they know how to react to, say, every third stimulus. It's natural to suppose that there are devices in the neural system that can count, so that they will react differently, for instance, to twofold and onefold actions. Our current understanding of the principles of neural function allows us to assert that such a "simple" (from the human point of view) operation as counting is beyond the reach of a single nerve cell. Lack of space prevents us from describing the device consisting of several neurons that is capable of responding to, say, every second stimulus.

Eyes and logs

Visual receptors, and others as well (acoustic, temperature, and so on) receive signals from the environment. They must convey visual information to the brain precisely and quickly. Transmission of the signals from the eye to the brain is initiated by threshold neurons—the analog approach is inappropriate for conveying signals over comparatively long distances. The impulses of threshold neurons, as was mentioned, are identical, and the data about the input value are conveyed by these neurons by changing the pulse frequency.

Here a problem arises. Luminosity at twilight, when things are barely visible, is about a billionth (10^{-9}) the luminosity in bright sunlight. But the maximum frequency of a neuron's operation is 1,000 impulses per second. It's easy to see that it's impossible to convey information by changing the frequency of neural operation in proportion to luminosity: if in bright light the impulse frequency is maximum (1,000 impulses per second), then after reducing the luminosity by a factor of one million a signal will be received only once every fifteen minutes. But by that time it's completely irrelevant!

Well, maybe such an arrangement of a visual system would be reasonable if different elements and neurons in the system operated in their own range of luminosity: some of them in twilight, others on cloudy days, still others in bright sunlight. Simple calculation shows that if one impulse per second is taken as the lower frequency limit of neural operation necessary for timely transmission of the information, then it will take one million neurons to cover a range of luminosity that varies by a factor of a billion. And all of this without any margin of error, without any duplication of neural operation! But the main thing is that at every moment only one cell from a million will be in operation and the remaining 999,999 cells "won't be earning their keep": living mechanisms, in contrast to technological ones, expend energy (their "gasoline") not only while they are in operation; and the efficient use of energy is one of the primary conditions for survival in the biological world.

So it turns out that the linear dependance between input and output in human vision isn't expedient. And, in fact, nature uses another relationship—one expressed by a function that is rather complicated by textbook standards.

This functional dependence was established experimentally in 1932 by the English scientist H. Hartline. Figure 1 shows the results of his research. Hartline recorded the impulses running along a single nerve fiber from the eye to the brain of horseshoe crabs (sea arthropods similar to the extinct trilobites). The graph shows the dependence of the impulse frequency on the brightness of the light.

"Excuse me!" you say. "The graph is a straight line, which means that it's a linear function." Not so fast! Look closely at the scale of the hori-

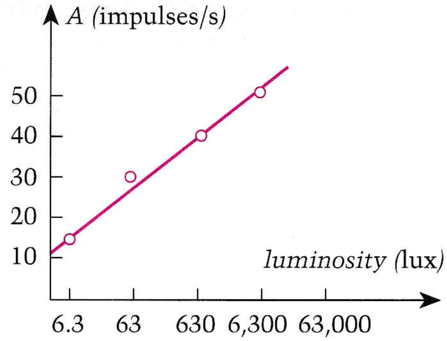


Figure 1

zontal axis—it is not linear. One increment corresponds to a change in the argument (brightness) not by the same value but by the same ratio. When there is a linear dependence, equal increments of the argument correspond to equal increments of the function—that is to say, a linear dependence converts an arithmetic sequence of arguments into an arithmetic sequence of function values.

When we deal with the exponential function $y = a^x$, a uniform *ratio of the values* of the function corresponds to equal increments of the argument. For instance, under constant habitat conditions and with unlimited resources, any population increases exponentially. The number of individuals increases yearly by some value, such as 10 percent—that is, the population at the end of the year is 1.1 times that at the beginning of the year. In other words, an exponential function "translates" an arithmetic sequence into a geometric one.

In our graph the situation is the reverse: the frequency of neural impulses changes by the same value when the stimulus changes by the same ratio. This means we're dealing with a function inverse to the exponential—that is, with a logarithmic function. In other words, the neurons of the horseshoe crab's eye convert a geometric progression of stimuli into an arithmetic sequence of signals.

This characteristic of the visual receptors, developed in the course of evolution, allows the eye to operate efficiently and economically, and it offers the possibility of correctly perceiving contrast. Suppose a light object and a dark one differ from each other by a factor of ten in their abilities to reflect light. Then a light object will reflect ten times more light than a dark object in both bright sun and twilight. Therefore, the comparative brightness of these objects does not change. The distance between corresponding points on the abscissa does not change either. And this means that the frequency difference of the operation of the receptors illuminated by these two objects will remain unchanged under different luminosities. So the ability to take the logarithm of a number allows the eye not only to operate over a wide range of luminosity but to distinguish, under conditions of low luminosity, objects whose absolute difference in luminosity is very small.

It's interesting that the dependence described above between the external signal (stimulus) and the signal perceived by the brain (sensation) was first found by psychologists. The eighteenth-century French scientist Pierre Bouguer made the discovery. At the beginning of the nineteenth century the German physiologist and psychologist Ernst Weber made a careful study of the relation between stimulus and sensation. He attempted to ascertain how some stimulus must change for a person to perceive this change. It turned out that the ratio of the change in the magnitude of the stimulus to its original magnitude is given by

$$\frac{\Delta I}{I} = k,$$

where I is a measure of the stimulus, ΔI is the incremental change in the stimulus, and k is a constant, called the *Weber constant.*

The Weber constant depends on which receptor is stimulated. For instance, in perceiving weight, $k = 1/30$. This means that when a person is holding a load of 100 g, a change is perceived when the weight is increased by 3.3 g; for a load of 200 g, an increase of 6.7 g is needed. For the pitch of sound, the Weber constant equals 0.003; for loudness, it equals 0.09. There are Weber constants for other sensations as well.

Working from Weber's results, another German physiologist and psychologist, Gustav Fechner, formulated the well-known Weber–Fechner law: *Sensation increases in an arithmetic progression when stimulation increases in a geometric progression.* This law was published in Fechner's book *Elements of Psychophysics* in 1859. In the same book the mathematical expression of the law is also presented:

$$E = a \log I + b,$$

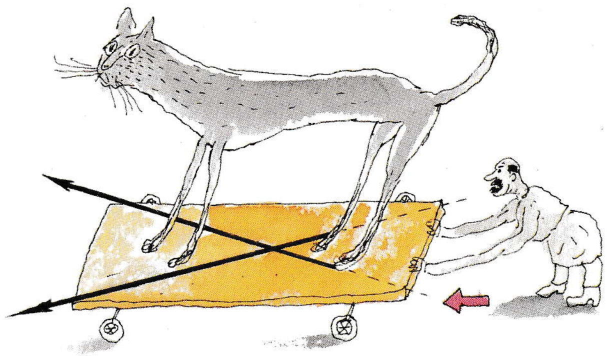
where E is the measure of a sensation, a and b are constants, and I is the measure of the stimulation.

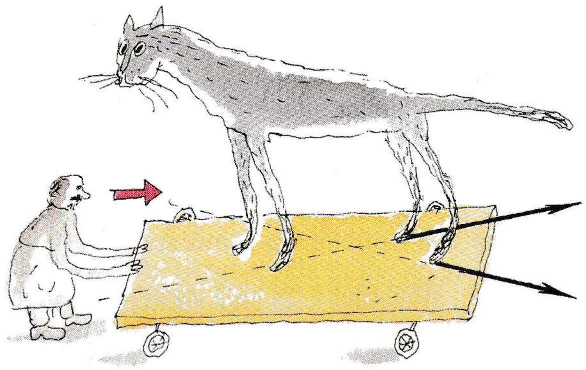
What does a cat need vectors for?

The word "vector" is just a baby, one might say. It seems to have appeared for the first time in 1845 in a work by the English mathematician Sir William Rowan Hamilton. But the corresponding concept was used in physics several centuries earlier in connection with investigations of the law of composition of forces (the parallelogram rule). We learned about vectors in the animal kingdom only very recently.

It all began with cats. In 1988 a Canadian scientist, J. Macpherson, carried out an interesting experiment. She put a cat on a special platform, pushed the platform in some direction, and watched how the cat kept its balance.

Suppose she pushes the platform forward. The cat's feet, together with the platform, begin to move forward, but its body remains in place. Then the cat, in order to restore the center of gravity to its initial correct position above the support points, activates the muscles in its legs and moves its body forward by pushing on the platform. If the platform is pushed to the right, the center of gravity will deflect to the left relative to the support and the cat's legs should produce a force directed to the right.





How do a cat's legs actually work as it tries to keep its balance?

The most natural thing is to suppose that each of its hind legs,[1] when pushing forward, produces a force directed forward; the sum of the two forces returns the body to the correct position (see figure 2a). If the platform is pushed to the right, each hind leg produces a force directed to the right, and so on. This hypothesis jibes with the fact that a cat has powerful muscles that move its hind legs forward or backward (used for walking and jumping) as well as muscles that turn the leg outward or toward the body's axis.

Macpherson elucidated what in fact occurs. It turned out that the actual scenario is quite different: when the platform is pushed, regardless of the direction, the cat's hind legs produce forces directed along two lines (one line for each leg) at angles of 45° to the body's axis. Even in the simplest case, when the platform is pushed straight forward, the forces produced by the legs are directed not forward but again at an angle of 45° to the body's axis (see figure 2b). Only their sum has the necessary direction and magnitude. Figure 2c shows how a force perpendicular to the body is produced. A force directed at an angle of 30° to the body's axis is shown in figure 2d.

So the cat's nervous system solves the following problem. When the platform is pushed, the information received from various receptors makes it possible to determine the needed vector (force). Then this vector is laid along the fixed coordinate axis. When this method is used, only one number needs to be transmitted to each of the cat's hind legs. This number is the magnitude of the force (positive or negative) that the leg must produce along its fixed axis.

This is a relatively simple picture. But life is full of unexpected things! Having figured out which muscles produce this fixed direction, one would think that it would always occur in the simplest way. The muscles that move the leg forward and in could produce forces in one direction, and the muscles that move the leg back and out could produce forces in the opposite direction. The different magnitudes would be produced by proportionally changing the forces developed by these muscles—"multiplying by a number." Macpherson got another unexpected result. She found that different muscles can take part in producing a given magnitude of force for a single leg. The combination changes, depending on the direction of the push.

What is the point of such a complicated (from our point of view) solution? That is a question that remains to be answered. But here we've seen an instance of a general biological principle: avoid rigid schemes, and always keep an overstock of degrees of freedom.

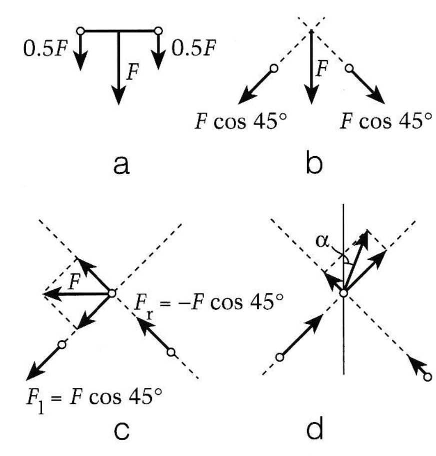


Figure 2

Vectors in the primate brain

The difficulties one encounters in explaining how this or that problem is actually solved have to do with the fact that it's very difficult to look inside the "control center"—that is, the brain itself. In this sense the brain is still a "black box." We can see the problem presented to the brain, and we can see what result it produces, but we have almost no idea what is happening inside the brain.

Researchers wanted to see more directly how neurons function in solving certain problems. The American scientist A. Georgopoulos recently advanced our understanding in this area with a series of clever experiments involving a trained monkey. The monkey's leg was placed at a certain point on a table, and electric lights were placed at various points around the table. The monkey was trained to move its leg toward any bulb that flashed. At the same time the researcher recorded through implanted electrodes the activity (impulse frequency) of the nerve cells of the cerebral cortex in the area that controls leg movement.

Georgopoulos found that the activity of most of the cells in this part of the brain depends on the direction of leg movement. This dependence

[1]It was found that when it restores its center of gravity, a cat uses its forefeet as passive supports. Only the hind legs work actively.

is clear-cut: for every cell there is a direction of movement such that the activity is maximal; with other directions the activity decreases approximately as the cosine of the angle between this direction and the direction of the maximum activity.[2] For the directions where the cosine is negative, the cell generally stops sending impulses.

It turns out that a certain vector of maximum activity $\mathbf{A}_{max}$ is linked with every cell in the cerebral cortex (fig. 3). When it's necessary to move the leg in another direction—that is, when some unit vector of the direction e is given, the cell finds the projection $\mathbf{A}_{max}$ in this direction—that is, it "computes" the scalar product $\mathbf{A}_{max} \cdot \mathbf{e}$. Having clarified this point, Georgopoulos posed the inverse problem: is it possible to determine the direction of leg movement by monitoring the activity of the nerve cells? Mathematically this problem could be formulated as a question about the existence of the function inverse to the given function. It's clear that it's impossible to determine the direction of movement by looking at the activity of a single cell: first, the cosine is an even function and has no inverse function. Indeed, if the direction of the maximum activity is straight ahead, for instance, and the neural activity is one half of the maximum, then it's known that the leg moves at an angle of 60° to the primary direction, but whether to the right or left it's impossible to determine. Second, one cell has too large a "dead zone"—the zone where the cell keeps silent. But if several cells are monitored, it's possible to determine the direction in which the leg is moving (and even to predict this direction, since the cells begin to work a tenth of a second before the leg starts to move). We leave it to our readers to solve the following problem: what is the minimum number of cells needed to determine the direction of movement for any case? (Of course, we posed the problem in mathematical garb, so to speak, which as always simplifies the situation—just as we have simplified throughout this article.)

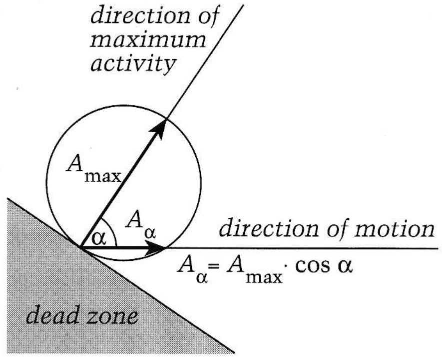


Figure 3

[2]The proportionality of the frequency of the activity of the nerve cells to the cosine of either angle was known before Georgopoulos's work. For instance, as early as 1981 neurons were discovered in the brain stem that were linked with twitching in the eye: the change in their activity depends on the direction of the twitch by the cosine law.

The fact that one can use neural activity not only to determine where the leg is moving but to predict where the monkey is about to move it—to read its thoughts about moving it—allowed Georgopoulos to conduct another experiment, and it was an elegant one.

As early as 1971 the American psychologists R. Sheppard and J. Metzler had discovered a phenomenon that they termed "mental rotation." Subjects in an experiment were shown two figures and asked: are they different figures or are they the same figure but rotated by some angle? The response time turned out to be a linear function of the magnitude of the angle of rotation of one figure relative to the other.

In another variant of the experiment, the letter R or its mirror reflection Я were shown. The subjects had to determine quickly what letter it was. In addition, this letter was shown in various orientations. And in this case the response time was proportional to the angle of rotation relative to the "normal" position.

The researchers supposed that the subject in such experiments mentally rotates an image of the figure being perceived (and, according to a series of psychological experiments, probably not the figure itself but a model of it stored in the memory) with a constant angular velocity. They even determined this velocity: 450°/s. It's impossible to prove the hypothesis of "mental rotation" by such experiments, however, since it's still not clear what's going on in the subjects' heads.

Now that he had a way of "spying" on the operation of neurons in a monkey's brain, Georgopoulos collected data in 1989 that support the hypothesis of mental rotation.

The monkey was subsequently trained to stretch its leg not toward a lit bulb but toward the one that is located at an angle of 90° to it. The researchers were able to discover what occurs in the monkey's brain from the moment the bulb lights up until the leg starts to move. After the flash, the vector is directed straight toward the bulb, then it begins to rotate, and when it turns 90° the leg begins to move. The velocity of the vector rotation was about 730°/s—that is, it was of the same order as in the psychological experiments with humans.

Thus, as these experiments have shown, the brain can also perform geometric transformations (in fact, not merely rotations but evidently many others—for example, the transformation of similitude).

We'll drop another hint about the mathematical abilities of the brain. Parallel processing in computers is undergoing furious development right now. But when a person picks up an object, the brain simultaneously controls the shoulder, elbow, and fingers—now that's parallel processing!

Conclusion

As we've seen, the processes of converting, transmitting, and using information for control purposes occur inside living organisms. Evolution has gradually selected successful forms of data processing, and these forms are rather like mathematical operations. And it is these clever products of evolution that we have called "mathematics in living organisms."

POSTER

Space physics: a voyage of adventure

A pictorial overview to close ISY 1992 and open new perspectives

by M. Frank Watt Ireton, Sue Cox Kauffman, Ron Morse, and Mark Pesses

THE AGE OF SPACE EXPLORATION HAS SEEN A dramatic increase in our knowledge of space physics—the scientific study of magnetic and electric phenomena that occur in outer space, in the upper atmosphere of planets, and on the Sun. Space physicists use ground-based observatories, balloons, rockets, satellites, and deep-space probes to study many phenomena both remotely and directly where they occur.

The origins and development of space physics are closely tied to the birth and growth of NASA itself. Research in space physics began at NASA in 1958, using data transmitted by the NASA satellites Explorer I and Explorer III. These data revealed belts of radiation trapped inside the Earth's magnetic field. These belts were named after their discoverer, Dr. James Van Allen of the University of Iowa. At about the same time, Dr. Eugene Parker of the University of Chicago demonstrated that the Sun's corona theoretically had to be continuously expanding at supersonic speeds. His conclusions were subsequently confirmed by NASA satellite data, and "solar wind"—the name for this extended solar atmosphere—became a commonplace term.

What does the poster show?

The poster includes a tour of the solar system and current research in various areas of solar physics. Beginning in the lower-left corner of the poster is the Sun, represented by the yellow surface we call the photosphere and red structures called prominences. The filamentary structures flowing outward from the Sun are coronal streamers; the longer structures are solar wind streams.

All nine planets are shown, from Mercury (closest to the Sun) to Pluto (farthest away). The wispy envelope flowing around each planet is its magnetosphere. Three comets are shown; Halley's comet is the one closest to Uranus. The solar wind causes the tails of both comets and magnetospheres to point away from the Sun.

Space is often referred to as a great "void," yet the poster shows a full, active scene. Is space truly empty anywhere? Why?

The Ulysses spacecraft is depicted flying by Jupiter in February 1992 on its way to the Sun (it will fly over its poles in 1994). Voyager 2 is shown as it flew by Neptune in August 1989 on its journey out of the solar system.

The arc near the top of the poster is the heliopause—the plasma boundary of our solar system. Beyond the heliopause, in interstellar space, are the Crab Nebula and a pulsar. The yellow and red "waves" emanating from these objects depict cosmic rays that are believed to be accelerated by the objects. The paths are "kinky" because cosmic rays are charged particles that scatter in interplanetary and interstellar space. Also visible in the poster are hundreds of stars.

Photos and figures related to areas of current research are shown on the right side of the poster.

Space shuttle aurora observations

This dramatic photo, taken from the space shuttle (mission STS-39), shows the northern auroral display. Most shuttle missions are flown at lower latitudes, and so the northern latitudes of the Earth are not visible. Mission STS-39 was flown at a higher latitude, making it possible to photograph the aurora through the cargo bay window. The photo plainly shows the white crescent

The poster inserted in this issue of *Quantum* rounds out our celebration of International Space Year. It is one of three backdrop panels created by Dale Glasgow for the NASA exhibit booth at NSTA's 1991 national convention in Boston. The poster was originally developed from the artwork and distributed at the Boston meeting. The National Science Foundation generously provided funds to reprint the poster for the benefit of *Quantum* readers. Our thanks to the American Geophysical Union for initiating this poster project and providing the accompanying text.

shape of an aurora generated by energetic electrons that have been accelerated along the Earth's magnetic field. (Notice also the glow visible on the shuttle engine pods and tail surface, produced by chemical interactions rather than the motion of charged particles.)

Ground-based solar studies

Solar research, such as that done at the pictured McMath solar telescope at Kitt Peak National Observatory, Arizona, is necessary for understanding the Sun's atmosphere and corona. Filtered photographs in the red light of H-alpha (6563 Å) show the intricate filamentary structure and dynamics at the Sun's surface and chromosphere, which form the underpinning of the coronal structures that we observe with ultraviolet and X-ray instruments from space.

Ground-based ionospheric studies

The Sondrestrom Radar Facility is located above the Arctic Circle in southwestern Greenland. The facility is a major hub for upper-atmospheric and solar-terrestrial research. The principal instrument is the 33-meter fully steerable parabolic antenna featured in the photo. The antenna is complemented by a wide range of instrumentation, including spectrometers, imagers, interferometers, magnetometers, riometers, and a lidar. The Sondrestrom facility is operated for the National Science Foundation by SRI International and is the largest NSF-supported facility north of the Arctic Circle. This radar is capable of measuring electron density, electron and ion temperature, Doppler velocities, and other parameters of the upper atmosphere and ionosphere.

Antarctic balloon flights to the edge of the atmosphere

Scientific experiments dealing with space can often be performed inexpensively through the use of balloons. The focus of the Antarctic balloon program has been on solar physics and cosmic rays. The largest balloons used in these studies reach an average altitude of 39 km and carry payloads weighing more than 2,250 kg.

The High-Resolution Gamma-Ray and Hard X-Ray Spectrometer (HIREGS) is an international collaborative investigation designed to make detailed measurements of the X rays and gamma rays produced when particles accelerated by solar flares collide with the solar atmosphere. The goal of the lengthy HIREGS balloon flights is to provide a better understanding of the conditions and mechanisms responsible for producing solar flares and accelerating particles to high energies.

In the first flight, on January 10, 1992, a 1,700-kilogram payload was launched on an 830,000-cubic-meter, helium-filled balloon and circumnavigated the Antarctic continent in 13 days, 17 hours. A second flight is planned for December 1992.

Why do you think such circumnavigation is possible with only a balloon?

Another investigation performed by balloon flights over Antarctica was designed to study the isotopic composition of the cosmic rays. A balloon-borne payload, Magnetic Passive Isotope Experiment (MAGPIE), was flown over Antarctica in December 1991. Measurements to determine the origin of cosmic rays and how they might have been accelerated to their remarkably high energies were conducted during this flight. Analysis of the data from the MAGPIE flight is currently underway.

Sounding rocket flights into near space

The multistage launch featured in the time-lapse photo took place in 1984 at NASA's Poker Flats launch facility in Alaska. A camera focused on the rocket, its shutter stopped open for over 20 minutes, records the dramatic liftoff, staging, and release of the experiments, along with the faint traces of stars and an auroral display. The first stage was a US Navy surplus Terrier rocket; the second stage was a Canadian-built Black Brant; and the third stage was propelled by a Canadian Nihka motor. The sounding rocket reached an altitude of approximately 1,000 km in 10 minutes. The upper stage, which contained the payload, was assembled at NASA's Wallops Island, Virginia, facility. The 135-kilogram payload contained instrumentation to track the rocket and to launch the two experimental devices released into space. The first release occurred 340 seconds into the flight; the second release occurred 685 seconds after launch. In the releases a high-temperature chemical reaction rapidly ejected barium, which is then ionized by solar ultraviolet radiation. Resonance fluorescence of these ions produces a faint glow of light, and the ions are aligned by the Earth's magnetic field. The space physicists participating in this experiment were studying the alignment and convection of ions in the Earth's magnetic field by tracking the glowing clouds. As you study the photograph, observe the arcs described by the first, second, and third stages and the location of the experiment releases. The rocket is actually gaining altitude during the flight, reaching an apogee of about 1,000 km. Why do you think it appears that the experiments are lower in altitude than the stages, even though you know they were conducted at a higher altitude? (Try to figure it out yourself, then compare your findings with ours on page 61.)

Data interpretation: Earth's Van Allen radiation belts

This figure is based on data provided by Dr. James Van Allen and shows the counting rate of charged particle detectors flown in Explorers I, III, and IV and Pioneer 3. In this figure, the blue sphere represents the Earth, the red zone the inner belt, and the yellow zone the outer belt of the Van Allen radiation belts.

The horizontal white lines represent the trajectory of Pioneer 3 (launched toward the Moon in 1959)—the lower line showing the outbound path and the upper line the return path. The numbers running horizontally are the distances from the center of the Earth in units of Earth radii.

What do you think a plot of Pioneer 3 data would look like?

Theoretical models: particle motion in the radiation belts

The formula shown on the poster is for the conservation of the first adiabatic invariant of a charged particle moving in a magnetic field. Any charged particle moving through any magnetic field will experience a force perpendicular to its direction of motion (direction of velocity vector). The angle between the velocity vector and the magnetic field vector is given the symbol α and is called the particle pitch angle. This perpendicular force causes the particle's motion to continually deflect to the right or left, depending on the sign of the electric charge and direction of the magnetic field. The resulting motion is spiral. The intensity of the force F depends on the amount of electrical charge q, the speed V, the strength B of the magnetic field, and the pitch angle α:

$$F = BVq \sin a.$$

These quantities are related by a variety of natural "laws" that can be expressed mathematically. An important mathematical relationship used by space physicists studying such phenomena is called "the conservation of the first adiabatic invariant." This relationship combines the laws of conservation of energy and conservation of momentum. It states that

$$\frac{\frac{1}{2} m V^2 \sin^2 \alpha}{|B|} = \text{constant},$$

where m is the mass of the particle; V is the speed, which is the magnitude of the velocity vector $\mathbf{v}$; α is the pitch angle between $\mathbf{v}$ and $\mathbf{B}$; and $|B|$ is the magnitude (strength) of the magnetic field vector $\mathbf{B}$. The formula is valid as long as changes in the magnetic field are slow on the time scale of one 360° spiral motion of the particle. As a charged particle moves in the Earth's magnetic field, it spirals toward one of the Earth's magnetic poles. The spiral path gets tighter as the magnetic field intensifies closer to Earth. (Why is this?) Eventually, but still far above the Earth's surface, the magnetic field gets strong enough that the particle "bounces" back, spiraling in a great curved path toward the other magnetic pole of the Earth. The bounce point occurs when $\mathbf{B}$ in the mathematical relationship above gets large enough so that $\sin \alpha$ equals 1. (At what angle would this be?) The process is repeated at the other pole, and the particles are trapped in the Earth's magnetic field at different locations, depending on their mass and charge. Areas where the particles are concentrated are called radiation belts or Van Allen belts.

Careers in space physics

To prepare for the graduate degree necessary for a career in space physics, you should take as many mathematics and basic science courses as possible in high school. Courses in algebra, geometry, trigonometry, calculus, computer science, chemistry, Earth science, and physics are very strongly recommended. In addition, it's important to develop the strong communication skills that will be essential in describing new discoveries for other scientists and writing proposals to obtain funding for research.

After receiving an undergraduate degree, a prospective space physicist should pursue a doctor of philosophy degree (Ph.D.) in physics with a specialization in space physics. Most doctoral programs take three to five years to complete, and graduate students very often earn their tuition, room, and board by working as either a research assistant or teaching assistant. After earning a Ph.D. the next step is usually a two- to four-year postdoctoral position, roughly equivalent to a medical doctor's residency.

M. Frank Watt Ireton, *manager of precollege education at the American Geophysical Union, coordinated this* Quantum *poster project.* **Sue Cox Kauffman** *is an education specialist and* **Mark Pesses** *is a senior staff scientist at Space Applications International Corporation in Washington, D.C.* **Ron Morse** *is an adjunct professor in the science teaching department at Syracuse University.*

The "assayer" weighs the facts

"The less people know and understand about such matters, the more positively they attempt to reason about them . . ."

by Yuli Danilov

AN OUTSTANDING MATHEMATICIAN OF OUR time, Hermann Weyl, once complained that the style of modern mathematical works reminds him of prison cells filled with dead electric light that renders all details equal because they don't cast any shadows. Weyl himself preferred "a soft landscape under the open sky" and wrote in perfect refined German reminiscent of Hermann Hesse, and while in exile in America he lamented "the fetters of a foreign language that he did not hear from the cradle."

The excerpt below is taken from a small but important work by one of founders of a new physics based not on a reference to someone's authority but on observation and mathematical calculation: Galileo Galilei (1564–1642). Galileo wrote the treatise in 1623 and gave it a title that might seem strange at first sight: "The Assayer, in which with a delicate and precise scale will be weighed the things contained in The Astronomical and Philosophical Balance of Lothario Sarsi of Siquenza." Reading it, we get an inkling of why Galileo's countrymen hold this relatively neglected work in high esteem, and why Italians consider Galileo one of the creators of the Italian literary language.

The excerpt is a parable presented as an argument in a scientific dispute on the nature of comets. To set the stage: at the end of 1618 three comets appeared one after another in the sky over Italy. The third comet was observed from November 1618 to January 1619 and was especially bright. The universal interest in these "tailed stars" stimulated the renewal of an old debate about the nature of comets. According to Aristotle, everything that is perishable and transient—capable of being born and passing away—belongs to the sublunar world. The celestial, translunar world contains all that is eternal, permanent, and perfect. If a comet is a real body, then, because it appears and disappears, in the framework of the Aristotelian conception it can belong only to the sublunar world. It's quite a different matter if a comet, as Galileo supposed, is an optical illusion, a play of light in evaporations rising to the upper layers of the atmosphere: an incorporeal vision could, without any contradiction, appear in the translunar world as well. A secondary question, it would seem, acquired primary importance, which explains why the ensuing discussion was so keen. The moral of Galileo's parable is clear: if the main character cannot understand how a cicada in his hands produces its song, how can one speak with certainty about the nature of comets that are out of reach, high above the Earth? "[T]he number of things known and understood," wrote Galileo, finishing the thought quoted above, "renders [people] more cautious in passing judgment about anything new."

What makes "The Assayer" so interesting is its passionate and at the same time inconspicuous profession of a new scientific method.[1] At one point Galileo lets drop a remark in the spirit of Democritus's atomistic theory, according to which all that exists is the result of a mixing of elements devoid of any qualitative differences. At another, he caustically mocks how his opponent adheres to proofs based on the opinions of others, even if these others happen to be the greatest Roman poets—Virgil, Ovid, Seneca, Horace. Elsewhere Galileo pens his famous phrase about the open book of the universe, accessible only to those who know the language of mathematics:

It seems to me that I discern in Sarsi a firm belief that in philosophizing it is essential to support oneself upon the opinion of some celebrated author, as if when our minds are not wedded to the reasoning of some other person they ought to remain completely barren and sterile. Possibly he thinks that philosophy is a book of fiction created by some man, like the *Iliad* or *Orlando Furioso*[2]—books in which the least important thing is whether what is written in them is true. Well, Sig. Sarsi, that is not the way matters stand. Philosophy is written in this grand book—I mean the universe—which stands continually open to our gaze, but it cannot be understood unless one first learns to comprehend the language and interpret the characters in which it is written. It is written in the language of mathematics, and its characters are triangles,

[1]At least it escaped Pope Urban VIII's (perhaps drowsy) attention—he would ask to have "The Assayer" read to him at mealtimes. And it even slipped past the unblinking eye of the Inquisition.

[2]This epic poem by Ludovico Ariosto (1474–1533) was a favorite of Galileo's and reflects most closely his own aesthetic views.

circles, and other geometrical figures, without which it is humanly impossible to understand a single word of it; without these, one is wandering about in a dark labyrinth.[3]

In the polemics that followed, the new scientific method was defended by Galileo and his pupil Mario Guiducci. The defender of the scholastic tradition was an influential professor of mathematics at the Collegio Romano, Horatio (Orazio) Grassi, who published a treatise entitled "The Astronomical and Philosophical Balance" under the pseudonym Lotario Sarsi. In criticizing certain of Guiducci's assertions, whose book resonated clearly with Galilean motifs (*ex ungue leonem*—"you can tell a lion by its claws"), Sarsi rhetorically "weighed" them. In response, the infuriated Galileo sliced his opponent's treatise into 63 fragments and figuratively weighed each of them with the especially sensitive balance used by assayers (persons who analyze substances for one or more specific components)—thus the title.

The assayer

(Excerpt)

There once lived, in a very solitary place, a man endowed by nature with extraordinary curiosity and a very penetrating mind. He raised many birds as a hobby, much enjoying their songs, and he used to observe with great admiration the happy contrivance by which they would transform at will the very air they breathed into a variety of sweet songs. Close to his house one evening, he chanced to hear a delicate sound, and, being unable to imagine what it could be except some small bird, he set out to capture it. Arriving at the road, he found a shepherd boy who was blowing into a kind of hollow stick and moving his fingers about on the wood, thus drawing from it a variety of notes similar to those of a bird, though by quite a different method. Puzzled, and led on by his natural curiosity, he gave the boy a calf in exchange for his recorder and retired to solitude. Realizing that if he had not chanced to meet the boy he would never have learned of the existence of two methods for forming musical notes and very sweet songs, he tried traveling far from his home in the hope of meeting with some new adventure. The very next day he happened to pass near a small hut, and, hearing a similar tone within, he went inside to find out whether it was a recorder or a blackbird. There he found a boy holding a bow in his right hand and sawing upon some fibres stretched upon a concave piece of wood. The fingers of the left hand (which supported the instrument) were moving, and without blowing the boy was drawing from it various notes, and most sweet ones too. Now, you who are participating in this man's mind and sharing in his curiosity, judge his astonishment! Finding himself to have two unexpected new ways of forming tones and melodies, he began to believe that still others might exist in nature. His wonder increased when upon entering a certain temple he glanced behind the gates to learn what it was that had sounded, and perceived that the noise had emanated from the hinges and fastenings as he had opened the gate. Again, impelled by curiosity, he entered an inn expecting to see someone lightly bowing the strings of a violin, and instead saw a man rubbing the tip of his finger round the rim of a goblet and drawing forth from it a very sweet sound. And later he observed that wasps, mosquitoes, and flies did not form separate notes from their breaths, as did his original birds, but made steady tones by the swift beating of their wings.

In proportion as his amazement grew, his belief diminished that he knew how sounds were created; nor could all his previous experience have sufficed to make him understand or even believe that crickets, which do not fly, could draw their sweet and sonorous shrilling not from breath but from a scraping of wings. And when he had almost come to believe that there could be no further ways of forming notes—after having observed in addition to what has been recounted numerous organs, trumpets, fifes, stringed instruments of various sorts, and even that little iron tongue which when placed between the teeth makes strange use of the buccal cavity as a sounding box and of the breath as a vehicle of sound—when, I say, he believed that he had seen everything, he found himself more than ever wrapped in ignorance and bafflement upon capturing in his hand a cicada, for neither by closing its mouth nor by stopping its wings could he diminish its strident sound, and yet he could not see it move either its scales or any other parts. At length, lifting up the armor of its chest and seeing beneath this some thin, hard ligaments, he believed that the sound was coming from a shaking of these, and he resolved to break them in order to silence it. But everything failed until, driving the needle too deep, he transfixed the creature and took away its life with its voice, so that even then he could not make sure whether the song had originated in those ligaments. Thereupon his knowledge was reduced to such diffidence that when asked how sounds are generated he used to reply tolerantly that although he knew some of the ways, he was certain that many more existed which were unknown and unimaginable. ◘

[3]*The Controversy on the Comets of 1618*, Stillman Drake and C. D. O'Malley, trans., Philadelphia: University of Pennsylvania Press, 1960, pp.183–84. This quotation from "The Assayer," and the excerpt that follows (pp. 235–36), were translated from the Italian by Stillman Drake.

Art by Yury Vashchenko

Squeaky doors, squealing tires, and singing violins

All brought to you by dry friction—don't leave home without it

by I. Slobodetsky

WE COME ACROSS FRICTION at every step. Or better yet, we can't take a step without friction. But in spite of the crucial role friction plays in our life, we still don't have a complete picture of how friction arises. It's not because friction is so complicated by nature; it's just that experiments with friction are very sensitive to the surfaces used—that's why they're so hard to reproduce.

Here's an example. The English physicist Hardy investigated the temperature dependence of friction between glass plates. He carefully washed the plates with bleach (calcium hypochlorite solution) and rinsed them with water, removing grease and other impurities. Friction increased with temperature. The experiment was repeated many times, and the results obtained were approximately the same. But once, while washing the plates, Hardy rubbed them with his fingers. Friction stopped depending on temperature. Hardy thought that by rubbing the plates he had removed a very thin layer of glass with altered properties due to the interaction with bleach and water.

When we talk about friction, we can distinguish three slightly different phenomena: resistance when a body is moving in liquid or gas—this is called *fluid friction*; resistance that arises when a body slides over a surface—*sliding* or *dry friction*; and resistance that arises when a body rolls—*rolling friction*. This article is devoted to dry friction.

The first investigations of friction that we know about were performed by Leonardo da Vinci approximately 450 years ago. He measured the frictional force acting on wooden blocks as they slid across a board. Placing the blocks on their various faces, he discovered the independence of the frictional force on the area of the surface touching the board. But da Vinci never published his results. They became known only after the classical laws of friction were discovered by the French scientists Guillaume Amontons and Charles-Augustin de Coulomb.

These laws were as follows. (1) The frictional force F is directly proportional to the force N—the normal (perpendicular) force of the body on the surface over which it is moving: $F = kN$, where k is a dimensionless coefficient called the coefficient of friction. (2) The frictional force does not depend on the area of contact between the surfaces. (3) The coefficient of friction depends on the properties of the surfaces that rub together. (4) The frictional force does not depend on the body's velocity.

Three hundred years of friction research have corroborated the validity of the first three laws proposed by Amontons and Coulomb. Only the last one proved to be incorrect. But that was discovered much later, when railroads came on the scene and engineers noticed that the trains didn't behave as they expected when they applied the brakes.

Amontons and Coulomb explained the nature of friction rather simply. Both surfaces are uneven—they are covered with small bumps and pits. When a body is in motion, the bumps slide over one another, and so the body keeps going up and down. A certain force must be applied to drag the body onto these "hills." If the bumps are larger, the force must be increased as well. But this explanation contradicts one very important phenomenon: mechanical energy is expended on friction. A cube sliding over a horizontal surface stops. Its mechanical energy is transformed into heat by the friction. But in going up and down, the body doesn't use up its energy. Think of a roller coaster. When the cars roll down, their potential energy is converted into kinetic energy and they go faster; when the cars go up the next hill, their kinetic energy is converted back into potential energy. The mechanical energy of the cars decreases because of friction, not because they are going up and down. The movement of one body on the surface of another is analogous. Here again the energy loss can have nothing to do with the fact that the bumps of one body climb up the bumps of the other.

There are still other objections. For example, simple experiments to measure the frictional force between two polished glass plates showed that af-

Art by Dmitry Krymov

ter the surfaces were polished even more, the frictional force at first didn't change, and then it increased—it didn't decrease the way we'd expect from the model proposed by Amontons and Coulomb.

The mechanism of friction is much more complicated. Because of the unevenness of the surfaces, they come in contact only at individual points on the tops of the bumps. Here molecules of the two bodies are separated by distances approximately equal to the intermolecular distance within the bodies themselves, and the molecules interact. A stable bond is formed, which is broken when one body is pushed. While the body is moving, bonds are continuously formed and broken. At the same time molecular oscillations arise. And that's where the mechanical energy goes—into these oscillations.

The actual area of contact is usually from 1,000 to 2,000 square microns. In practice it doesn't depend on the size of the body—it's determined by the nature of the surfaces, their finish, the temperature, and the normal force. If we press the surfaces together, the bumps are crushed and the actual area of contact increases. So does the frictional force.

If the roughness of the surfaces is considerable, the mechanical catching of the hills begins to play a large role in increasing the frictional force. During movement the hills are crushed, and this gives rise to molecular oscillations as well.

Now the experiment with the polished glass plates makes sense. When the surfaces were still rough, the number of contact points was small; after a good polishing, that number increased.

Here's another example of increasing the frictional force by working the surface. If we take two metal bars with clean polished surfaces, they stick together. Here friction is great, because the actual area of contact is large. The molecular bonding forces responsible for friction turn the two bars into a monolith![1]

You can do the following experiment at home. Tie a string to the stem of a wine glass. Put it on a piece of glass and pull it. Now moisten the piece of glass and the base of the wine glass with water, which will wash off any grease and dirt. Pull the wine glass again. Now it's much more difficult to do it. It's hard to break a clean glass–glass contact. If you look closely at the surface, you may even notice scratches. It was easier to rip out pieces of glass than to break contact!

The model of friction we've been using is rather crude. We didn't stop to look at molecular diffusion—that is, the penetration of molecules of one body into the other; the role of electrical charges that arise on touching surfaces; the role of lubrication and how a lubricant works. These are still open questions, to some extent. We can only wonder why such a complicated phenomenon as friction is described by such a simple law: $F = kN$. And though the coefficient of friction isn't really constant and changes a bit from one point on the surface to another, we can make pretty good estimates of frictional forces for the many surfaces we come across so often in technical applications.

Dry friction has one important peculiarity: *static friction*. Whereas friction arises in a liquid or gas only when the body is in motion, and it's possible to move it with very little force, with dry friction a body starts to move only when the component of the applied force F parallel to the surface exceeds a certain magnitude (see figure 1). Until the body starts to slide, the frictional force is equal to the parallel component of the applied force but in the opposite direction. When the applied force is increased, the static frictional force also increases until it reaches its maximum value, usually greater than kN, at which point the body begins to slide. After that the frictional force is equal to kN.

This is often forgotten when the time comes to solve problems. Question: "What frictional force acts on a table weighing 300 N if the coefficient of friction is 0.4?" Most students will confidently answer: "120 N." Wrong! The frictional force is equal to zero—otherwise the table would move in the direction of the frictional force, since there are no other horizontal forces.

So, to move a body at rest, we must apply a force greater than the maximum possible static frictional force caused by the strength of the molecular bonds. But what if the body is already moving? What force must be applied to make the body move in a different direction? It turns out to be infinitesimal. This can be explained by the fact that the frictional force can't exceed the static frictional force.

Try this simple experiment. Take a book and put one edge on another, thicker book. You'll get an inclined plane. Now attach a thread to a matchbox and put the matchbox on the inclined plane. If the box slides, decrease the angle of inclination—use a thinner book for support. Pull the box to the side. It will slide downward too! Decrease the angle of inclination and pull the thread again. The same phenomenon occurs. The box slides down even with very small angles of inclination. The frictional force that held the box on the plane has somehow become very small.

Let's try to figure out what's going on here. If the matchbox moved only horizontally, it would be acted upon by a frictional force antiparallel to the velocity and equal to kN. To prevent the box from sliding down, there must be a frictional force acting on it that is directed up the inclined plane and equal to the component of the weight of the box parallel to the inclined plane. The resultant of these two frictional forces would be more than kN, and that can't be true. So the box must slide down the inclined plane.

Let's take a block, attach a string to it, and put the block on a horizontal plane. Pull it by the string so that it has a constant velocity $\mathbf{v}_1$ (fig. 2). Applying a force perpendicular to $\mathbf{v}_1$, we can make the block move in this direction with the velocity $\mathbf{v}_2$. The frictional force will be equal to kN and will have the direction opposite to the velocity $\mathbf{v}$ of the block relative to the plane ($\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$). Let's break the frictional force into two components along the directions of velocities $\mathbf{v}_1$ and $\mathbf{v}_2$:

$$F_1 = F_{fr} \cos \beta, \quad F_2 = F_{fr} \sin \beta,$$

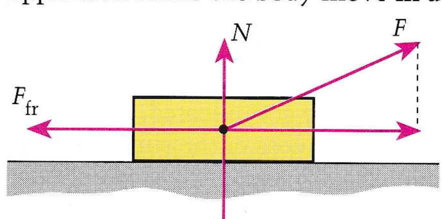


Figure 1

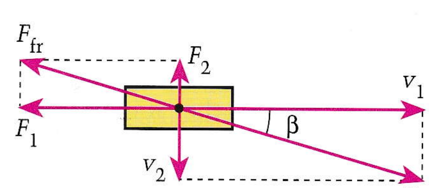


Figure 2

[1]In its literal sense (a "single stone").—*Ed.*

where β is the angle between the velocities $\mathbf{v}_1$ and $\mathbf{v}_2$, and $\tan\beta = v_2/v_1$.

The component F_1 of the frictional force counteracts the tension in the string, and the component F_2 counteracts the "sideways" force applied to the block. Since

$$\sin\beta = \frac{\tan\beta}{\sqrt{1+\tan^2\beta}},$$

then

$$F_2 = F_{fr}\frac{\dfrac{v_2}{v_1}}{\sqrt{1+\left(\dfrac{v_2}{v_1}\right)^2}} = F_{fr}\frac{v_2}{\sqrt{v_1^2+v_2^2}}.$$

If $v_2 << v_1$, then angle β is small and $\sin\beta \cong \tan\beta$. In this case

$$F_2 = F_{fr}\tan\beta = kN\frac{v_2}{v_1},$$

and the component of the frictional force that prevents the block from moving sideways is proportional to the speed of this movement. So we get a scenario identical to the one for small velocities with liquid friction. This means that a moving block can be made to move also in the perpendicular direction by applying an infinitesimally small force.

An interesting conclusion can be drawn for a box sliding on an inclined plane (fig. 3). Here $F_2 = W\sin\alpha$ and $N = W\cos\alpha$ (W is the weight of the box and α is the plane's angle relative to the horizontal). So

$$W\sin\alpha = k\cdot W\cos\alpha\frac{v_2}{\sqrt{v_1^2+v_2^2}},$$

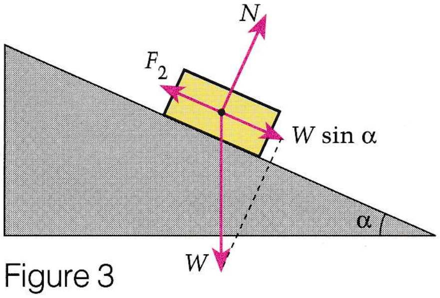


Figure 3

from which we get

$$v_2 = v_1\frac{\tan\alpha}{\sqrt{k^2-\tan^2\alpha}}.$$

(Of course, this holds only if $\tan\alpha < k$, because at larger angles of inclination the frictional force can't hold the box on the plane.)

At small angles of inclination (such that $\tan\alpha << k$),

$$v_2 = v_1\frac{\tan\alpha}{k}$$

—that is, the speed at which the box slides down the inclined plane is proportional to the speed at which it moves across the inclined plane and to the tangent of the angle of inclination.

This conclusion can easily be proved by experiment. In uniform motion the displacement of a body is proportional to its velocity, and the ratio of the speeds v_2 and v_1 will equal the ratio of the distances it travels in these directions.

This phenomenon is far from rare. For example, it's known that when an electric motor is suddenly stopped, the drive belt often jumps off the pulleys. Why? When the motor is stopped, the belt starts to slide relative to the pulleys, and it doesn't take much force to pull the belt sideways. Since the pulleys and the belt are usually set a little askew, this force is a component of the tension in the belt. Here are some other examples. When you want to pull a nail out of the wall without pliers, you bend it and pull, simultaneously turning it about its axis. For the same reason, a car skids and goes out of control when you brake hard. The wheels are sliding over the road, and a lateral force arises due to the unevenness of the road.

Now let's look at the last Amontons–Coulomb law: "the frictional force does not depend on the speed of the body." This isn't quite true.

The question of how the frictional force is related to speed is of great practical significance. And though experiments in this area encounter many specific complications, they are justified by the application of the information obtained to the theory of metal cutting, to calculations of the movement of bullets and shells in the barrels of weapons, and so on.

It's generally thought that to start a body moving you must apply a greater force than is necessary to pull it at a constant speed. In most cases this has to do with impurities on the surfaces of the bodies being rubbed together. For example, such a sharp increase in the frictional force is not observed with pure metals.

Experiments with the motion of a bullet in a barrel demonstrated that as the bullet's velocity increases, the frictional force decreases—at first quickly, then more and more slowly; but at speeds exceeding 100 m/s, it starts increasing. Figure 4 shows the graph of frictional force versus velocity. It can be roughly explained by the fact that in the area of contact a great deal of heat is released. At speeds of about 100 m/s, the temperature of the contact area may be several thousand degrees, and a layer of molten metal forms between the surfaces. Dry friction turns into fluid friction. And at high speeds, liquid friction is proportional to the square of the velocity.

It's interesting that the dependence of the frictional force on velocity is approximately the same for a violin bow and string. That's why we can listen to music performed on bowed instruments like the violin, cello, or viola.

When the motion of the bow is uniform, it pulls the string aside and stretches it (fig. 5). As the stretch increases, the frictional force between the bow and string increases. When the frictional force reaches the maximum, the string starts to slide relative to the bow. If the frictional force didn't depend on the relative speed of

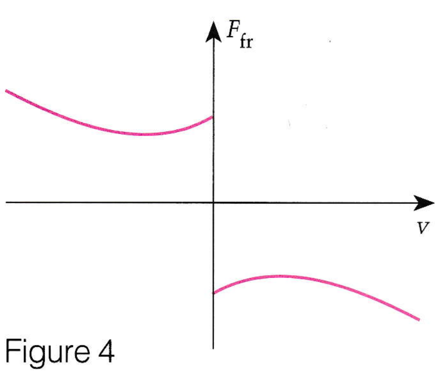


Figure 4

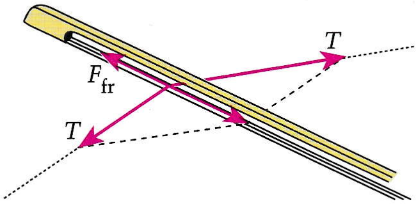


Figure 5

the bow and string, then obviously the displacement of the string from the equilibrium position would not change.

But when sliding occurs, the friction decreases. That's why the string starts to move to the equilibrium position. The relative speed of the string increases, which decreases the frictional force. And when the string rebounds, its speed relative to the bow decreases, and the bow grabs the string again. Then everything repeats. And that's how vibrations are generated in the string. These are sustained vibrations, because the energy lost by the string while sliding is continuously replenished by the work of the frictional force pulling the string to the position at which the string breaks free.

On that happy note, I'll bring this article on dry friction to a close. We still don't understand it completely, but we can describe it with laws that give us accurate enough results. And this allows us to explain many physical phenomena and to make the calculations necessary to build machines.

Problems

1. Why does a car turn when its front wheels turn?

2. Balance a stick horizontally on your outstretched index fingers. Slowly move your right hand to the left. Why does the stick also move, and why does it remain balanced?[2]

3. Draw a graph of the dependence of the frictional force acting on a block placed on an inclined plane on the plane's angle of inclination.

4. Draw a graph of the dependence of the frictional force acting on a block placed on a horizontal plane on the angle between the applied force and horizontal. The applied force is less than the weight of the block, and the angle varies from 0 to $\pi/2$.

5. Pipes that break during the drilling of wells can be lifted out by means of the device shown in figure 6. The hinged arms AB and AC are attached to a cable at point A. The pipe moves upward because of the friction of the arms against the pipe. Determine the condition under which this mechanism can lift pipes of any weight. Naturally, you can assume that the cable is strong enough.

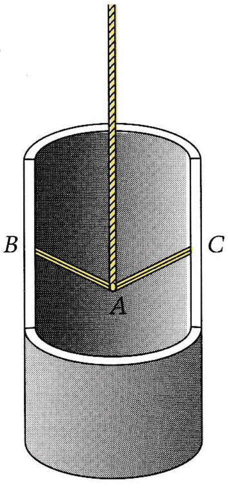


Figure 6

6. A small cube of mass m is placed on a rough inclined plane whose angle of inclination is α. The coefficient of friction is $k = 2 \tan \alpha$. Define the minimum horizontal force F (see figure 7) needed to move the cube.

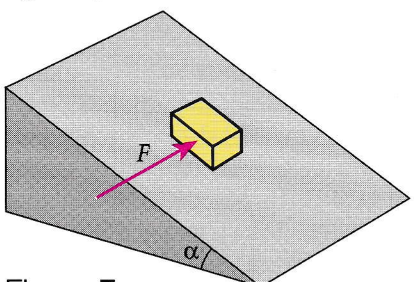


Figure 7

7. A flywheel of radius $R = 20$ cm is mounted on a fixed axle of radius $r = 2$ cm. You can remove the flywheel by pulling it with a force $F = 1{,}000$ N. To make it easier to remove the flywheel, a force $F_1 = 80$ N is applied to its rim, creating a torque relative to the axle. What is the minimum force F_2 needed to pull the flywheel along the axle?

8. Why does a badly lubricated door squeak?

ANSWERS, HINTS & SOLUTIONS ON PAGE 59

[2]See also "What the Seesaw Taught" in the January/February 1991 issue of *Quantum*.—*Ed.*

US wins gold at the International Physics Olympiad

Exciting challenges and reindeer steaks in Helsinki

by Larry D. Kirkpatrick

AMERICAN PHYSICS STUDENTS once again proved themselves to be among the best in the world at the XXIII International Physics Olympiad held in Helsinki, Finland, on July 5–13, 1992. All five members of the US Physics Team won awards, including 2 gold medals, 1 silver medal, and 2 honorable mentions. In this year's Olympiad 177 competitors from 37 countries won a total of 13 gold medals, 19 silver medals, 25 bronze medals, and 33 honorable mentions.

The US effort was led by the gold medal performance of Eric Miller (San Rafael, California), who scored 41 out of a possible 50 points. Eric's score was fifth in the world and quite close to the top score of 44. As a junior Eric won a bronze medal at the Olympiad held in Havana, Cuba. Next year Eric will attend Harvard University. The second gold medal was won by Szymon Rusinkiewicz (Houston, Texas). The fact that Szymon came to the United States at the age of five from Poland was not lost on the Polish team as they tried to claim a share of the gold medal. Szymon will be attending Cornell University in the fall.

The silver medal was won by Michael Schulz of Baldwin, New York, who will be attending MIT this fall. The two juniors on the team, Carwil James (East Cleveland, Ohio) and Dean Jens (Ankeny, Iowa), were awarded honorable mentions. Carwil has decided to pass up his last year of high school to attend Northwestern University, but Dean hopes to return to next year's team and bring along his twin brother Steve.

Team leaders for the US Physics Team are Larry Kirkpatrick (Montana State University) and Avi Hauser (AT&T Bell Laboratories). They are assisted by Ted Vittatoe (Libertyville High School, Libertyville, Illinois). The US Physics Team is organized by the American Association of Physics Teachers. Funding is organized by the American Institute of Physics, which received major contributions from AT&T and IBM.

Left to right: Avi Hauser, Michael Schulz, Szymon Rusinkiewicz, Dean Jens, Eric Miller, Carwil James, Larry Kirkpatrick.

The exam

The Olympiad exam consists of a theoretical exam that counts for 60% of the total and an experimental exam. On the first exam day, students were asked to solve three very difficult problems in five hours without the aid of any books or tables. The first problem involved a satellite that had four masses in a ferris wheel arrangement. The masses rotated about the center of mass as the entire satellite orbited the Earth. The masses could be pulled or pushed by motors

A report on the 1992 International Mathematics Olympiad will appear in the next issue.

so that their distances to the center changed. The students were asked to calculate the forces on each mass, the work performed by the motors, and the resultant effects on the motion of the satellite.

The second problem required the students to analyze the vibrational modes of linear molecules using a model in which the forces between nearest atomic neighbors are approximated by springs. The third theoretical problem concerned the heating and cooling of a satellite in near-Earth orbit. After calculating the equilibrium temperature of the satellite, assuming that it was a black body, the students calculated what changes would occur with various suggested coatings on the surface. The top theoretical score was 26.5 out of a maximum of 30 points.

Two days later the students were asked to solve two experimental problems in 2.5 hours each. The first problem used an ingenious device in which a falling mass compressed a piezoelectric device that, in turn, produced a spark across an adjustable gap. The students investigated the relationship between the potential difference generated by the piezoelectric device and the spacing of the gap. The second experimental problem required the students to investigate the transmission properties of colored filters after determining the groove spacing of a piece of a compact disk. They were also required to investigate the diffraction produced by a fine wire mesh. The top experimental score was 19.25 out of a maximum of 20 points.

Only eleven countries were able to win five awards (gold–silver–bronze–honorable mention) each: Australia (0–1–1–3), China (5–0–0–0), Czechoslovakia (0–2–2–1), Germany (1–0–4–0), Great Britain (0–4–1–0), Hungary (0–0–2–3), Netherlands (0–2–2–1), Romania (1–1–1–2), Russia (3–1–1–0), Ukraine (1–3–1–0), and the US (2–1–0–2).

One of the highlights of this year's Olympiad was the participation of teams from several nations that had only recently gained their independence: Croatia, Estonia, Lithuania, Russia, Slovenia, and Ukraine. Our world is certainly changing!

Reindeer and saunas

The US Physics Team arrived in Helsinki on July 3 in time to take a walking tour of Helsinki and attend the Fourth of July picnic sponsored by the American Embassy. The next day was spent on a second walking tour, a practice experiment in the hotel, and dinner at a Finnish restaurant, where some of the team tasted reindeer meat for the first time. The team leaders also enjoyed participating in the Finnish tradition of the sauna. (There is approximately one sauna for every four people in Finland.)

Besides the examinations and the opportunity to meet talented physics students from around the world, the students were treated to tours, receptions, visits to the Heureka Science Center and the fortress on Suomenlinna Island, and musical and dance entertainment. One night the students and leaders were treated to a traditional sauna party.

Selection and training

The selection of the team began in November when the AAPT sent invitations to physics teachers across the US soliciting nominations of the best physics students in the country. The first selection examination was administered to 444 students. The exam consisted of 30 very challenging multiple-choice questions and four open-response problems. The top 75 students were given a second, harder examination in March. This exam consisted of four open-response questions to be answered in 60 minutes and two very difficult problems to be completed in two hours. The top 20 students were then invited to take part in a weeklong training camp at the University of Maryland during the last week in May. (In contrast, the Chinese train for five months, and the Russian team holds three training sessions for a total of five weeks.)

During the training camp the students enjoyed problem-solving sessions; tutorial lectures on optics, interference, thermodynamics, AC electricity, and selected topics in modern physics; frequent examinations; and a trip to Washington, D.C., to visit the Presidential Science Advisor, the Secretary of Education, and the Associate Director of the National Science Foundation. The students also had the opportunity to hear about the frontiers of physics from prominent scientists from the University of Maryland, AT&T Bell Laboratories, and IBM's Thomas J. Watson Research Center.

The 1993 competition

The United States will host the XXIV International Physics Olympiad at the College of William and Mary in Williamsburg, Virginia, from July 10–18, 1993. Students interested in becoming members of the US Physics Team should contact their high school physics teacher or Maria Elena Khoury, American Association of Physics Teachers, 5112 Berwyn Road, College Park, MD 20740 by December 31.

1992 Physics Team

The 1992 US Physics Olympiad Team drew its members from all across the country. In the list below, members who represented the team in Helsinki are marked by an asterisk; each member's physics teacher is noted in parentheses.

Ibrahim Abdullah, Stuyvesant High School, New York City (Mr. Tarrendash)

Michael S. Agney (alternate), Melbourne High School, Melbourne, Florida (Carolyn A. Ronchetti)

Keith Bradley, Parkway North High School, Creve Coeur, Missouri (David Lay)

Robert T. Brockman II, St. John's School, Houston, Texas (Mark Kinsey)

Mary Pat Campbell, NC School of Science and Math, Durham, North Carolina (Hugh Haskell)

Chang Shih Chan, Northeast High School, Philadelphia, Pennsylvania (Raj G. Rajan)

Eric Scott Dickson, Long Beach Polytech, Long Beach, California (James R. Outwater)

***Carwil James**, *honorable mention*, Hawken School, Gate Mills, Ohio (Robert Shurtz)

***Dean W. Jens**, *honorable mention*, Ankeny High School, Ankeny, Iowa (D. F. Savage)

Glenn L. Kashan, Riverdale Country School, Bronx, New York (Sankar Sengupta)

Jose M. Lorenzo, Miami Sunset Senior High School, Miami, Florida (David Jones)

***Eric Miller**, *gold medal*, San Francisco University High School, San Francisco, California (Tucker Hiatt)

Daniel J. Rabinowitz, Cherry Hill High School West, Cherry Hill, New Jersey (Hirendra Chatterjee)

***Szymon Rusinkiewicz**, *gold medal*, Bellaire High School, Bellaire, Texas (John E. Beam)

***Michael B. Schulz**, *silver medal*, Baldwin High School, Baldwin, New York (Dominick J. Capozzi)

David Scott, New Trier High School, Winnetka, Illinois (Alan A. Brix)

Noam Shomron, Mt. Olive High School, Flanders, New Jersey (Ronald Gounaud)

Michail Sunitsky, Stuyvesant High School, New York City (Mr. Tarrendash)

Milorad Todorovich, Libertyville High School, Libertyville, Illinois (Theodore W. Vittitoe)

Victoria Yung, Cherry Hill High School West, Cherry Hill, New Jersey (Hirendra Chatterjee)

Bulletin board

First step toward a Nobel Prize

The Institute of Physics of the Polish Academy of Sciences has organized a research competition in physics to encourage high school students to conduct research projects on their own. The competition is open to all students who have not begun their university studies or reached the age of 21 before March 31, 1993.

Papers describing the research do not have a specified format or content, but must have a single author and not exceed 20 typewritten (double-spaced) pages. Each paper must be in English and contain the name, birthdate, and home address of the author and the name and address of the author's school. All papers will be evaluated by the organizing committee and prizes awarded to the best papers. Winners will receive an invitation to spend one month at the Institute of Physics with all expenses paid. However, winners will be responsible for their own travel expenses to and from Warsaw, Poland.

Two copies of the research paper must be sent by March 31, 1993 to

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

Additional information is available by writing to Dr. Gorzkowski or contacting him via e-mail at gorzk@planif61.bitnet.

From egg to—what?

How does a microscopic fertilized egg turn into a fly, a chicken, you or me? Using the latest techniques of molecular biology, scientists are beginning to find new answers. Their findings are described in the 56-page report *From Egg to Adult*, published by the Howard Hughes Medical Institute. The full-color report tells about the dialogue between sperm and egg that leads to fertilization—one of the first examples of cell–cell communication. It describes the mother's role in enabling a fly embryo to know its head from its tail. The report also shows how scientists are discovering what makes individual cells of an embryo migrate, divide, or die. Such research is bringing new insights to the old question of how much is determined by heredity, how much by the environment. A separate article describes the development of the world's most complex system—the brain.

For a free copy of *From Egg to Adult*, write to the Howard Hughes Medical Institute, Communications Office, 6701 Rockledge Drive, Bethesda, MD 20817.

Conservation tips

The Appalachian Mountain Club has published *The Conservation**works** Book: Practical Conservation Tips for the Home and Outdoors*, written by Lisa Capone and illustrated by Cady Goldfield. This easy-to-read ecology handbook combines advice for conservation at home with tips for protecting the countryside when hiking or camping. There are dozens of suggestions for simple, effective changes everyone can make to reduce waste, save energy, and conserve resources. Humorous cartoons serve to illustrate the book's concepts while entertaining the reader.

*The Conservation**works** Book* is available for $7.95 in bookstores, or by contacting the publisher directly. Write to Appalachian Mountain Club Books, 5 Joy Street, Boston, MA 02108.

Computational chemistry

Chemists are obtaining better tools with which to study the properties and structure of molecules. Instead of spending months developing, screening, and examining molecules in the lab, they can model and analyze alternative molecular structures using HyperChem, a computer-aided chemistry package for PCs. HyperChem integrates sophisticated model-building and visualization capabilities with advanced computational methods, including molecular dynamics and classical and semiempirical quantum mechanics.

HyperChem runs under the Microsoft Windows™ interface on 386- and 486-class PCs. It features an open architecture and a built-in scripting language that allows users and software developers to extend and tailor HyperChem to meet specialized needs.

To request a demo disk or more information about HyperChem, write to Autodesk, Inc., Education Department, 2320 Marinship Way, Sausalito, CA 94965, or call 800 424-9737.

ANSWERS, HINTS & SOLUTIONS

Math

M66

Since the total number of victories of all teams equals the total number of defeats, and each team played seven games, the average number of wins per team is 3½. So at least one team—we'll call it A—has won at least four games. Similarly, the average number of victories per team in the games between the four teams beaten by A is 1½, so one of them—team B—has defeated at least two of the others. Of these two teams, one—team C—defeated the other, team D. The teams A, B, C, D constitute the desired four.

The statement of the problem is easily generalized: *if 2^{n-1} teams played in a tournament, then n teams $A_1, \ldots, A_n$ can be chosen such that each of them has beaten all the teams with greater numbers.*

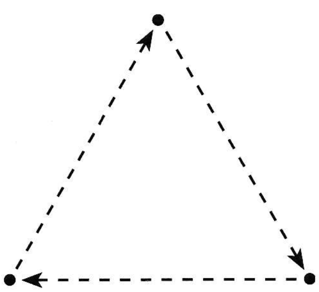


Figure 1

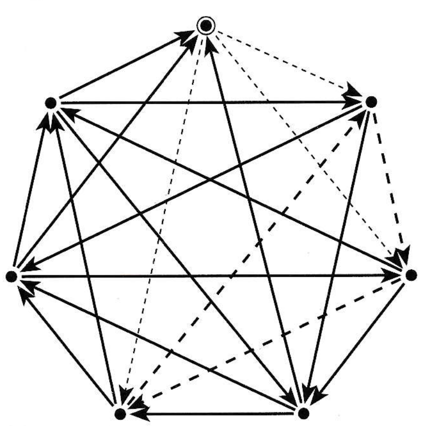


Figure 2

An interesting question is whether the total number N of teams in this statement can be *less* than 2^{n-1}. The tournament diagrams in figures 1 and 2 (where each arrow is drawn from a winner to a loser) show that for $n = 3$, $N = 3 = 2^2 - 1$, and for $n = 4$, $N = 7 = 2^3 - 1$, our statement proves to be wrong. (Note that in both figures the arrangement of arrows is the same for all teams, and in figure 2 any three teams beaten by a fourth form the diagram in figure 1.) So in these two cases, 2^{n-1} is the smallest admissible value of N. The cases of $n = 5, 6, \ldots$ are left to the reader.

M67

We can view the parabolas' lines of symmetry as the rectangular coordinate axes (see figure 3). Then the equations of the parabolas will be $x = ay^2 + b$ ($a > 0$, $b < 0$) and $y = cx^2 + d$ ($c > 0$, $d < 0$). The coordinates (x, y) of any common point of the parabolas

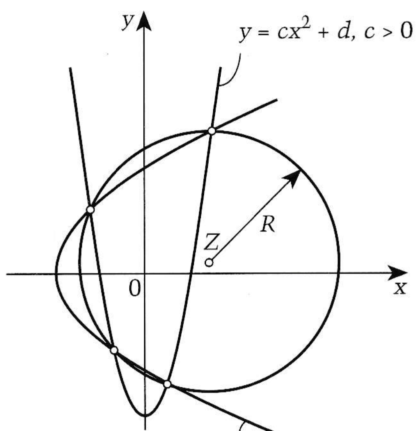


Figure 3

satisfy both these equations and, therefore, the equation obtained by dividing the first equation by a, the second one by c, and adding them together:

$$\frac{x}{a} + \frac{y}{c} = y^2 + \frac{b}{a} + x^2 + \frac{d}{c}.$$

Completing the squares, we arrive at

$$\left(x - \frac{1}{2a}\right)^2 + \left(y - \frac{1}{2c}\right)^2 = \frac{1}{4a^2} + \frac{1}{4b^2} - \frac{b}{a} - \frac{d}{c}.$$

Since all the terms on the right side are positive, this equation defines a circle with center $Z(1/2a, 1/2c)$ and radius equal to the square root of the right side. So all the common points of the parabolas lie on this circle.

M68

It is always possible to choose x and y such that $a = x^{10}$ and $b = y^{15}$. Plugging these into the inequality in question—

$$2x^5 + 3y^5 \geq 5x^2y^3$$

—and dividing this by y^5, we arrive at the inequality

$$f(t) = 2t^5 - 5t^2 + 3 \geq 0,$$

where $t = x/y > 0$. We must prove that $f(t) \geq 0$.

A standard approach is to use calculus. The derivative $f'(t) = 10t(t^3 - 1)$ is negative for $0 < t < 1$, positive for $t > 1$, and equals 0 for $t = 1$. So this function takes its minimal value at the point $t = 1$, and this value is $f(1) = 0$ (see the graph in figure 4); for all the other positive values of t, $f(t) \geq f(1) = 0$.

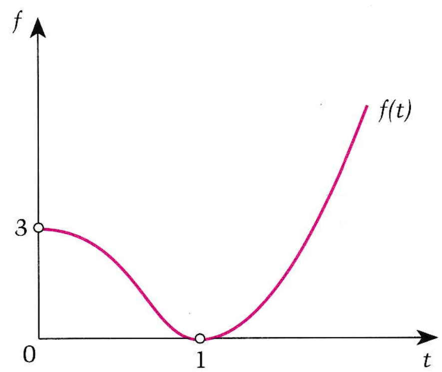


Figure 4

A more elementary proof, though a bit trickier, involves factoring the polynomial $f(t)$:

$$f(t) = (t-1)^2(2t^3 + 4t^2 + 6t + 3)$$

—both factors are obviously nonnegative. Perhaps the simplest proof, however, starts with the (not quite elementary) fact that the arithmetic mean of n positive numbers is never less than their geometric mean.[1] Taking the five numbers $\sqrt{a}$, $\sqrt{a}$, $\sqrt[3]{b}$, $\sqrt[3]{b}$, $\sqrt[3]{b}$, we find that

$$\frac{2\sqrt{a}+3\sqrt[3]{b}}{5} = \sqrt[5]{\left(\sqrt{a}\right)^2\left(\sqrt[3]{b}\right)^3} = \sqrt[5]{ab},$$

which is equivalent to the above result. This last method of proof can be generalized for the case of k positive numbers $a_1, a_2, \dots, a_k$ to show that

$$p_1 a_1^{1/p_1} + p_2 a_2^{1/p_2} + \dots + p_k a_k^{1/p_k} \geq p\left(a_1 a_2 \dots a_k\right)^{1/p},$$

where $p_1, p_2, \dots, p_k$ are positive integers, p is their sum.

M69

Write out all the divisors of n in increasing order: $1 = d_1 < d_2 < \dots < d_k = n$. Note that $n/d_2 = d_{k-1}, \dots, n/d_k = d_1$, and $d_k = n$. The number of positive integers not exceeding n and divisible by d_i is equal to n/d_i, so the number of integers that are not coprime with n—that is, $n - \phi(n)$—is not greater than

$$\frac{n}{d_2} + \frac{n}{d_3} + \dots + \frac{n}{d_k} = d_{k-1} + d_{k-2} + \dots + d_1 = \sigma(n) - n$$

Thus,

$$n - \phi(n) \leq \sigma(n) - n,$$

and we're done.

Note that for a prime p this estimate is precise: $\sigma(p) = p + 1$, $\phi(p) = p - 1$, $\sigma(p) + \phi(p) = 2p$. Are there any other numbers for which equality holds?

M70

Let K be the intersection of the variable line with CD, and let line AB intersect CP and CQ at M and N, respectively. We will show that $AM = BN$. Knowing this, we can use considerations of symmetry (or congruent triangles) to show that angles ACM and BCN are equal. The angles ACP and BCQ (given in the problem) are either identical to these angles or supplementary to them.

To show that $AM = BN$, we can use various pairs of similar triangles to find that

$$\frac{DC}{AM} = \frac{DP}{AP} = \frac{DK}{AH}$$

and

$$\frac{DC}{BN} = \frac{DQ}{BQ} = \frac{DK}{BH}.$$

Since $AH = BH$, it follows that $AM = BN$.

This problem can also be solved without any calculations at all, by going "off into space" (see the article of that name in the January/February 1992 issue). We consider a central projection of our plane onto some other plane p, passing through line AB. We choose the center of projection O such that plane OCD will be parallel to p. Under this central projection (see figures 5a and 5b), both C and D map onto points at infinity, so the images of any two lines passing through point C will be parallel, as will the images of any two lines passing through D. Figure 5b shows the parallel lines that are images of MC and NC. It also shows parallel lines $P'A$ and $Q'B$, which are the images of AD and BD, respectively.

Therefore, the images P' and Q' of points P and Q will be symmetric with respect to H. Using figure 5b, we can see that points M and N (which are left fixed by our central projection) are therefore also symmetric with respect to H. This means that $AM = BN$, and our result follows.

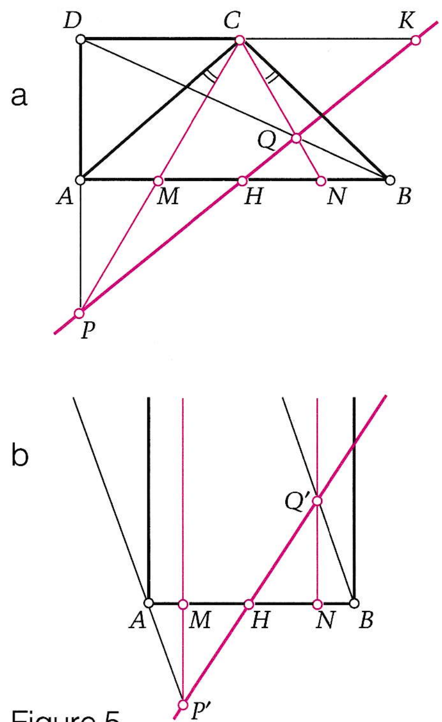


Figure 5

To find out more about properties of central projections, see, for example, *Geometric Transformations* by I. M. Yaglom (Washington: Mathematical Association of America, 1973). (I. Sharygin, V. Dubrovsky)

Physics

P66

Since there is no friction, there are no external forces acting on the system in the horizontal direction (fig. 6). In order to determine the velocity v of the left wedge and the velocity u of the washer immediately after the descent, we can use the energy and momentum conservation laws:

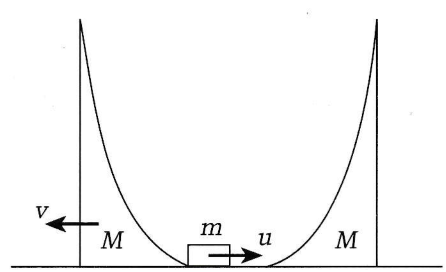


Figure 6

[1]An elegant proof of the arithmetic–geometric mean inequality was given by Cauchy. See, for example, Beckenbach and Bellman, *An Introduction to Inequalities* (Washington: Mathematical Association of America, 1961, pp. 47–60).

$$\frac{Mv^2}{2} + \frac{mu^2}{2} = mgh, \quad Mv = mu.$$

At the moment the washer reaches its maximum height h_{max} on the right-hand wedge, the velocities of the washer and the wedge are equal. Therefore, conservation of momentum can be written in the form

$$mu = (M + m)V,$$

where V is the velocity of the washer and the right-hand wedge. Let's also use conservation of energy on the right-hand side:

$$\frac{mu^2}{2} = \frac{M+m}{2}V^2 + mgh_{max}.$$

Solving the four simultaneous equations leads to the expression for the maximum height h_{max}:

$$h_{max} = h\frac{M^2}{(M+m)^2}.$$

P67

The mass has a weight $W = mg$ = 60 N that is larger than the applied force F_1 = 40 N. The difference of 20 N must be supplied by the frictional force between the rope and the beam. You might guess that F_2 = 60 N + 20 N = 80 N because the frictional force reverses direction when you try to pull the mass upward. This is a good first approximation, but the value is too small. As you pull harder on the rope, the rope is pulled tighter against the beam, increasing the frictional force. Therefore, we need to take a closer look at the problem to get a better answer.

A complete calculation of the frictional forces as a function of the position around the beam is rather complex as the tension in the rope changes from F_1 at the left end of the rope to $W = mg$ at the right end. (This must be true since each end of the rope is in equilibrium.) However, to solve this problem it is sufficient to note that the frictional force at each point is proportional to the tension in the rope at that point. Therefore, we can take the total frictional force F_{fr} as being proportional to the largest tension—that is, $F_{fr} = kW$. This means that $F_1 = W - kW$, and the ratio of the largest tension to the smallest tension is a constant given by $W/F_1 = 1/(1-k)$.

In the second case, when we want to lift the load, the two ends of the rope reverse roles as the frictional force is now opposite the applied force F_2. The ratio of the largest tension F_2 to the smallest tension W must be the same as it was in the first case: $F_2/W = W/F_1$. From this we get

$$F_2 = \frac{W^2}{F_1} = 90 \text{ N}.$$

P68

The total amount of heat Q emitted into space per unit time remains unchanged, since it is determined by the energy transformed into heat in the operation of the station's equipment. Since only the outer surface of the shell radiates into space and this radiation depends only on its temperature, the temperature of the shell must be equal to the initial temperature T = 500 K of the station. However, the shell radiates the same amount of heat Q inward. This radiation reaches the surface of the station and is absorbed. Therefore, the total amount of heat supplied to the station per unit time is the sum of the heat Q from the equipment and the amount of heat Q absorbed at its inner surface—that is, it is equal to $2Q$. According to the conservation of energy, the same amount of heat must be radiated from the surface of the station. Therefore,

$$\frac{Q}{2Q} = \frac{T^4}{T_x^4},$$

where T_x is the required temperature of the station's shell. Finally, we get

$$T_x = \sqrt[4]{2}T \cong 600 \text{ K}.$$

P69

The electrostatic potential energy of a charged body is known to be equal to the total potential energy of the interaction of all possible pairs of small charged areas constituting the charged body.

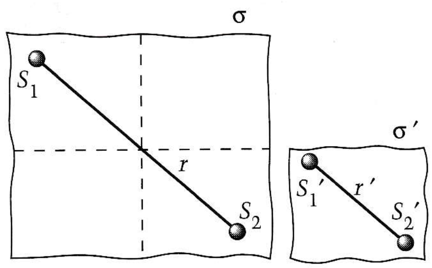


Figure 7

Take two arbitrary small areas (fig. 7) on a charged square sheet; the areas are S_1 and S_2, respectively, and the distance between them is r. Then the interaction energy of the two selected areas is

$$\Delta W_{12} = \frac{1}{4\pi\varepsilon_0}\frac{(\sigma S_1)(\sigma S_2)}{r},$$

where σ is the surface density of the charge.

For the folded sheet, we'll take two areas corresponding to the first pair; the areas are $S_1' = S_1/4$ and $S_2' = S_2/4$, respectively, and the distance is $r' = r/2$. Then the interaction energy of the new pair of areas is

$$\Delta W_{12}' = \frac{1}{4\pi\varepsilon_0}\frac{(\sigma' S_1')(\sigma' S_2')}{r'} = \frac{1}{4\pi\varepsilon_0}\frac{S_1 S_2 \sigma^2}{r/2} = 2\Delta W_{12}$$

(we've taken into consideration the fact that after folding the sheet the surface density of the charge is four times larger: $\sigma' = 4\sigma$).

Therefore, the contribution to the total energy of any pair of areas in the second case will be twice that of the corresponding pair of areas in the first case. Thus, the total energy of the sheet after folding is twice the total energy before folding.

We'll obtain the same answer if we use dimensional analysis. (See "The Power of Dimensional Thinking" in the May/June 1992 issue.) We'll take into account the fact that the total energy of a uniformly charged square will be proportional to the square of its total charge and inversely proportional to the length of its side. After folding it in half we'll get a situation

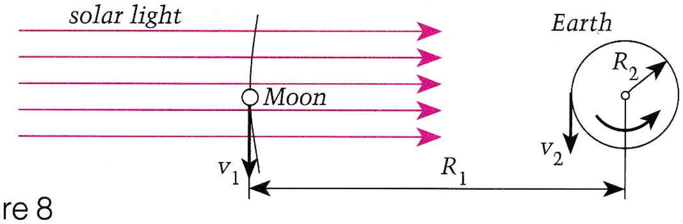


Figure 8

similar to the initial one. The total charge of the square is unchanged, while the dimensions have been halved. Thus, the total energy of the square is doubled.

P70

To calculate the velocity of the Moon's shadow we'll take the distance between the Moon and the Earth to be R_1 = 384,000 km and the period of rotation around the Earth to be T_1 = 28 days = 2,419,200 s (see figure 8). Similarly we'll assume that the Earth's radius R_2 = 6,400 km and that the Earth completes a full rotation about its axis in the time T_2 = 1 day = 86,400 s. Then the velocity of the Moon's revolution along a circular orbit is

$$v_1 = \frac{2\pi R_1}{T_1} \cong 995 \text{ m/s},$$

and the linear velocity of points on the Earth's surface is

$$v_2 = \frac{2\pi R_2}{T_2} \cong 465 \text{ m/s}.$$

(In the area on the Earth's surface where the Moon's shadow is observed, both velocities are in the same direction.)

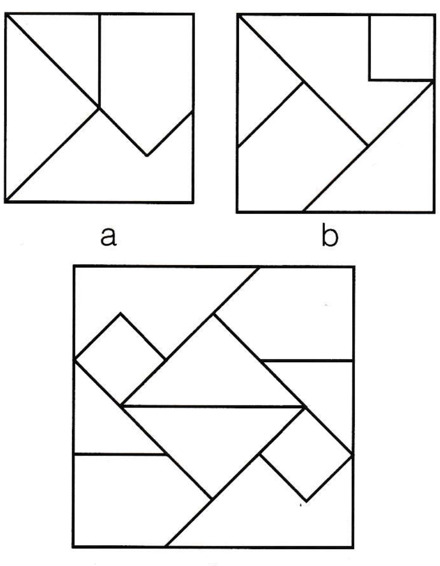


Figure 9

As the Sun's rays can be considered parallel and the Moon's dimensions are small, the shadow of the Moon moves over the Earth's surface at noon at the velocity

$$v = v_1 - v_2 = 2\pi\left(\frac{R_1}{T_1} - \frac{R_2}{T_2}\right) \cong 530 \text{ m/s}.$$

This is a speed of approximately 1,200 mph!

Brainteasers

B66

Let HE = x; then $x^2 - x$ is a three-digit number ending in two zeros—that is, it's divisible by 100. Since $x^2 - x = x(x - 1)$ is a product of two coprime numbers, one of them must be divisible by 25 and the other one by 4. If x or $x - 1$ equals $25k$, where $k \geq 2$, then the product $x(x - 1) \geq 50 \cdot 49$ and consists of more than three digits. So either $x - 1 = 25$ and $x = 26$ (which isn't divisible by 4), or $x = 25$ and $x - 1 = 24$, which yields the unique solution: $625 = 25^2$. (V. Dubrovsky)

B67

See figures 9a through 9c.

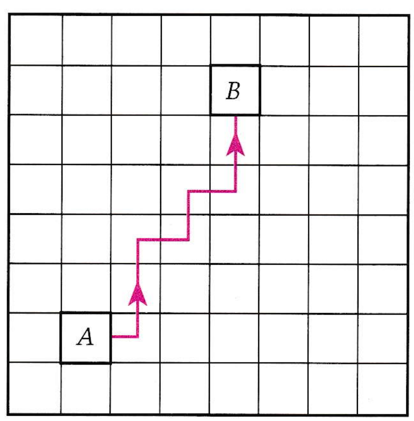


Figure 10

B68

The cavity of the thermos can be considered a resonator that amplifies sound frequencies close to the natural frequencies of the cavity. When you fill the thermos, the noise produced by the liquid has a wide spectrum of frequencies; but only the frequencies close to the resonant frequencies are amplified and are therefore audible to us. While the thermos is being filled, the length of the cavity decreases and the wavelengths of the resonant frequencies become shorter as well. As a result, the pitch of the tone should get higher.

B69

If the last two digits of a number are not 99, then, in passing to the next higher number, we either increase the digit-sum by 1 or (if the last digit was 9) decrease it by 8. So at least one of two such "next numbers" can't be divisible by 17. When a number ending in 99 (but not 999) is increased by 1, the digit-sum decreases by 9 + 9 – 1 = 17, which is just what we need. However, numbers ending in 999 (but not 9999), when increased by 1, decrease their digit-sum by 26, which is not divisible by 17. Therefore, we want the smallest number ending in 00 whose digit-sum is divisible by 17. This is 8,900, so the answer is 8,900 – 1 = 8,899.

B70

We can rotate the checkerboard so that the square A with the sum of 3 cents will be located below and to the left of the square B with 17 cents (fig. 10). Let's move from A to B by

B

10	11	12	13	14	15	16	17
9	10	11	12	13	14	15	16
8	9	10	11	12	13	14	15
7	8	9	10	11	12	13	14
6	7	8	9	10	11	12	13
5	6	7	8	9	10	11	12
4	5	6	7	8	9	10	11
3	4	5	6	7	8	9	10

A

making one-square steps up or to the right. Since every step changes the sum in the current square by one and the numbers of right steps and up steps are both not greater than 7, the total increase of the sum is not greater than 7 + 7 = 14. In fact, it is exactly 14 (17 – 3). Therefore, squares A and B must be the bottom left and top right squares of the checkerboard, and whichever path we choose, each step must increase the sum by one; so after *any n* steps we find ourselves in a square with the sum of 3 + n cents. This uniquely determines the arrangement of sums—see figure 10. So the answer is (3 + 5 + ... + 17) + 8 · 10 = 160.

Poincaria

1. The first two assertions are simply particular cases of the fundamental properties of inversion mentioned in the article.

Further, recall that p-distance $d(A, B)$ for the circular version of the Poincaré model was defined in the article "Inversion" (in the last issue) as

$$d(A, B) = \ln \{AB, ab\},$$

where $\{Ab, ab\} = (Aa \cdot Bb)/(Ab \cdot Ba)$ is the "cross ratio" of four points: A, B, and the "endpoints" a and b of p-line AB (B lying between A and a). First note that exactly the same formula can be applied in the case of the model in the upper half-plane to compute the p-distance $\rho(A, B)$ between points A and B on p-line $L(a, b)$ (fig. 11). Indeed, by definition,

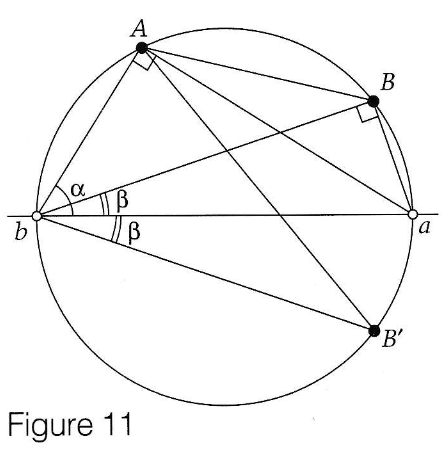


Figure 11

$$\rho(A, B) = \ln\frac{AB' + AB}{AB' - AB},$$

where B' is the reflection of B in the x-axis. If $L(a, b)$ is a semicircle with diameter $D = ab$, then in the notation of figure 11, by the extended law of sines $AB = D \sin(\angle AbB) = D \sin(\alpha - \beta)$, $AB' = D \sin(\alpha + \beta)$, and

$$\begin{aligned}\rho(A, B) &= \ln\frac{\sin(\alpha+\beta)+\sin(\alpha-\beta)}{\sin(\alpha+\beta)-\sin(\alpha-\beta)}\\ &= \ln\frac{\sin\alpha\cos\beta}{\cos\alpha\sin\beta}\\ &= \ln\frac{\tan\alpha}{\tan\beta}\\ &= \ln\frac{Aa/Ab}{Ba/Bb}\\ &= \ln\{AB, ab\}.\end{aligned}$$

The case when $L(a, b)$ is a ray ($a = \infty$ or $b = \infty$) is left to the reader (we make the agreement that $A\infty/B\infty = 1$).

All that remains is to note that the inversion I transforming one model into the other preserves cross ratios (like any inversion) and takes the "endpoints" of a p-line in one model into those of its image, a p-line in the other model.

2. Since the inversion I considered in the above solution preserves p-distances, computed by the formulas for the corresponding models, I takes p-circles of one model into those of the other. Like any inversion, it also takes Euclidean circles into Euclidean circles. So we can use either model to prove that p-circles are Euclidean circles. The circular model is more convenient. If O is the center of the circle representing the entire p-plane in this model, then any p-circle with center O is a Euclidean circle, because the condition $d(O, X)$ = constant is equivalent to OX = constant. A p-circle ω with center C other than O can be obtained from some p-circle ω_0 with center O under the p-reflection taking O into C. We know that ω_0 is a Euclidean circle. So its p-reflection ω is also a Euclidean circle.

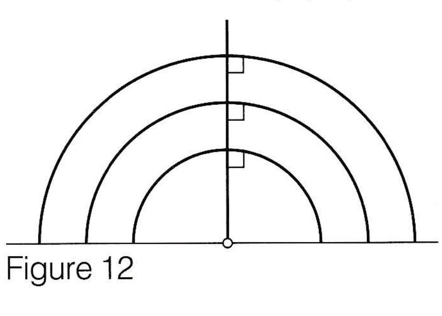


Figure 12

3. One of two perpendicular p-lines can always be represented as a ray (by applying an appropriate p-isometry, if necessary). Then the other one, clearly, will be a semicircle centered at the endpoint of this ray. So all p-lines perpendicular to a given p-line (problem (a)) can be represented as concentric semicircles (fig. 12). In problem (b), if one of the given p-lines is taken to be a ray, then the other one is semicircle (fig. 13), and the required p-line perpendicular to both the given ones is uniquely constructible as the semicircle whose center O is the endpoint of the ray and whose radius OT is the tangent to the given semicircle from O.

4. The statement about bisectors can be proved in exactly the same way as the similar Euclidean theorem, because the proof does not involve the parallel postulate. But the latter *is* involved in the Euclidean proof that perpendicular bisectors of two sides of a triangle intersect. And in fact, the p-circumcircle of a p-triangle, coinciding with the Euclidean circle through its vertices, may not fit entirely into the upper half-plane: in this case the p-triangle has no p-circumcircle.

5. Apply the Euclidean proof—parallels are not involved.

6. As in problem 2, it is consistent and more convenient here to use the circular model. The proof was given in "Inversion" (see exercise 8 there).

7. Suppose p-triangles ABC and AB_1C_1 do not coincide. To be definite, let B_1 lie on p-segment AB, and suppose first that p-lines BC and B_1C_1 have a common point D (possibly infinite)—see figure 14a. Then $\angle BB_1D + \angle DB_1A$ is equal to π (this is easy to

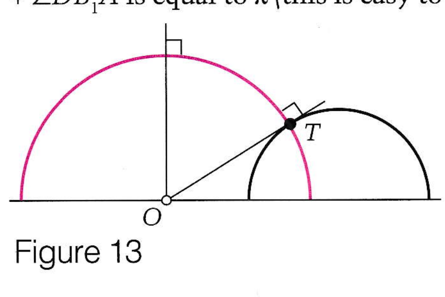


Figure 13

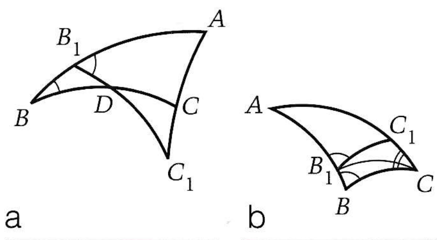


Figure 14

see if we draw the Euclidean lines tangent to the two circles at B_1). But $\angle DB_1A$ is equal to $\angle B$, so $\angle B + \angle BB_1D = \pi$. This makes the sum of the angles of p-triangles BB_1D greater than π. But in non-Euclidean geometry this sum must be less than π. If BC and B_1C_1 do not intersect (fig. 14b), the sum of the angles of p-quadrilateral BB_1C_1C is equal to 2π, whereas it must be less than 2π (because the quadrilateral is composed of two p-triangles BCB_1 and CB_1C_1).

In the general case, for any two p-triangles ABC and $A_1B_1C_1$ having equal corresponding angles, we can always find a p-isometry that fits angle BAC onto angle $B_1A_1C_1$. Then, by the above argument, vertices B and C must hit B_1 and C_1, so the p-triangles are p-congruent, thus proving the AAA test.

8. Suppose the p-isometry taking $L(a, b)$ into ray $L(0, \infty)$ maps $L(a, b_1)$ and $L(a, b_2)$ into $L(0, c_1)$ and $L(0, c_2)$, respectively (fig. 15). If the numbers c_1 and c_2 are of the same sign, the p-shift D_k, $k = c_1/c_2$, leaves $L(0, \infty)$ in its place and maps $L(0, c_1)$ onto $L(0, c_2)$. If c_1 and c_2 are of different signs, the p-reflection in $L(0, \infty)$ must be performed before the p-shift.

9. P-reflections are inversions, and as was explained in the solution to problem 1, inversions preserve the p-distance ρ.

To prove the equality $\rho(A, C) = \rho(A, B) + \rho(B, C)$, we can use the following formula, which was proven incidentally during the solution of problem 1 above: $\rho(A, B) = \ln(\tan\alpha/\tan\beta)$. For the case when the p-line is a semicircle, this formula quickly leads to the desired result. If the p-line is a ray, the equality directly follows from the definition of ρ.

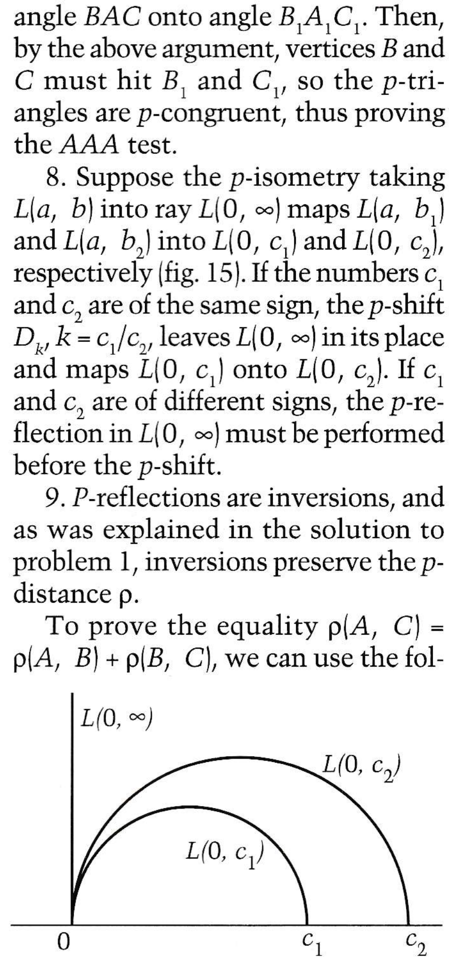


Figure 15

10. Again, as in problems 3 and 8, we can represent p-line L_0 as a ray $L(0, \infty)$. Then all p-lines perpendicular to L_0 are semicircles centered at 0, so the statement of the problem follows from figures 16a for intersecting p-lines, 16b for parallel p-lines, and 16c for superparallel p-lines.

Squeaky doors

1. Let's split the frictional forces acting on the car's front wheels into two components: F_1, which are in the wheel's plane, and F_2, which are perpendicular to the wheel's plane (see figure 17). The forces F_1 cause the wheels to rotate, and the forces F_2 turn the car.

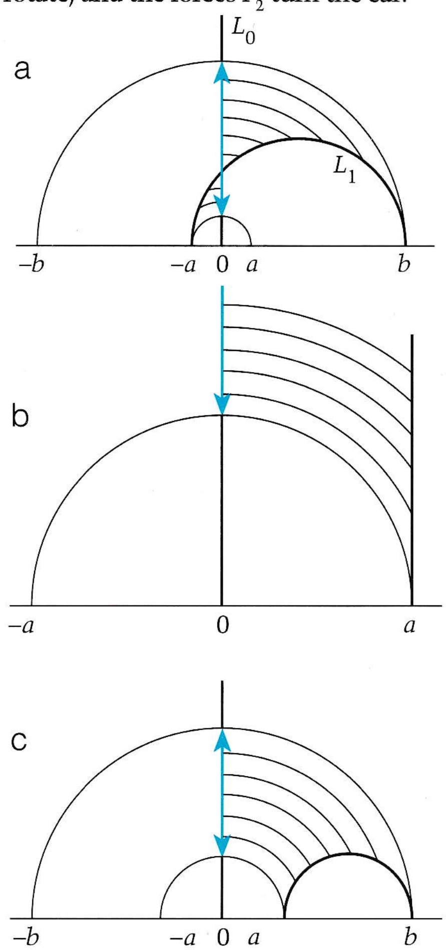


Figure 16

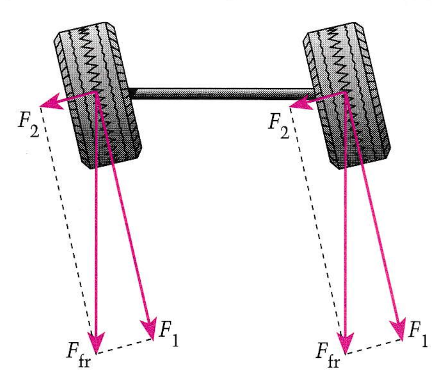


Figure 17

2. If the stick's center of gravity is not halfway between the fingers, then the pressure of the stick is different for the two fingers. The frictional forces of the fingers acting on the stick are different as well. The stick is shifted toward the side where the friction is less.

3. As long as the block is not sliding along the plane, the frictional force is equal to the projection of the block's weight onto the inclined plane $F_{fr} = W \sin\alpha$ (fig. 18a). The block begins to slide when the frictional force reaches the maximum value of static friction $F_{fr} = kN = kW\cos\alpha$. The condition $kW\cos\alpha = W\sin\alpha$ is satisfied. Therefore, the block begins to slide when the slope of the plane is $\alpha = \arctan k$. After that the frictional force will be equal to $F_{fr} = kN = kW\cos\alpha$ (fig. 18b).

The angle $\alpha = \arctan k$ at which the block begins to slide is called the "angle of repose." It also has another geometric meaning: if a force that makes an angle less than the angle of

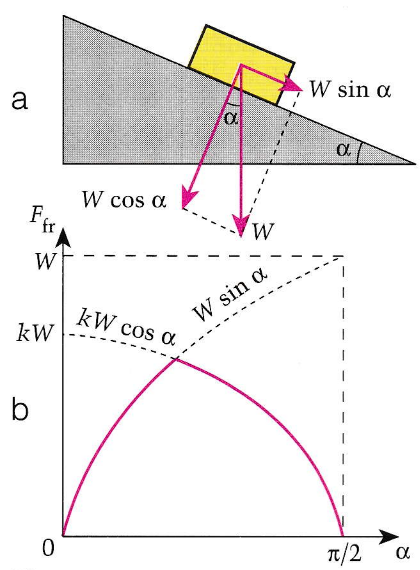


Figure 18

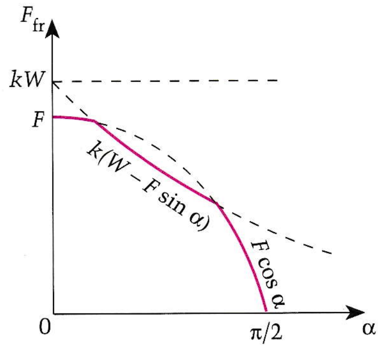


Figure 19

repose with the vertical is applied to a block on a horizontal plane, it's impossible to move the block no matter how great the force is. Here's how we can prove it. Let's place an observer on an inclined plane with a block resting on it, and let's increase the slope of the plane. In the observer's coordinate system, a force at an angle α to the plane acts on the body lying on the plane, which the observer takes to be horizontal. If $\tan \alpha < k$ (that is, $\alpha < \arctan k$), then the block does not slide along the plane no matter how great a force is applied.

4. See figure 19.

5. The cable acts on the rods with forces F_1 and F_2 that depend on the pipe's weight. These forces are directed along the rods (fig. 20). Therefore, if the angle between the rods and the perpendicular to the pipe's surface is less than $\alpha = \arctan k$ (k being the coefficient of friction of the rods with the pipe), then the rods will not slide along the pipe no matter how great the forces F_1 and F_2 are and no matter how heavy the pipe is (see the solution to problem 3). This is sometimes referred to as "wedging."

You could use the same principle

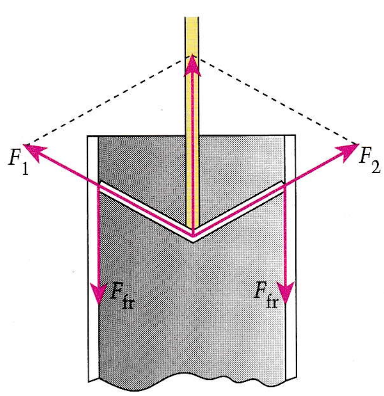


Figure 20

to make a simple and handy clasp for holding maps or blueprints. You would need to make two metallic yokes (see figure 21) and nail them to the wall. Put a small ball in each of them. A blueprint inserted into the slot will be pressed by the ball against the board. If the angle α formed by the inclined face of the yoke relative to the vertical is such that $\tan 2\alpha < k$ (where k is the coefficient of friction between the ball and the blueprint), then the blueprint will be held by the force of friction against the ball no matter how heavy the blueprint and no matter how small the friction of the blueprint with the wall. To take the blueprint out of the clasp, just push the ball up slightly with the tip of a pencil or pen.

6. The cube begins to slide when the resultant of force F and the component of the cube's weight that is parallel to the inclined plane $W \sin \alpha$ is equal to the maximum static frictional force:

$$kW \cos \alpha = \sqrt{W^2 \sin^2 \alpha + F^2}.$$

From this we get

$$F = \sqrt{k^2 W^2 \cos^2 \alpha - W^2 \sin^2 \alpha}$$
$$= \sqrt{3} W \cdot \sin \alpha.$$

7. 600 N.

8. The explanation for a door's squeaks is analogous to that for the sound emitted by a violin string.

Inversion

(See the September/October issue of *Quantum*)

1.(a) The result follows from the similarity of triangles OAX and OAX'.

(b) The triangles XOA and AOX' in figure 3 in the article are isosceles and have a common angle at O, so they are similar, which implies that $OX/OA = OA/OX'$.

2. If the given inversions take point X successively into X_1 and X_2, then $OX_2 = R_2^2/OX_1 = (R_2^2/R_1^2)OX$.

3. The first statement is a combination of two well-known theorems:

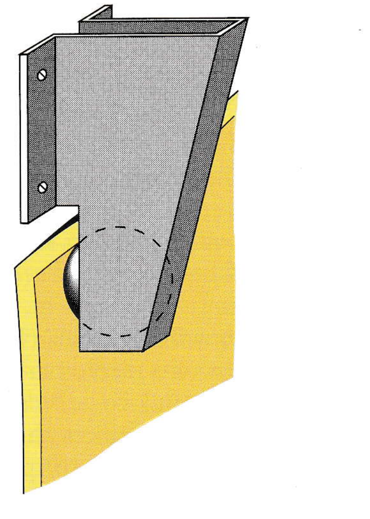


Figure 21

one on intersecting chords, the other on extended chords. It follows from the similarity of triangles PAA' and $PB'B$, where A' and B' are the points of intersection of another line through P with a given circle ω. If the line PAB is drawn through the center O of ω, then $PA \cdot PB$ equals $(OP + r)(OP - r) = OP^2 - r^2$ for P outside the circle and $r^2 - OP^2$ for P inside the circle (r is the radius).

Now imagine that point P in the situation above is the center of some inversion I. Consider another inversion I_1 with the same center and the radius $R = \sqrt{PA \cdot PB} = \sqrt{|OP^2 - r^2|}$. Then, if P lies outside ω, $B = I_1(A)$ for any line through P, because B is on the ray PA and $PA \cdot PB = R^2$. So in this case $I_1(\omega) = \omega$, and $I(\omega) = I(I_1(\omega))$. But by exercise 2, two successive inversions with the same center result in one dilation, so the inverse $I(\omega)$ of ω is a circle coinciding with a dilation of ω. If P lies inside ω, the reasoning is the same except that in this case $I(\omega) = I(I_1\omega'))$, where ω' is the circle symmetric to ω about point P: the point B' symmetric to B about P lies on ray PA, satisfies $PB' \cdot PA = PB \cdot PA = R^2$, and so is the I_1-inverse of A. Thus, $I(\omega)$ is the dilation of ω' relative to center P; and so the dilation of ω by some *negative* factor.

4. After the suggested inversion we get figure 22, in which the radius r' of ω', the inverse of ω, is evidently equal to $BC/4 = a/4$, and the power of A with respect to ω' is $AT^2 = (5a/4)^2 =$

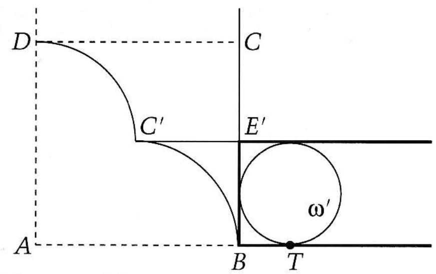


Figure 22

$(25/16)a^2$. So the radius of ω equals $r'(AB^2/AT^2) = (4/25)a$.

5. (a) Circle ω through A is orthogonal to ω_1 if and only if it passes through the inverse A' of A. So if $A' \neq A$ (A does not lie on ω_1), the required circle is the one through A, A', and B. If A lies on ω_1 and B does not, we draw ω through A, B, and B' (the inverse of B). Finally, if both points A and B lie on ω_1, the center of ω can be found as the intersection of the tangents to ω_1 at A and B. Note that the solution is always unique.

(b) Circle ω must pass through the inverses of A in ω_1 and ω_2. Special cases are left to the reader.

(c) The construction for the general case is shown in figure 23. We draw an arbitrary circle through A and B intersecting ω_1 in M and N and extend AB and MN to intersect at O. Then O is the center of the desired circle ω, and OT—the tangent to ω_1—is its radius ($OA \cdot OB = OM \cdot ON = OT^2$).

6. Let B', C', D' be the inverses of B, C, D in a circle with center A. Then $B'D' \leq B'C' + C'D'$, the exact equality being true if and only if C' lies on segment $B'D'$. What remains is to express the side lengths of triangle $B'C'D'$ in terms of the lengths of the sides and diagonals of $ABCD$.

7. Reflection in the equator is a particular case of inversion on the sphere, so the reasoning in the paragraph preceding the exercise is applicable. The reader is invited to work out a direct proof using similar triangles and suchlike.

8. Verifying the equality $d(A, B) = d(B, A)$ is straightforward if we take into account that when the labels A and B are swapped, so are A_0 and B_0. By the definition of $d(A, B)$, the inequality in question reduces to $R(A, B) < R(A, C) \cdot R(C, B)$. We may think of point C as the center of circle α—by exercise 5c, there is always a p-line that reflects C into the above center, and p-reflection preserves p-distances. For this particular central location of C, in the notation of figure 24, $R(A, C)$ reduces to a single ratio AM/AN, since $CM = CN$. It's not difficult to prove that $AM > AA_0$ and $AN < AB_0$. These inequalities imply that $R(A, C) > AA_0/AB_0$. Similarly, $R(C, B) > BB_0/BA_0$. The product of the right sides of the last two inequalities is just $R(A, B)$.

9. As in the preceding exercise, it will suffice to consider a p-triangle ABC with a vertex at the center of α. It's not hard to see, in figure 24, that in this case the sum of the angles of the p-triangle ABC is less than that of the conventional triangle ABC.

10. Let p-triangles ABC and $A_1B_1C_1$ have equal angles at the vertices A and A_1, and equal corresponding sides issuing from these vertices ($d(A, B) = d(A_1, B_1)$, $d(A, C) = (A_1, C_1)$). We must present a combination of p-reflections that map ABC onto $A_1B_1C_1$. Suppose first that $A = A_1 = O$, where O is the center of α. Then sides AB and A_1B_1, as well as AC and A_1C_1, are normal straight segments of equal (Euclidean) length, all issuing from O; so triangles $ABC = OBC$ and $A_1B_1C_1 = OB_1C_1$ are congruent by the Euclidean side–angle–side test. Therefore, they can be put in coincidence by at most two reflections in lines through O (which, in this case, are p-reflections as well): the reflection in the bisector of angle BOB_1, and then, if needed, in line OB_1. In the general case, by exercise 5c, we can find p-lines l and l_1 p-reflecting points A and A_1, respectively, into O (and triangles ABC and $A_1B_1C_1$ into $OB'C'$ and $OB''C''$). Now p-reflect triangle ABC in l to get $OB'C'$, then map the latter onto $OB''C''$ by means of one or two p-reflections (as above), and p-reflect $OB''C''$ in l_1. The result is triangle $A_1B_1C_1$, obtained in four or fewer p-reflections. In fact, three p-reflections will always suffice: a little thought will show that the reflection from $OB'C'$ to $A_1B_1C_1$ can be done directly (rather than via $OB''C''$).

Space physics

As the rocket travels downrange, it goes over the Earth's horizon. To demonstrate this, construct a semicircle with a 12-centimeter radius and label the center O. Construct a line at 90° from the base line, starting at point O, that bisects the arc at point L. Construct second line 15-centimeters long from point O at an angle of 30° from line OL and label its other endpoint A. Line OA should intersect the semicircle and extend beyond it (see figure 25). Measure 8 cm from point O along OA, mark a point, and label it X. The distance from point X to point L should be about 6.5 cm. Using that distance as the radius and point X as the center, start at the point L and construct an arc

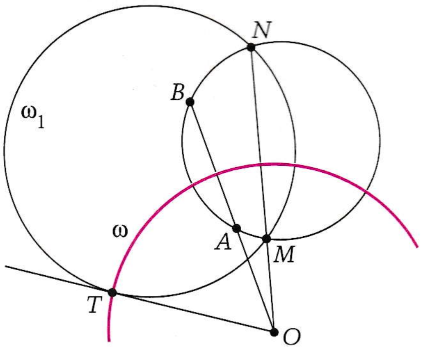


Figure 23

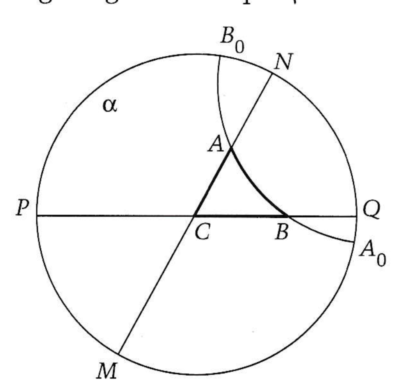


Figure 24

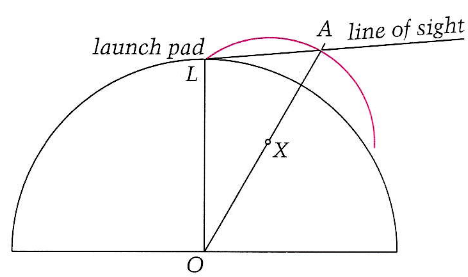


Figure 25

that intersects the semicircle a second time. That arc will represent the path of the rocket after liftoff. The rocket's apogee would be at the point where OA intersects the second semicircle. If you are standing near the liftoff point L, your line of sight extended to the apogee falls below the path of the rocket. Figure 25 is not drawn to scale and does not represent the actual path of the rocket, but it serves to demonstrate the illusion created by the rocket's flight.

Corrections

Vol. 2, no. 5:

p. 25, col. 2, ll. 33-34: *for* $(n^2 + 1)(n - 1)$ *read* $(n^2 + 1)/(n - 1)$; *for* $(n^5 + 3)(n^2 + 1)$ *read* $(n^5 + 3)/(n^2 + 1)$

p. 56, col. 2, l. 20 (counting the displayed equation as one line): *for* circle O_2 *read* circle O_1

Vol. 2, no. 6:

p. 14, col. 1, l. 30: *for* counted by S_8 *read* counted by S_9

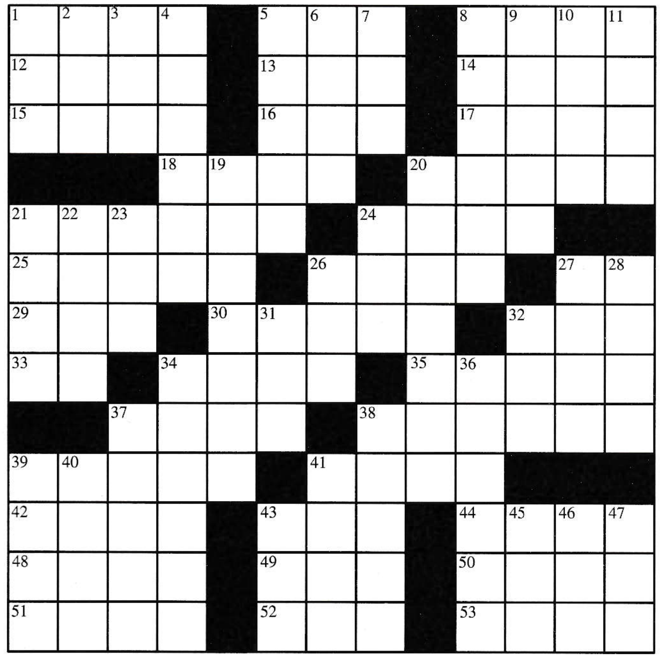
Vol. 3, no. 1:

p. 39, col. 1, ¶2: in the second sentence Martin Gelfand was inexplicably renamed Andrew; he remains Martin.

Index of Advertisers

Criss✕cross science

by David R. Martin



Across

1 Fund. particle
5 Remainder after deductions
8 Producers of eccentric rotation
12 Nat. space org.
13 Southern constellation
14 Leave out
15 One who leaves
16 Period of time
17 Food grain
18 Large wading bird
20 Submarine locator
21 Flat
24 Atomic number (abbr.)
25 Hole borer
26 Region
27 Computer smarts (abbr.)
29 A product (suff.)
30 Greek letter
32 Almost a Ph.D. (abbr.)
33 Tenth element
34 Transmitted
35 Hypothetical representation
37 Break _____ (2 wds.)
38 Sell door to door
39 Noble gas
41 Indian dress
42 Harvest
43 Energy unit (abbr.)
44 Chooses
48 After and, or or
49 Energy unit (abbr.)
50 Not new
51 Snow vehicle
52 ____ Lanka
53 ____ Valley (city in CA)

Down

1 Unit of length (abbr.)
2 *The ___ of Physics*
3 Carbohydrate (suff.)
4 Of the sea
5 Lowest point
6 Geologic time periods
7 River in Scotland
8 ____ Borealis
9 ____ acid
10 Complex silicate
11 After gang, or trick
19 Superconductivity theorist
20 It uses water vapor
21 Hurt feeling
22 Old stringed instrument
23 Ripen
24 Creative work
26 Alternate (abbr.)
27 Adam's son
28 Disengaged
31 Designer (abbr.)
32 Sum
34 Inclined
36 Disgusting
37 Semiprecious stone
38 ____ exclusion principle
39 Independent variables (abbr.)
40 Number system
41 Large ball of plasma
43 Edge conditions (abbr.)
45 Unit of pressure (abbr.)
46 Wave guide mode (abbr.)
47 Star Wars program

SOLUTION IN THE NEXT ISSUE

Slinking around

Put some spring in your walk—and vice versa

by Diar Chokin

MOST OF THE TOYS PRESented so far in our Toy Store have been mathematical. But physical toys are just as interesting. And the most interesting are those that look plain and simple but, when put into action, behave in a strange and unexpected way. The explanation for such behavior sometimes requires a full-fledged minitheory, often far from trivial.

The Slinky™ is just such a toy—a loosely coiled spring with 50–100 coils of diameter 2–4 inches. I'm sure most if not all of you have seen and played with a Slinky. Its very low spring constant makes it possible to perform interesting and instructive experiments that wouldn't be possible with an ordinary spring.

The most curious thing is that a Slinky can "walk" down steps or an incline. All you have to do is to put the Slinky upright at the edge of a step and gently push its upper end toward the lower step. The Slinky will immediately start walking down, by gradually "flowing" over the upper step down to the lower one. When the top of the spring flows over the edge, it describes an arc in the air, falls down onto the next step, and the motion continues (fig. 1).

Clearly, the main explanation of this effect is that, because of its low elasticity, a Slinky has no time to damp the horizontal component of the velocity of its upper end, which enables it to tumble over onto the next step. Such a walking spring can be likened to a self-oscillating system drawing its kinetic energy from the potential. I'll evaluate some dynamical parameters of a Slinky.

First, let's consider the portion of the spring's mass that participates in the motion. To this end I'll start with the following problem.

If a Slinky is suspended by its upper end (fig. 2), how is its mass distributed along its length?

Consider the nth coil of the spring (counting from the bottom). If its height is Δx_n, and its spring constant and mass are Δk and Δm, respectively (they are the same for all coils), then, by Hooke's law,

$$\Delta k \cdot \Delta x_n = n\Delta m \cdot g.$$

So the mean linear density λ_n of the nth coil equals

$$\lambda_n = \frac{\Delta m}{\Delta x_n} = \frac{\Delta k}{ng} = \frac{Nk}{ng},$$

where N is the total number of coils and k is the spring constant of the entire Slinky (think why $\Delta k = Nk$). Let's find the height x_n of the nth coil above the lower end of the spring. Since $\Delta x_n = ng\Delta m/\Delta k$, the height is equal to

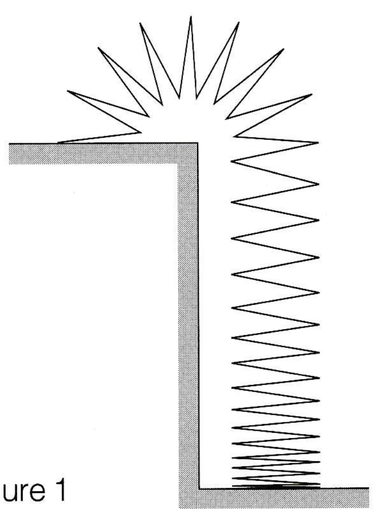


Figure 1

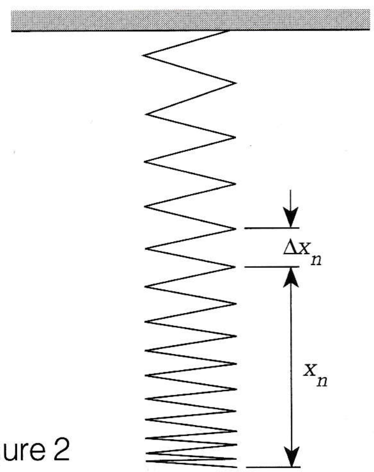


Figure 2

$$\begin{aligned} x_n &= \Delta x_1 + \Delta x_2 + \cdots + \Delta x_{n-1} \\ &= \frac{\Delta m \cdot g}{\Delta k}(1 + 2 + \cdots + n - 1) \\ &= \frac{(M/N)g}{kN}\frac{n(n-1)}{2} \\ &\cong \frac{Mg}{2k}\frac{n^2}{N^2}, \end{aligned}$$

where M is the mass of the entire spring. (It's interesting that the full length of the suspended spring, $L = Mg/(2k)$, turns out to be half the length of a similar massless spring with a mass M suspended at its end.)

So we see that the section of a suspended Slinky of height $x = x_n$ above its bottom end will consist of approximately $n = N\sqrt{2kx/Mg}$ coils. Therefore, its mass $m(x)$ equals

$$m(x) = n\Delta m = n\frac{M}{N} = \sqrt{\frac{2Mkx}{g}}.$$

(Now the linear density $\lambda(x)$ at height x can be found by differentiating $m(x)$:

$$\lambda(x) = m'(x) = \sqrt{\frac{Mk}{2gx}}.$$

Or, if you haven't studied the derivative yet, you can obtain this formula simply by comparing the above equations for $\lambda_n = \lambda(x_n)$ and x_n.)

To estimate the mass of the spring that participates in the motion, assume that the portion in question is equivalent to a freely suspended spring whose length equals the height h of one step. Using the expression for $m(x)$, we find that the unknown mass is equal to

$$m = m(h) = \sqrt{\frac{2Mkh}{g}},$$

and the ratio of this mass to the total mass is

$$\frac{m}{M} = \sqrt{\frac{2kh}{Mg}} = \sqrt{\frac{h}{L}}.$$

The smaller the portion of the Slinky's mass that participates in the motion, the more stable its walk. So to improve the "walking ability" of the toy, our equation indicates that we should diminish the spring constant, increase the spring's mass, and start the spring from steps that aren't too high. It would be even better to do it, say, on Jupiter—to increase the acceleration of gravity g. For a real spring, taking h = 10 cm and L = 1 m, we get $m/M \cong 0.3$.

Another question I tried to answer was this: what is the time T it takes for a Slinky to take one step? I don't want to tire the reader with calculations any longer (they may not be very appropriate in the Toy Store)—I'll just announce my result,

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

and invite you to prove (or, maybe, disprove) it yourself. It's interesting that this time doesn't depend on the height of the step and is of the same order as the period of the spring's free oscillations, or as the free-fall time of a body dropped from a height L. Setting, for instance, L = 1 m, we get $T \cong$ 0.5 s, which is supported by experiment.

Another no less important experiment with a Slinky is modeling longitudinal mechanical waves. Stretch the spring and compress one of its coils. You'll initiate a propagating wave of alternating stretches and compressions, which will be reflected at the ends of the spring. One of the properties that can be illustrated in this experiment is that a decrease in the medium's density leads to an increase in the speed of the longitudinal waves. To verify this, stretch the Slinky slightly—the wave speed will increase noticeably.

When he wrote this article, **Diar Chokin** *was a senior in the Republic High School for Physics and Mathematics in the city of Alma-Ata (Kazakhstan). He is now a student in the physics department at Moscow State University.*

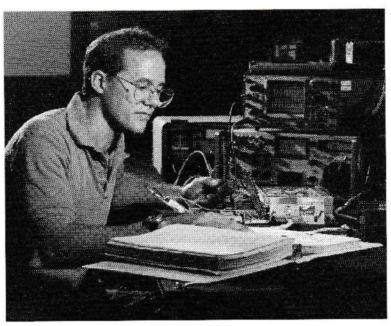
Art by Sergey Ivanov

When we say, "Let's see if it will fly," we mean it!

Embry-Riddle engineering students use CADD resources to aid in the design of the their two senior team projects. The curriculum includes a mix of theory and practical applications that fully prepare students for aerospace and aviation careers.

Over 67 international university teams entered the 7th Annual Aero Design Competition, co-hosted by the Embry-Riddle Society of Automotive Engineering Chapter. The objective of the competition is to lift as much weight as possible with a R/C aircraft designed and built by students.

Aerospace engineering students at Embry-Riddle don't have to wait until tomorrow to apply what they're learning today. The curriculum builds to two senior design projects that test your knowledge of aerodynamics, controls, materials, propulsion and structure as your team designs the internal and external details of an airplane or spacecraft. Both your imagination and abilities are put to the test, just as they will be in your professional life. We believe that's what education is all about.



Electrical Engineering students at Embry-Riddle receive a complete, broad-based education with full laboratory resources and equipment to complement academic study.

You may know about us, ***but did you know***:

- We offer over 20 different accredited aviation/aerospace majors at the Bachelor and Master's levels including aerospace engineering, electrical engineering, engineering physics, avionics engineering technology, aircraft engineering technology and aviation computer science.
- Our aerospace engineering enrollment at our Daytona Beach, Florida and Prescott, Arizona campuses is the largest in the United States. We are listed as one of the 18 best engineering schools in the nation by *U.S. News and World Report's 1992 College Guide* called "America's Best Colleges."
- Your university life is enriched through many social and professional clubs and organizations, like our SAE Chapter, that provide the opportunity for a variety of living and learning experiences, related to your major, outside the classroom.

Want to test your career design? Write to us today.

EMBRY-RIDDLE
AERONAUTICAL UNIVERSITY
Look at us now.

Mail to: Embry-Riddle Aeronautical University
Director of University Admissions
600 South Clyde Morris Boulevard
Daytona Beach, FL 32114-3900
Or Call 1-800-222-ERAU

Circle No. 1 on Reader Service Card

Please send me more information about Embry-Riddle.

Engineering Related Programs:
☐ Undergraduate
☐ Graduate

Other Aviation/Aerospace Majors:
☐ Undergraduate
☐ Graduate

Name ______________________

Address ______________________

City ______________ State ______ Zip ______________

Phone ( ) ______________________

MZ