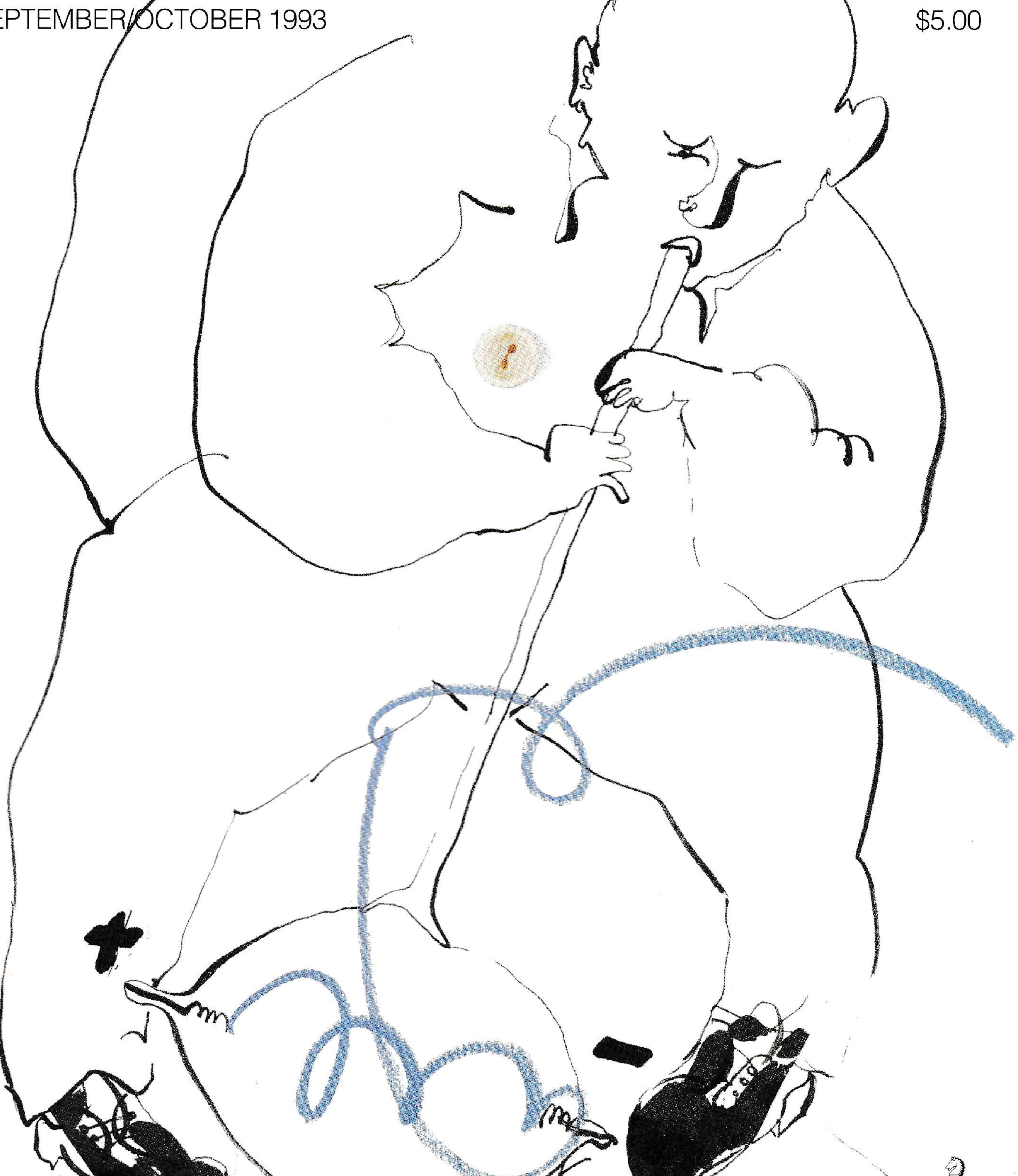
QUANTUM

SEPTEMBER/OCTOBER 1993

$5.00



SPRINGER INTERNATIONAL

GALLERY Q

National Gallery of Art, Washington (Samuel H. Kress Collection) © 1993 NGA

Saint Francis Receiving the Stigmata (ca. 1445) by Domenico Veneziano

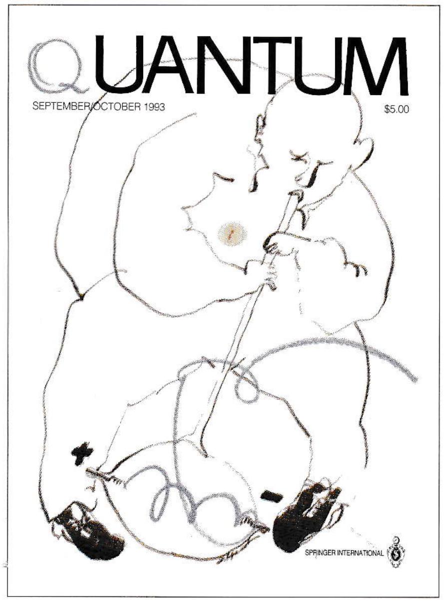
THE WORD "STIGMA," WHICH LITERALLY MEANS A mark, almost always implies disgrace or infamy. The Greek plural form, however, has just the opposite connotation. In the Christian tradition, "stigmata" are a sign of special grace and holiness. The person receiving them—in this case, Saint Francis of Assisi—bears mysterious wounds in the five places where the crucified Christ was pierced by nails and by the spear of the Roman centurion.

One curious aspect of this painting is the object in the upper right-hand corner from which the five lines are emerging. In Medieval and Renaissance art, the source of the stigmata is usually the dying Christ or a crucifix. In the first case, the saint stands at the foot of the cross in the actual historical setting; in the second, a small artistic representation of the Crucifixion hovers above the saint in his or her own time. In Domenico Veneziano's version, St. Francis and his companion are in a timeless landscape, and the Crucifixion has been transformed almost beyond recognition. There is something awful about the form suspended in the air, and for a modern viewer, the level of abstraction is wonderfully unexpected. We are used to seeing, for instance, the Holy Spirit represented as a dove or (in representations of the Pentecost) as tongues of flame. But this painting hints at the spiritual realm in a way that most art of the period does not.

Once you've read the article in this issue by the great Russian mathematician Andrey Kolmogorov, our reason for choosing this work of art for Gallery Q will become so obvious that we prefer to leave it unsaid.

QUANTUM

SEPTEMBER/OCTOBER 1993 VOLUME 4, NUMBER 1



Cover art by Dmitry Krymov

The physicist on our cover looks a little frustrated. That's because he's working with a very tricky substance: plasma. The allure of plasma physics has been like a will-o'-the-wisp, leading deeper into the woods where the secret of controlled thermonuclear fusion resides. The three requirements for a thermonuclear reactor—high particle density, high plasma temperature, and long confinement time—are proving more difficult to achieve than was originally thought. But the research continues, because many see fusion as the energy source of the future. (This explains the big blue *Q* that has muscled its way into our logo.)

If the only plasma you know about is the stuff flowing through your veins, turn to the cover story on page 4 for a "conceptual transfusion."

FEATURES

DEPARTMENTS

The view from the masthead

You have to get into it to read the fine print!

WHEN I SAY "MASTHEAD," do you think of a crow's nest? I wouldn't blame you if you do. Only editors—and perhaps a peculiar breed of masthead aficionados—actually read the dense block of information that usually appears on page 3—right across the "gutter" from my little soapbox! But I'd like to draw your attention to a few changes that have been made there.

The first is a statement of the mission of the National Science Teachers Association. It will appear in all NSTA journals. Next to it is a seal commemorating the fiftieth anniversary of NSTA, which will appear in the masthead throughout this jubilee year. *Quantum* is one of the many ways in which NSTA strives to fulfill its stated goal of stimulating and improving science teaching and learning.

If you have sharp eyes, you may have noticed that I skipped something. The line under "QUANTUM" now reads: The Magazine of Math and Science. Although we've dropped the word "student" from the subtitle, we haven't changed our goal of producing a magazine that students will enjoy. In fact, we hope to include in each issue more material that will engage students with less math and physics background. In making this change, we're simply acknowledging that *Quantum* appeals to a broad range of readers, from high school students and teachers to researchers and graduate students. It seems there is no upper limit to *Quantum*'s appeal.

Finally, a new address is hiding somewhere in all that fine print. NSTA has purchased a new headquarters building in a Washington suburb. Operations that are now scattered in several locations will soon be brought together under one roof—specifically, at 1840 Wilson Boulevard, Arlington, VA 22201. Use this address for any editorial correspondence. (Subscription inquiries should be directed to our copublisher, Springer-Verlag New York, at the address and phone numbers given in the masthead.)

Thanks, E'beth!

One change to the masthead won't appear until the next issue, and it is a sad one. Because of budgetary pressure, the position of art director at *Quantum* has been eliminated. Some of the tasks performed by Elisabeth Tobia will be transferred to the NSTA Publications Production Department. But E'beth (as she is known to her coworkers) is irreplaceable. She was one of a handful of people who brought *Quantum* into existence and who worked tirelessly to make it better and help it grow. Because the staff is small, each member wears several hats. E'beth lent her expertise to promoting *Quantum* and was our principal liaison with our production and marketing colleagues at Springer.

E'beth has taken another position at NSTA—project manager of student competition sponsored by NYNEX and administered by NSTA. We wish her well, and congratulate NYNEX on their good fortune. We will miss her skillful hand at layout, her good humor in tight situations, and her fierce devotion to *Quantum*.

Did I say that?

Sometimes the words just don't come out the way you meant them. In my last Publisher's Page, I explored the "circuitous route" I took to a kind of science I found "relevant." On the way, I described the view from a high window onto the elevated plain of advanced physics, and wrote that I "figuratively lowered myself from that window . . . I became a teacher." Some readers took the time to point out how bizarre that sounds. Even with the words restored to the ellipsis—"and pursued those things for which I was fully capable"—the text as printed doesn't quite work.

Let me try again. I feel strongly that teaching is as important as research (look at the name of my organization!). Each requires different abilities, though. I had spent time in the world of Einstein, Feynman, and Co., and understood the profound theoretical work being done there—but only with great difficulty. Realizing that I could not succeed at their level, I still wanted to excel, not just do time. So I tried to become the best in another field: teaching.

My apologies to teachers everywhere who may have been offended by what I wrote. I've met bright people in research labs, and I've known geniuses who work in classrooms. And none of them was pursuing watered-down science. That was my larger point, which I hope came through loud and clear.

—Bill G. Aldridge

QUANTUM

THE MAGAZINE OF MATH AND SCIENCE

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

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

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

Advertising:

Advertising Representatives: (Washington) Paul Kuntzler (202) 328-5800; (New York) Patrick Ferencz (212) 460-1575; and G. Probst, Springer-Verlag, Heidelberger Platz 3, D-14197 Berlin, 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.

Printed in U.S.A.

Subscription Information: North America

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

All Countries Outside North America

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

The fourth state of matter

Neither solid nor liquid nor gas

by Alexander Kingsep

LET ME ASK YOU A STRAIGHT-forward question: what state of matter is most prevalent in the universe? Solid, you say? Not so. Liquid? Guess again. Gas? Sorry. You may be surprised by the answer. The most typical state of matter throughout the universe is *plasma*, and the bulk of creation exists in just this phase. That might strike you, a solid–liquid being, as a bit unlikely.

Plasma used to be defined as ionized gas. As a matter of fact, that's not quite right—or, more exactly, it's as correct as these definitions: "liquid is melted solid," or "gas is vaporized liquid." The properties of plasma (at least, of typical plasmas) are essentially different from those of gases, so physicists generally look upon plasma as the fourth state of matter, occupying the highest place on the temperature scale.

You've seen plasmatic objects virtually every day of your life—the Sun and stars. You've seen plasmatic light sources—neon lights. Everyone is familiar with the short-lived plasmatic phenomenon of lightning. And almost everyone is aware of shortwave radio broadcasting. In its older variants, this means of long-distance communication was based to a great extent on the so-called plasma mirror provided by the ionosphere, which reflects the electromagnetic waves broadcast by shortwave radio stations (fig. 1). It's not surprising that such a broad range of objects and phenomena have attracted the attention of the physics community. Many instruments and devices have been invented as a direct result of progress in plasma physics. But the main problem stimulating progress in this field is that of controlled nuclear fusion, which many consider the energy source of the future. If this problem were solved, fusion would give humanity a power source no less efficient than our current fission reactors but much less dangerous. It would be ecologically pure and almost inexhaustible. Most plasma physicists nowadays are involved in several advanced fusion programs.

In this article, though, we'll be looking at just a few aspects of plasma—its most fundamental properties.

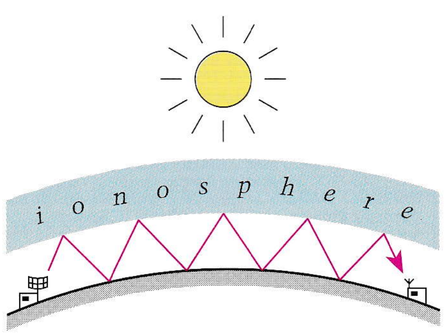


Figure 1
Two natural plasmatic objects and a schematic rendering of long-distance radio broadcasting.

Plasma as a continuous medium

At first glance, as noted above, plasma should be more gaseous than gas. Indeed, its place on the temperature scale is higher. Let's take hydrogen plasma as an example. To transform the neutral H atom into the pair $H^+ + e$, we need energies as high as $\varepsilon_{ion} = 13.6$ eV $\cong 2.18 \cdot 10^{-18}$ J. (The dissociation threshold is much lower, so we don't take it into account.) Such ionization can occur in many ways, but to keep a hydrogen cloud in such a state, the temperature must be at least $T_{min} \cong \varepsilon_{ion}/k$ (where k is the Boltzmann constant, $k = 1.38 \cdot 10^{-23}$ J/K = 1 eV/11,600 K). In actuality the temperature may be a little lower, but it's still of the order of several electron volts.[1] This temperature is close to that on the surface of stars (it's funny, but they call these plasmas "cold"!). Fusion plasmas are substantially hotter (temperatures as high as 10^8 K); the plasmas in neon lights are much cooler, so they're only partially ionized. This is an argument in favor of the notion that plasma is "more gaseous than gas."

All gases, even metal vapors, are always poor conductors. This situation changes drastically, however,

[1]Plasma physicists often give temperatures in energy units using the relationship $T = E/k$. An energy of 1eV is equivalent to a temperature of 11,600 K.

Art by Dmitry Krymov


QUANTUM
NOVEMBER/DECEMBER 1991
$5.00
The student magazine of math and science
Д. Крымов

when ionization occurs. Depending on the parameters (primarily temperature and particle density), plasma can be a very good conductor, no worse than such metals as copper and silver in the solid state. As a result, its macroscopic dynamics become rather complicated, since its mechanical properties are closely linked with its electrodynamic properties. This might be considered the main feature of plasma—that electrodynamics plays an extremely important role in any problem having to do with plasma and in any of its effects that you come across. Specifically, plasma flow produces an electromagnetic field, while the motion and equilibrium of the plasma are determined by the fields acting on it.

This property turns out to be very useful for holding plasma. The point is, very hot plasmas—say, fusion plasmas—can't be put in a container where it comes into contact with the walls. Such contact would provoke both very rapid energy losses and a very rapid recombination of the plasma particles (for example, $H^+ + e \to H$). So, if we want to heat plasma, we have to trap it some other way. A strong, properly designed magnetic field is just what we need. The simple principle underlying a magnetic trap (the so-called mirror trap, or Budker's trap) is shown in figure 2.

To understand the physical mechanism of magnetic confinement (in particular, the principle of a magnetic mirror), let's turn to figure 3. As was just mentioned, typical plasmas are good conductors. This means that an external magnetic field has difficulty penetrating into a space occupied by plasma. Researchers in the field of superconductors know this effect well—it's called the Meissner effect. Now, plasmas are very good conductors, but they're not ideal conductors, so such penetration isn't completely prohibited. It's just that it requires too much time compared to the typical time scales of plasma dynamics. If the magnetic field doesn't penetrate the plasma, it has to be kept away from it by a thin layer of current, as shown in figure 3. The surface charge density j—and the very fact that such a layer exists—are direct consequences of Maxwell's electrodynamics. And the Amperean force per unit area is jB and acts perpendicular to the plasma's surface, from outside in, as if some pressure were being applied to this surface. And that's why magnetic trapping occurs.

Another example of this sort, perhaps the simplest and most visual, is the so-called pinch effect (fig. 4a). Imagine a plasma cylinder with a current running along its length (it doesn't matter whether the current is concentrated on the plasma's surface or is distributed inside the cylinder). It's known that conductors with parallel currents in the same direction attract one other. This is one of the basic tenets of electrodynamics. In our case, it means that such a current-carrying plasma cylinder must draw tighter along its radius. On the other hand, ordinary "thermal" or "gas" pressure opposes this tightening. The resulting equilibrium determines the radius of the cylinder, the thermal pressure, and the magnetic field at any point in space. And that is what we call a "pinch." By the way, if this pinch is shaped into a torus to eliminate the effects of the ends and placed in the external magnetic field in a toroidal conducting chamber, it forms the so-called tokamak configuration, at present the most effective type of plasma trap.

The pinch is also useful in demonstrating another fundamental property of hot plasmas—one that gives plasma physicists fits. I'm talking about plasma instability. It's been responsible for sending many a promising idea or project to an early grave. A simple instance of this instability is shown in figure 4b. It's usually called a "neck" or "constriction instability." Let's assume that the radius of the plasma cylinder has become less than the average at some point. Such a perturbation doesn't disrupt the thermal pressure (assuming, of course, that the pinch is long enough). As for the magnetic field, it is drastically disturbed in some region of the plasma's surface, because the total current must be conserved, while its magnetic field (as for any straight conductor) decreases with distance: $B \propto jr^{-1}$. Since the attractive Amperean force increases

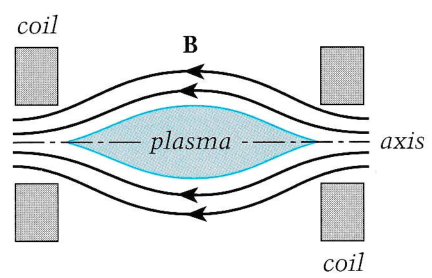


Figure 2
The mirror plasma trap.

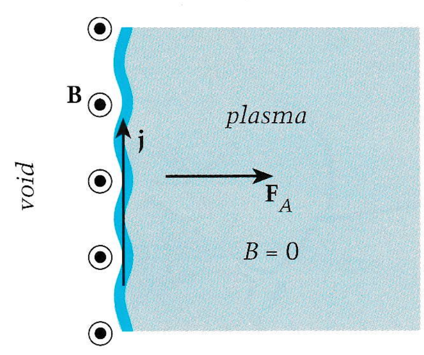


Figure 3
Magnetic confinement.

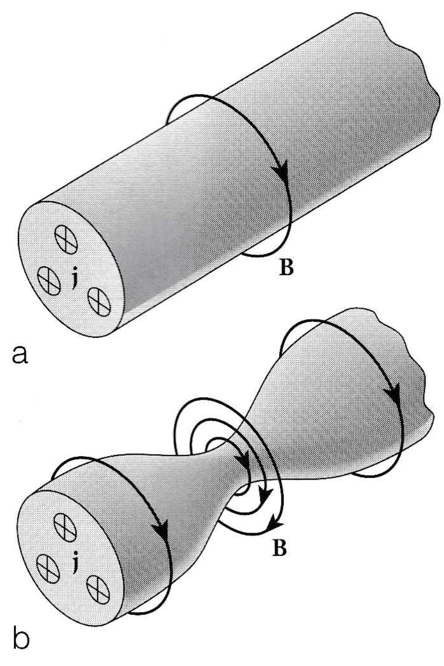


Figure 4
The pinch effect: (a) steady, (b) unstable.

while the repulsive gas pressure stays almost constant, our equilibrium is broken.

Instead of a neck, we could imagine a "bubble" with similar consequences. Such a bubble must expand, thereby destroying the equilibrium. For decades the physics community has struggled with plasma instabilities on route to controlled fusion. Some instabilities can be eliminated by properly arranging the experimental setup and adjusting the plasma parameters; some can be suppressed with external fields; some can be delayed during plasma experiments. Plasma instabilities essentially determine the physical scenario in outer space and must be studied to understand many astrophysical problems.

I'd like to point out one final consequence of the influence that plasma mechanics and electrodynamics have on one another. This is the phenomenon of "freezing in a magnetic field." As was mentioned above, rapid penetration of a magnetic field into a plasma is difficult, but it can still occur as the result of a rather slow evolution. After that, though, if some sort of rapid dynamics sets in, penetration of the magnetic field again becomes difficult. (You may have seen the popular physics demonstration in which a silver or copper coin falls between the poles of a powerful magnet—it drops very slowly compared to the free-fall velocity.) So the rapid mechanical motion of a plasma—such as the compression of a plasma "cloud"—causes the magnetic field to be "glued" to the plasma. You can see a very good example of this in figure 5, which explains the huge values for the magnetic field in neutron stars. Indeed, the neutron star (fig. 5b) results from the gravitational collapse of a star (fig. 5a), and this collapse occurs fast enough to keep the magnetic field of the star "frozen in." As a result of this collapse, the density of the magnetic field lines and, hence, the strength of the magnetic field are greatly increased.

It's interesting that this connection between the mechanical and the electromagnetic properties is typical of molten metals. So in this respect plasma looks like condensed matter, not a gas.

Plasma as an ensemble of particles

Let's try to compare the different states of matter on the microscopic level (see figure 6). The solid state is lowest on the temperature scale. All the atoms in this state are ordered in space, forming a lattice. Their equilibrium positions are fixed, and their motion is essentially restricted—only rather small thermal oscillations near these positions are allowed. In the liquid state, particles are given much more freedom to move—this is the so-called Brownian motion. However, this motion isn't completely free. Nearest neighbors still interact, and this interaction affects how all the particles move about. The next phase on the temperature scale, the gaseous state, corresponds to maximum freedom of particle motion (atoms or molecules, depending on the kind of gas). These particles are almost completely free to move about. The only restriction occurs when they collide, since the intermolecular forces decrease sharply with distance:

$$F_{ij} \propto r_{ij}^{-7}$$

(the Lennard–Jones approximation of the van der Waals force). The more rarefied the gas, the rarer the collisions between particles, and the less influence the particles have on one another.

What about plasma? As rule, hot plasma is even more rarefied than gases and its particles move much faster, since the characteristic thermal velocity

$$v_{T\alpha} = \left(\frac{kT}{m_\alpha}\right)^{1/2}, \quad \alpha = i, e,$$

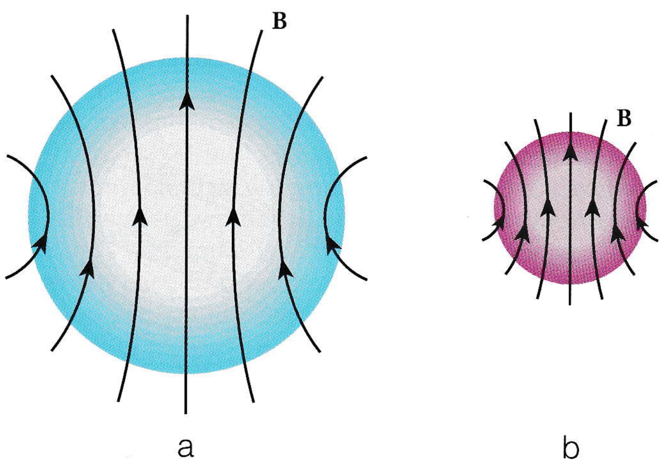


Figure 5
The "freezing-in" effect (a) in the initial state of a star; (b) in its final state (a neutron star).

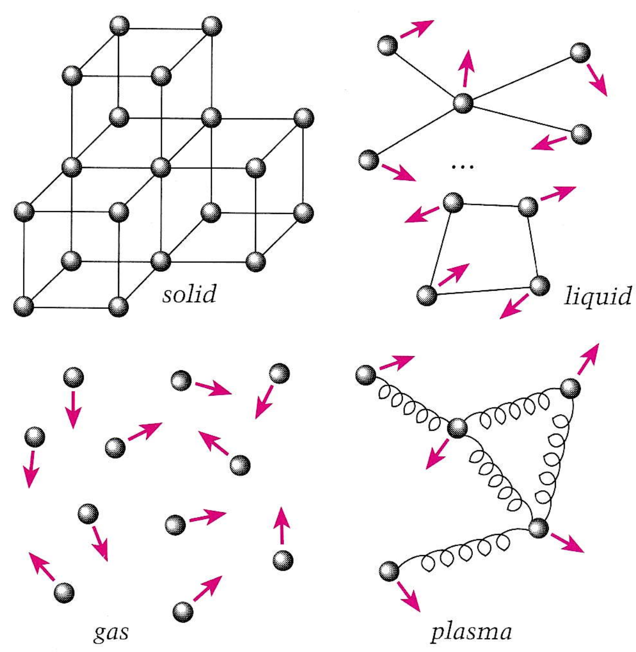


Figure 6
The four states of matter.

increases with temperature. In addition, the electron's mass is less than the least massive atoms by a factor of almost 2,000, so electrons travel faster than ions by a factor of $(2{,}000)^{1/2}$. Nevertheless, their motion doesn't become "freer"—on the contrary, long-distance interactions come into play here, radically changing the entire microdynamics. This occurs because the Coulomb forces between particles don't decrease very rapidly with distance:

$$F_{ij} \propto r_{ij}^{-2}.$$

This slow decrease is enough to completely change the mechanical interactions, forming great ensembles of particles despite the thermal motion. Though Coulomb collisions remain substantial at times, the *collective* behavior predominates in hot, rarefied plasmas just as in solids. So in this respect plasma is also closer to condensed matter than to a gas.

One of the most important properties of plasma is its so-called quasi-neutrality. Although it consists of charged particles, plasma as a whole has no charge. That's to be expected—since all the electrons and ions are produced by the ionization of neutral atoms, their net charge is zero. Not only that, the plasma's neutrality is maintained with great accuracy because it takes a very large voltage to separate the charges.

For example, let's look at a cube of plasma with a volume $V = 1$ cm^3 and a particle concentration that's rather typical for laboratory plasmas: $n = n_e = n_i = 10^{20}$ m^{-3}. If we want to disrupt the charge equilibrium in this small volume ever so slightly—say, by 1 percent—we need to bring a very strong external electric field to bear on this volume. Indeed, let's assume that there is some excess of electrons (ions) in the cube with a relative concentration $\delta n \cong 0.01n = 10^{18}$ m^{-3}. To estimate the resulting electric field, we can use Gauss's theorem:

$$\int E\, dS \cong ES = \frac{q}{\varepsilon_0} = \delta n \cdot V \cdot \frac{e}{\varepsilon_0}.$$

Then, taking the cube's surface $S \cong 6 \cdot 10^{-4}$ m^2, $V \cong 10^{-6}$ m^3, $e = 1.6 \cdot 10^{-19}$ C, and $\varepsilon_0 = 8.9 \cdot 10^{-12}$ C^2/N $\cdot$ m^2, we can readily obtain $E \cong 3 \cdot 10^7$ V/m.

In addition, we can modify the problem. Let's keep $\int n_e\, dV = \int n_i\, dV$, but violate local equality. If $n_e - n_i \cong 10^{-2} n_{e,i}$ near one side of the cube and $n_e - n_i \cong -10^{-2} n_{e,i}$ near the opposite side, then, as follows from the exact solution, E is again of the order of 10^7 V/m in the entire volume. In the first example, such a huge field induces a repulsive pressure that causes the plasma to blow apart, while in the second case the strong attraction equalizes the electron and ion densities inside the cube. To maintain the initial disruption of neutrality, we have to use an external electric field, again with a strength of 10^7 V/m. If our external field can't be this strong, then all the charge perturbations (as a whole) within this volume have to obey the inequality

$$\frac{\delta n}{n} \equiv \frac{|n_e - n_i|}{n_{e,i}} << 10^{-2}.$$

Thus, quasi-neutrality consists not only of net neutrality but also of the local relation

$$n_e = n_i = n,$$

or

$$n_e = Zn_i,\ Z \neq 1,$$

where Z is the ionic charge number. Nevertheless, charge oscillations are allowed (and do occur) near charge equilibrium. Imagine a plasma layer and let all the electrons be displaced a bit with respect to the ions (fig. 7). If x is the spatial displacement ($x << l$), then the charge per unit area in the narrow sheets on the boundary is equal to

$$\frac{dQ}{dS} = nex.$$

This picture reminds us of a parallel plate capacitor, and we can immediately calculate the electric field inside the volume of the plasma:

$$E = \frac{V}{l};\ V = \frac{Q}{C};\ C = \frac{\varepsilon_0 S}{l} \Rightarrow E = \frac{nex}{\varepsilon_0}.$$

Now we take into account that $m_e/m_i << 1$. This ratio is approximately equal to 1/1,840 for hydrogen, which allows us to neglect ion motion. Electron motion can be described by Newton's second law:

$$m\frac{d^2x}{dt^2} = -eE = -\frac{ne^2x}{\varepsilon_0},$$

where $m \equiv m_e$ and the negative sign of the electron's charge is taken into account. As a result, we obtain the equation for the harmonic oscillator:

$$\frac{d^2x}{dt^2} = -\omega^2 x,\quad \omega \equiv \omega_p = \left(\frac{ne^2}{\varepsilon_0 m}\right)^{1/2}.$$

This is the most typical collective motion for a plasma, called Langmuir oscillations or simply plasma oscillations.

Any spontaneous or induced disruption of quasi-neutrality leads to such oscillations. On the other hand, many interesting and important properties of plasma are based on this effect. In particular, it offers the possibility of nonlinear plasma diagnostics. Two Langmuir waves can merge into a single transverse electromagnetic wave with a fixed

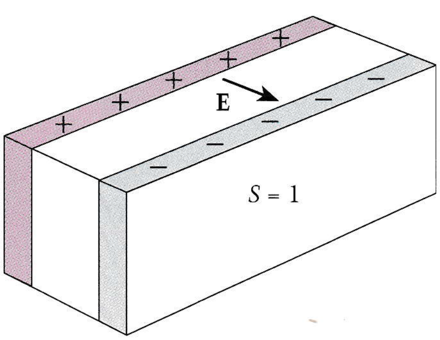


Figure 7
Deriving the Langmuir frequency.

frequency very close to $2\omega_p$. Such a wave can be detected. Since this frequency $\omega_t \cong 2\omega_p = (4ne^2/\varepsilon_0 m)^{1/2}$ can be measured by studying the radiation, this is one of very few possibilities for measuring the particle concentration of plasmas in outer space, especially of intergalactic plasma.

The reasoning and the computation given above both show that some perturbations that disrupt quasi-neutrality are still possible. But they must be small compared to unity (or, which is the same thing, δn must be much less than $n_{e,i}$). Also, their characteristic time scale cannot be too long:

$$\Delta t \le \omega_p^{-1}.$$

During this time the fastest plasma particle (the electron) can travel a distance of not more than $v_{Te}\Delta t$. Consequently, this estimate can serve as the upper limit for the radius of interaction:

$$\Delta x = v_{Te}\Delta t \le \frac{v_{Te}}{\omega_p} \cong r_{De} = \left(\frac{\varepsilon_0 kT}{ne^2}\right)^{1/2}.$$

The magnitude r_{De} is the Debye radius (or radius of Debye screening). The field due to this charge separation acts effectively only over distances less than r_{De}. In other words, r_{De} is the maximum length of the springs in figure 6. And now we're ready to introduce an effective criterion for the predominance of plasma's collective properties. If the number of charged particles inside the Debye sphere is great enough—$nr_{De}^3 \cong (e_0 kT)^{3/2}e^3 n^{-1/2} >> 1$—then collective interactions dominate. In the opposite case, we have something like a gas with modified interactions between particles. From this equation we can readily see that collective interactions become more dominant as the plasma becomes hotter and more rarefied. For example, the plasma corona of a pellet irradiated by a powerful laser is very dense—$n_e \cong 10^{27}$ m^{-3}—but extremely hot: $T \cong 10^7$–10^8 K. Interstellar plasmas are rather cold—$T \cong 10^4$ K—but extremely rarefied: $n \cong 10^6$ m^{-3}. In both cases, $nr_{De}^3 >> 1$ and collective interactions dominate.

Conclusion

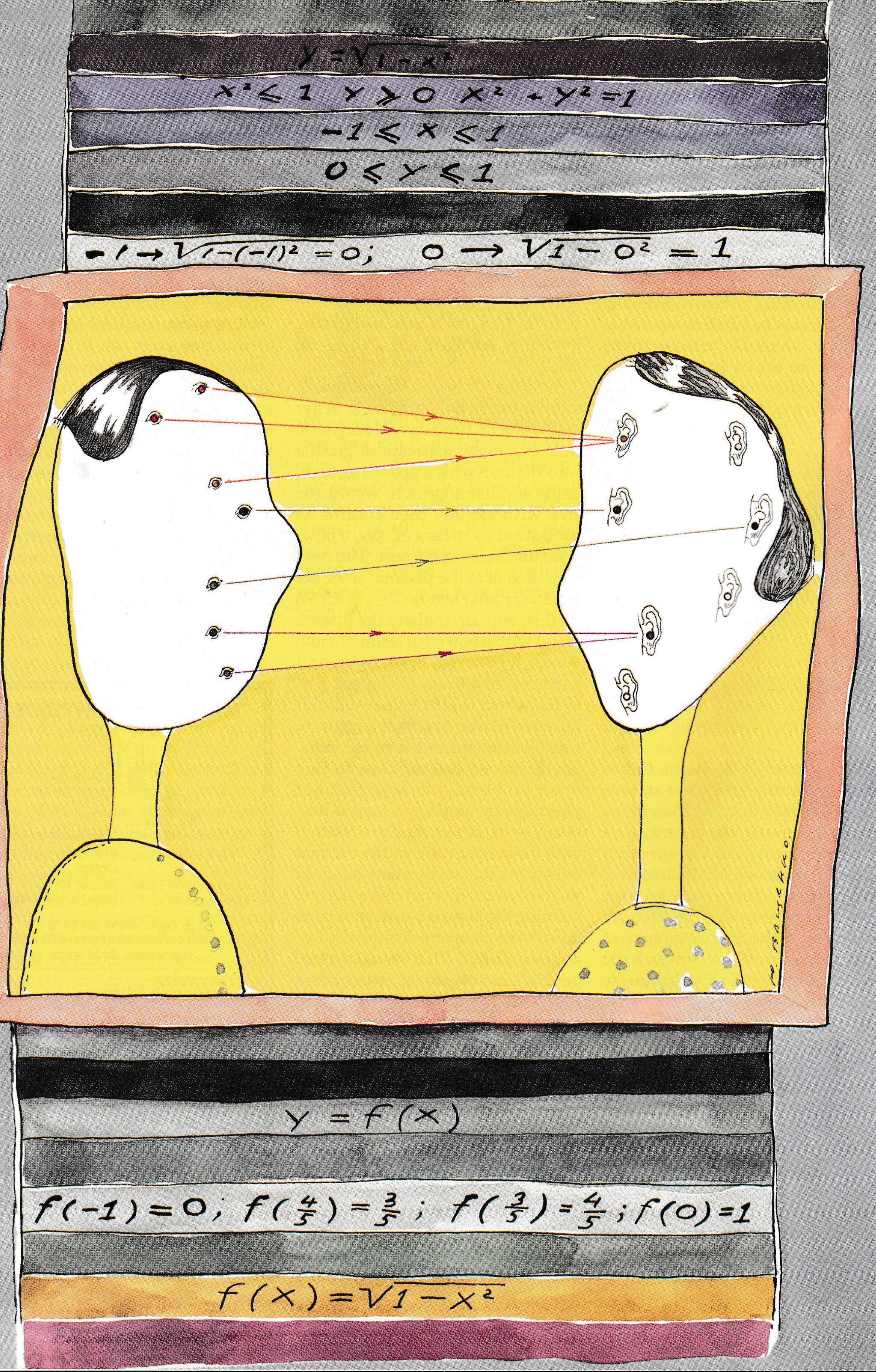
We should avoid the dangers of the *too* popular interpretation of plasma physics. It is, in fact, rather complicated and full of facts, details, and possibilities for viewing it from different angles. Let's try to come up with an imaginary scenario for the dynamics of a plasma in a magnetic trap.

Imagine we have a magnetic trap similar to the one in figure 2. After switching on the external magnetic field, we inject a stream of plasma into the trap with a plasma gun; or, using another approach, we fill the trap with gas and then control its ionization by means of, say, a powerful discharge of current. The next step is to heat the plasma, since we need very hot plasma ($T \cong 10^8$ K). To do this, we can irradiate the plasma cloud with a powerful beam of radio waves, a powerful beam of neutral particles, or a beam of charged particles (which is a little more difficult because of the external magnetic field). It's also possible to use a discharge of a very strong current. One other problem is to keep this hot plasma in the trap for as long as necessary—that is, we need to maintain both the plasma itself and its thermal energy. At this stage, many different kinds of instability enter the picture, coloring the plasma's existence right down to its complete dissolution. I've oversimplified this scenario, of course—in actual fact, all of these events occur simultaneously.

If we take space plasma as another example, we can't add any artificial elements to the scenario, since all the effects and processes occur naturally. But this doesn't make its nonlinear dynamics any simpler. In both space and laboratory plasmas, all the basic elements of the collective scenario are operative: oscillations and collective motion; disorder in its macrophysics accompanied by additional order in its microphysics; chaos arising out of structure and structure based on chaos. All these features are typical for both particle motion and the dynamics of the electromagnetic field.

So it turns out that everything in plasma physics is rather complicated and hard to investigate. But if this weren't the case, what would the poor physicist have to do? Look for another job? (Or, at least, another problem?) By the way, this is one of the few areas that deal with fundamental processes while remaining classical—that is, non–quantum-mechanical. (In fact, plasma physics and a sizable chunk of astrophysics may be the only areas.) The great advantage of plasma physics is that it's a very visual process. Alas, this quality—so attractive in the physics of the 19th century—has been completely lost in modern physics.

So, if you prefer visual yet up-to-date science, try getting acquainted with plasma physics!



$Y=\sqrt{1-x^2}$
$x^2 \leq 1 \quad Y \geq 0 \quad x^2+y^2=1$
$-1 \leq x \leq 1$
$0 \leq Y \leq 1$
$-1 \to \sqrt{1-(-1)^2}=0; \quad 0 \to \sqrt{1-0^2}=1$
$Y=f(x)$
$f(-1)=0; \; f(\frac{4}{5})=\frac{3}{5}; \; f(\frac{3}{5})=\frac{4}{5}; f(0)=1$
$f(x)=\sqrt{1-x^2}$

Home on the range

And in the domain as well

by Andrey N. Kolmogorov

THIS ARTICLE EXPLAINS what mathematicians mean nowadays when they use the word "function." It won't make for easy reading—you'll have to pay attention to every word, although it doesn't presuppose any special knowledge beyond the scope of the high school curriculum. Also, it's assumed that you understand the words "set" and "element of a set" and know how to handle them.

Introduction

When asked what a function is, a student will often say, "A function can be given by a table, graph, or formula." But, of course, this isn't a *definition*. Then again, students who avoid an explicit general definition and try to describe directly the ways of specifying a function aren't entirely wrong. Mathematics can't begin with definitions. When we formulate a definition of some concept, we inevitably have to use some other concepts in the definition itself. Until we understand the meaning of some concepts, we can't really get started and can't formulate a single definition. And so the exposition of any mathematical theory begins with the acceptance of some *basic concepts* that aren't defined. Then they can generally be used to formulate the definitions of subsequent *derived notions*.

So how do we explain to one another our understanding of what the basic concepts mean? There's no way to do this other than by illustrating the things to be defined with *examples* and by comprehensively describing their characteristic properties. These descriptions can be a bit unclear in some details, and they may not be exhaustive at first. But gradually they etch the meaning of the concept with sufficient clarity. This is how we'll approach the concept of a *function*, regarding it as one of the basic mathematical concepts that cannot be formally defined.

[Now, it's true, later on I'll say that a function is nothing but a *mapping* of one set onto another set (the *domain* of the function onto its *range*). But the word "mapping" here is simply a synonym of "function." These are two names for one concept. And the explanation of one word by means of another that's equivalent can't replace the definition of the concept it expresses.]

Example 1. Let the letters x and y denote *real numbers*; the radical sign "$\sqrt{}$" denotes the extraction of the (positive) square root. Then the equality

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

means that the conditions

$$x^2 \le 1,\ y \ge 0,\ x^2 + y^2 = 1 \qquad (2)$$

are valid. The points with coordinates that obey these conditions constitute the semicircle shown in figure 1.

Figure 1 graphically illustrates the following facts, which you may prove in a purely algebraic way.

1. For any x satisfying

$$-1 \le x \le 1, \qquad (3)$$

equation (1) enables us to calculate the corresponding y, which satisfies the inequalities

$$0 \le y \le 1. \qquad (4)$$

2. For every y satisfying inequalities (4), there is at least one x to which equation (1) assigns this y.

We can say that equation (1) defines a *mapping* of the set of numbers x satisfying inequalities (3) onto the set of numbers that obey inequalities (4). To denote a mapping, mathematicians often use an arrow. Thus, the mapping in question can be written as

$$x \to \sqrt{1-x^2}. \qquad (5)$$

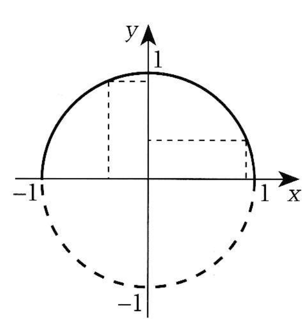


Figure 1

Art by Yury Vashchenko

For instance,

$$-1 \to \sqrt{1-(-1)^2} = 0,$$

$$-\frac{4}{5} \to \sqrt{1-\left(-\frac{4}{5}\right)^2} = \frac{3}{5},$$

$$\frac{3}{5} \to \sqrt{1-\left(\frac{3}{5}\right)^2} = \frac{4}{5},$$

$$0 \to \sqrt{1-0^2} = 1. \qquad (6)$$

Notice that *a mapping is completely defined if (a) the set E that is mapped is given; (b) for every element x of the set E an element y onto which x is mapped is given.*

We denote by M the set of all values of y. In example 1, E is the set of numbers satisfying condition (3), and M is the set of numbers satisfying condition (4).[1]

Example 2. The two rules

$$1.\ x \to \sqrt{x^2}$$

and

$$2.\ x \to \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0 \end{cases}$$

define the *same* mapping

$$x \to |x| \qquad (7)$$

of the real numbers onto their absolute values (fig. 2).

The mapping (7) sends the set of all real numbers

$$\mathbf{R} = (-\infty, \infty)$$

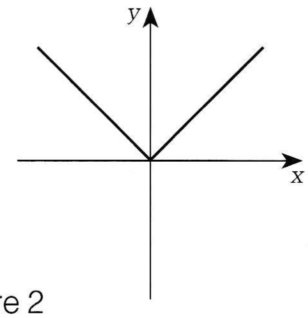


Figure 2

onto the set

$$\mathbf{R}_+ = [0, \infty)$$

of the nonnegative real numbers.

Instead of *mapping* we can say *function* and write the mapping (5) as

$$f(x) = \sqrt{1-x^2} \qquad (8)$$

and the mapping (7) as

$$f(x) = |x|. \qquad (9)$$

Then the special values of function (8) listed by formulas (6) will be written in the form

$$f(-1) = 0,$$

$$f\left(-\frac{4}{5}\right) = \frac{3}{5},$$

$$f\left(\frac{3}{5}\right) = \frac{4}{5},$$

$$f(0) = 1.$$

The domain of function (9) is the set of all real numbers $\mathbf{R}$. Its range is the set $\mathbf{R}_+$ of nonnegative real numbers.

Example 3. Petya, Kolya, Sasha, and Volodya live in the same room in a dormitory and take turns with the housekeeping. They've established a "duty roster" for February (fig. 3). The similarity between this table and the ordinary graphs of functions that we know from high school algebra immediately catches the eye. Does this analogy have an exact logical sense? Have the students established here a *mapping* of one set onto another—that is, have they defined a certain *function*? And haven't they drawn the *graph* of this function?

The general concept of a function

We see in example 3 that a certain student is assigned to duty for each of the 28 days of February. In other words, the set of the days in February is *mapped* onto the set of the students distributing the duties among themselves. As a kind of shorthand, we can say that the letter x denotes any day in February and $y = f(x)$ denotes the student on duty for that day. There's no reason why the mapping

$$\text{day } x \to y = \text{student on duty on day } x$$

shouldn't be called a *function*. We can write this mapping as

$$y = f(x).$$

Any mapping f of a set E onto a set M will be called a function with domain E and range M.

Don't forget that, when we talk about a mapping f of a set E onto a set M, we keep in mind that $y = f(x)$ is defined for *any* x from E and *only* for an x from this set, while the values y of function f necessarily belong to M and every y from M is a value of f *for at least one value of the argument x.*

	1	2	3	4	5	6	7	8	9	10	11	12	13	...	28
Petya	✓				✓				✓				✓	...	
Kolya		✓				✓				✓				...	
Sasha			✓				✓				✓			...	
Volodya				✓				✓				✓		...	✓

Figure 3

[1]A set can be labeled with any letter. Here we chose the letters E, from the French word *ensemble*, and M, from the German *Menge*—both words mean a set or collection of items. Coincidentally, the Russian word for "set" also begins with M: множество (MNOH-zhest-vuh). But one needn't stick with these letters—in the next example we'll denote the set of real numbers by the letter $\mathbf{R}$, as is customary.

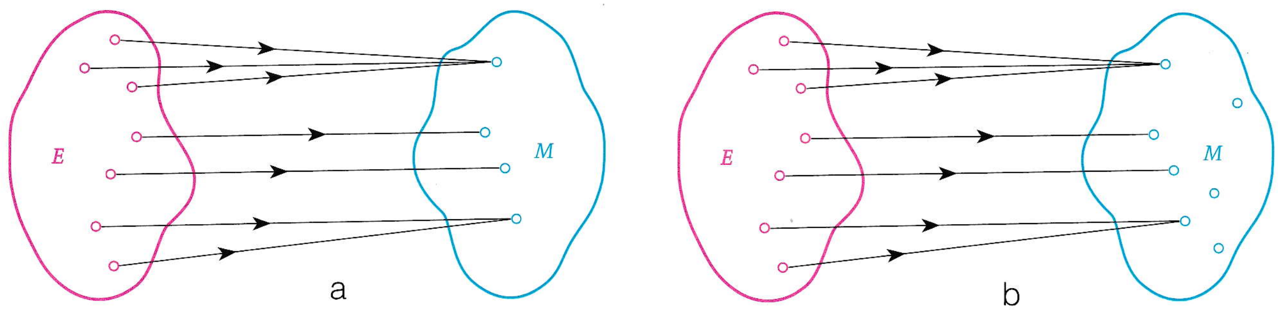


Figure 4
Mappings of (a) E onto M, (b) E into M.

If all we know is that the values of function f necessarily belong to M, but it's not stated that *any* element of this set is a value of function f, the function is said to map its domain E *into* the set M, or that the mapping f is a mapping of E *into* M.

So we must strictly distinguish between the two expressions: "mapping *onto* set M" (fig. 4a) and "mapping *into* set M" (fig. 4b).[2]

For example, the mapping

$$x \to |x|$$

can be described as a mapping of **R** *into* **R**, but not of **R** *onto* **R**.

From the purely logical point of view, the case of the finite domain is the simplest one. Clearly a function with a domain that consists of n elements cannot take more than n different values. So functions with finite domains map finite sets onto finite sets. Such mappings are studied by an important branch of mathematics—*combinatorics* (see problems 8, 11, 18, 19).

Example 4. Let's consider the functions whose domain is the set $M = \{A, B\}$ of two letters A and B, and whose values belong to the same set—that is, the mappings of the set M *into itself*.

There are exactly four such functions:

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

Functions f_1 and f_2 are *constants*—the range of either consists of a single element.

Functions f_3 and f_4 map M *onto* itself. Function f_3 can be given by the formula

$$f_3(x) = x.$$

This is the *identity* mapping: each element of E (= M) is mapped into itself.

To conclude our examination of the general concept of "function," we just need to make note of the fact that it's absolutely insignificant which letters are chosen as notations of the "independent variable"—that is, an arbitrary element of the domain, and the "dependent variable"—that is, an arbitrary element of the range. When we write

$$x \xrightarrow{f} \sqrt{x},$$

$$\xi \xrightarrow{f} \sqrt{\xi},$$

$$y \xrightarrow{f} \sqrt{y},$$

$$f(x) = y = \sqrt{x},$$

$$f(\xi) = \eta = \sqrt{\xi},$$

$$f(y) = x = \sqrt{y},$$

we define *one* function f that maps a nonnegative number into its square root. Using either of these notations, we get

$$f(1) = 1,\ f(4) = 2,\ f(9) = 3,$$

and so on.

Invertible functions

A function

$$y = f(x)$$

is called *invertible*[3] if it takes each of its values once and only once. Functions $f_3(x)$ and $f_4(x)$ in example 4 are invertible; functions $f_1(x)$ and $f_2(x)$, as well as those in examples 1, 2, and 3, are *noninvertible*.

To prove that some function is noninvertible, it suffices to indicate two values of the argument $x_1 \neq x_2$ such that

$$f(x_1) = f(x_2).$$

In example 3 we simply note that Petya is on duty on both February 1 and February 5. So the function in this example is noninvertible.

Example 5. The function

$$x \xrightarrow{f} y = -\sqrt{x}$$

is *invertible*. It is defined on the set $\mathbf{R}_+$ of the nonnegative numbers. Its range is the set

$$\mathbf{R}_- = (-\infty, 0]$$

of all nonpositive numbers. Given any y from the set $\mathbf{R}_-$, we can find the corresponding x by the formula $x = y^2$.

The function g—

$$y \xrightarrow{g} x = y^2 \quad \text{for } y \leq 0$$

[2]Notice also that any "mapping onto" can be called a "mapping into," but not vice versa.

[3]The origin of this name becomes clear below: a function is invertible if it has an inverse.

—is the *inverse function* (or just the *inverse*) of f. It maps $\mathbf{R}_-$ onto $\mathbf{R}_+$. As already noted, the choice of notations for the independent and dependent variables is irrelevant.

Functions f and g can be written

$$f(x) = -\sqrt{x} \quad \text{for } x \geq 0,$$
$$g(x) = x^2 \quad \text{for } x < 0.$$

Figure 5 shows the graphs of the inverse functions f and g.

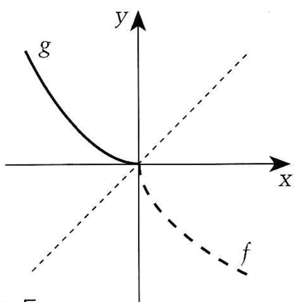


Figure 5

Example 6. The function f is defined on the set of the first five letters of the English alphabet, and its range is the set of the first five natural numbers (positive integers):

x	A	B	C	D	E
$y = f(x)$	3	1	2	5	4

The inverse function g is specified by the following table:

x	1	2	3	4	5
$y = g(x)$	B	C	A	E	D

The graphs of these functions are given in figure 6.

Let's give exact *definitions*. Let f be a mapping of a set E *onto* a set M. If for any element y of the set M there exists a *unique* element

$$x = g(y)$$

of the set E such that

$$f(x) = y$$

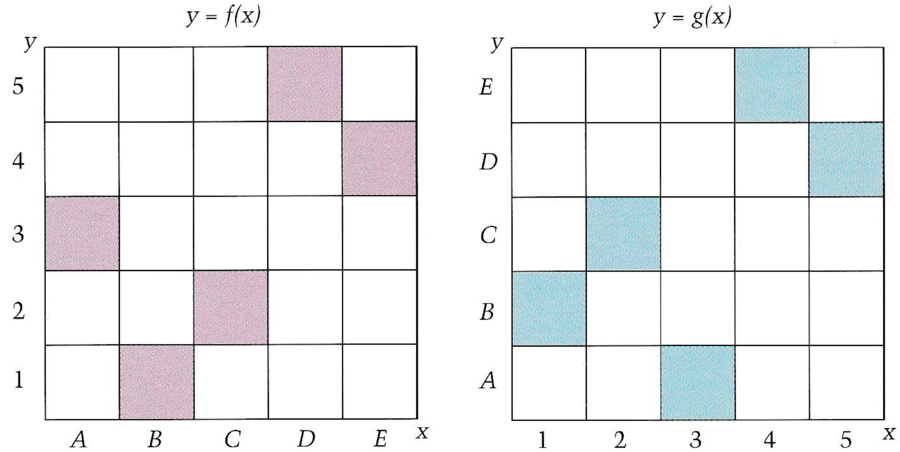


Figure 6

(fig. 7), then the mapping f is *invertible*,[4] and

$$y \xrightarrow{g} x$$

is called the *inverse* mapping for f.

So, the invertibility of a mapping f means that it has an inverse g. The inverse mapping of f is usually denoted by f^{-1}. For instance, if

$$f(x) = x^3,$$

then

$$f^{-1}(x) = \sqrt[3]{x}.$$

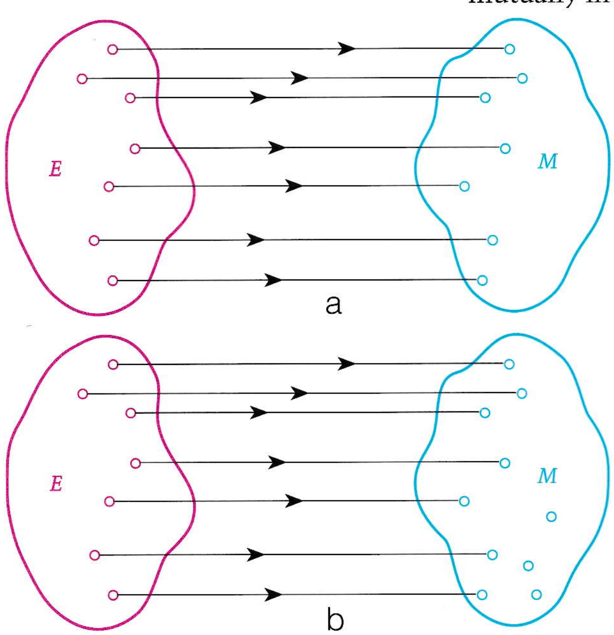


Figure 7
Invertible mappings of (a) E onto M, (b) E into M.

[4]Another name for such a mapping is a *one-to-one* mapping of E *onto* M.

Since "function" is a synonym of "mapping," we have thus defined the term "inverse function" as well. Try to repeat what was said above using the word "function" instead of "mapping."

It's clear that the domain of the inverse function f^{-1} is the range of f, while the range of f^{-1} is the domain of f.

The inverse of the inverse f^{-1} is the original function f:

$$(f^{-1})^{-1} = f.$$

So the functions f and f^{-1} are always mutually inverse.

Example 7. There exist functions that are their own inverses. These are the functions

(a) $f(x) = x,$

(b) $f(x) = \dfrac{1}{x},$

(c) $f(x) = \dfrac{x}{x-1}.$

Check them yourself! The graphs of these functions are plotted in figure 8. Notice that all these graphs are symmetric about the bisector of the first and third quadrants—that is, the line $y = x$.

Figure 9 shows the relationships between the different ways of mapping set A *onto* set B and mapping set

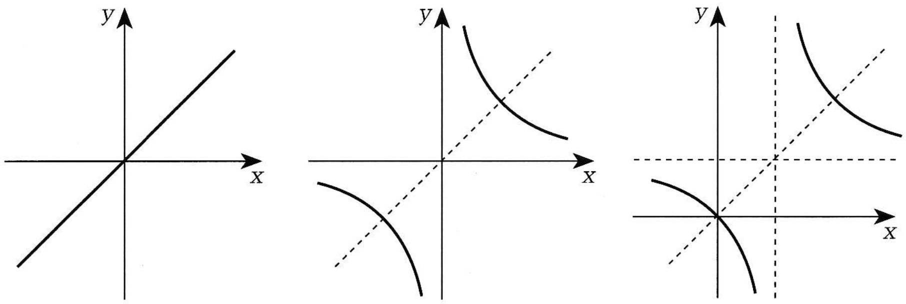


Figure 8

A into set *B*.

Let me remind you once again that the concept of a mapping of *A into B* is the most general. If the image of *A* in such a mapping coincides with *B*, we say it's a mapping of *A onto B*.

Invertible mappings are also called *one-to-one* mappings. You'll encounter this term often in math books. But it's not so customary to speak of "one-to-one functions." Since we think of "functions" and "mappings" as synonyms, we preferred to use "invertible function" or, what is the same thing, "invertible mapping."[5]

Lately, the French terminology has also found acceptance:

1. The mappings of *A onto B* are called *surjections*;
2. The invertible mappings of *A into B* are called *injections*;
3. The invertible mappings of *A onto B* are called *bijections*.

Notice that the accurate use of the prepositions "into" and "onto" makes this abundance of terms superfluous.

Problems

Degree signs indicate very easy questions intended to help you verify whether you've understood this article. More difficult problems are marked with a star. It's not necessary to solve them all.

Introduction

1°. Find the domains and ranges of the following functions: (a) $y = f(x) = 1/x^2$, (b) $y = f(x) = \sqrt{x^2 - 1}$.

2. The *integer part* of a number x is the greatest integer not exceeding x. The integer part of x is denoted by $[x]$. For instance,

$$[0] = 0, [7.5] = [7] = 7,$$
$$[-0.3] = -1, [-\pi] = -4.$$

The difference $x - [x]$ is called the *fractional part* of x and is denoted by $\{x\}$. Sketch the graphs of the following functions and find their domains and ranges: (a) $f_1(x) = [x]$, (b) $f_2(x) = \{x\}$, (c) $f_3(x) = \{x\} - 1/2$, (d) $f_4(x) = |\{x\} - 1/2|$, (e)* $f_5(x) = [1/x]$, (f)* $f_6(x) = 1/[x]$, (g)* $f_7(x) = \{1/x\}$, (h)* $f_8(x) = 1/\{x\}$.

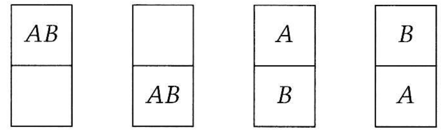
3*. For any natural n we'll define $s(n)$ as the sum of the divisors of n (not counting n itself). For instance,

$$s(1) = 0, s(2) = 1, s(6) = 6,$$
$$s(12) = 16, s(28) = 28.$$

Prove that $s(n)$ doesn't take the values 2 and 5.

The general concept of a function

4°. Two persons *A* and *B* can occupy two rooms in four different ways:



How many ways are there to place (a) two persons in three rooms; (b) three persons in two rooms; (c) three persons in two rooms, so that none of the rooms is left unoccupied?

5°. A set *M* contains three elements and a set *N* contains two elements. Find the number of mappings of (a) *M* into *N*, (b) *M* onto *N*, (c) *N* into *M*, (d) *N* onto *M*.

6. How many seven-digit telephone numbers are there? How many of them consist of the digits 0, 1, 2, and 3 only?

7. Prove that there are more than a million functions taking only two values 0 and 1 and defined on the set of the first twenty natural numbers.

8. The set *M* consists of m elements and the set *N* consists of n elements. How many functions defined on *M* and taking values in *N* are there?

NOTE. Problems 8, 11, 18, and 19 are basic problems in combinatorics. They are given here to show that combinatorics is in large part a matter of counting the number of mappings of various kinds of finite sets into finite sets.

9. How many ways can n guests

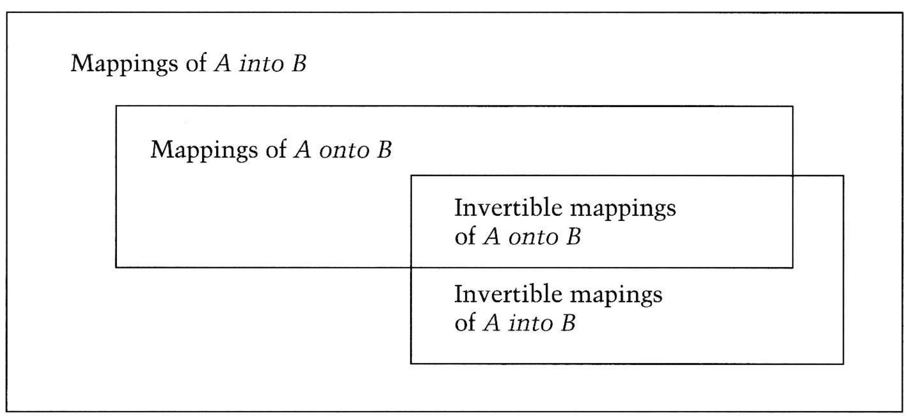
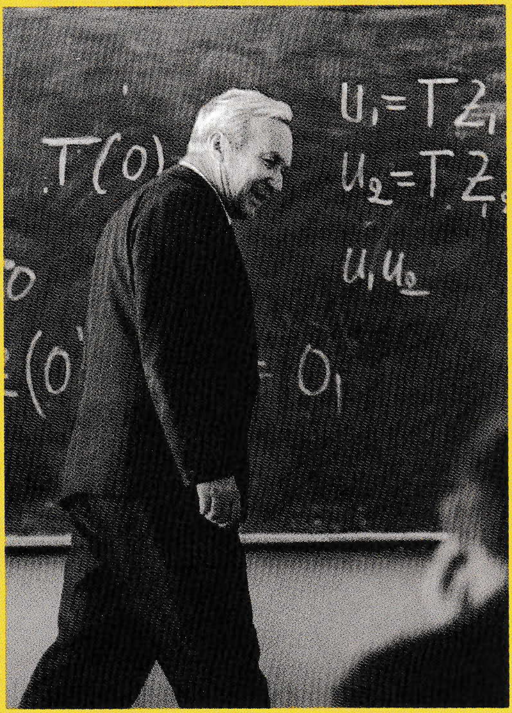


Figure 9

[5]The last remarks are, perhaps, somewhat out-of-date. Nowadays the term "one-to-one function" can hardly be called uncommon (mind you, the article was written in 1969). A. N. Kolmogorov was particular about terminology and its justification because at the time he was working on new high school textbooks and was engaged in many discussions about what, how, and in what terms math should be studied in school.— *Ed.*

A founder's legacy



APRIL 25, 1993, WOULD HAVE been the 90th birthday of Andrey Nikolayevich Kolmogorov, one of the greatest mathematicians of modern times. He blazed a trail in so many mathematical fields—the theory of functions, topology, differential equations, the theory of turbulence, dynamic systems, and so on—it's difficult to name a branch of mathematical analysis in which he didn't make a substantial contribution, where he didn't solve old (sometimes 200-year-old) problems.

Many of Kolmogorov's works were devoted to diverse applications of mathematics. He published articles on geology, mechanics, biology, metal crystallization, information theory, even versification. But most of his world fame is due to his fundamental work in probability. His superior role in the formation of this important branch of mathematics is vividly illustrated by the following quotation from the preface to a textbook by Moscow University professor V. N. Tutubalin: "It is striking that an elementary course for university students, intended to sum up the development of the theory of probability and random processes and provide its simplest and most essential parts, ended up consisting of results more than half of which belong to this great scientific personality. . . . Without his works, personal influence, and example we would be unable to understand how to effectively apply the theory of random processes, and in what areas of science such applications are feasible. . . . We must be grateful to A. N. Kolmogorov for establishing a truly bright ideal of scientific achievement."

These words express the typical attitude of Kolmogorov's students toward their outstanding teacher. And, without exaggeration, these scientists are the pride of the Russian school of mathematics. But Kolmogorov's fostering of new research mathematicians wasn't the only manifestation of his teaching ability. From the mid-sixties on he devoted most of his time to math education at the high school level. He taught in regular schools and was one of the founders of the Special Math and Science Boarding School at Moscow University (which now bears his name). He spearheaded the reform of math education in the Soviet Union and wrote a number of textbooks.

In a certain sense, it can be said that Kolmogorov is the grandfather of the magazine you're reading: together with the eminent physicist I. K. Kikoyin he founded *Kvant*, the Russian predecessor of *Quantum*, in 1970. He was its permanent editor in chief for mathematics until his death in 1987. (You can find more about Kolmogorov's life and work in the Innovators department of the January 1990 issue.)

The accompanying article was written specially for the very first issue of *Kvant*, when Kolmogorov was working most actively on reshaping the mathematics curriculum. Since that time its subject matter—the concept of a "function" and the functional approach in general—has become commonplace in textbooks throughout the world. Still, we think that a presentation by this great mathematical pioneer will be of interest to our readers—both teachers and students.

be seated on n chairs for (a) $n = 2$, (b) $n = 3$, (c) $n = 6$?

10. A set E consists of six elements. Prove that there are exactly 720 functions for which E is both the domain and the range.

11. A mapping of a finite set *onto* itself is called a *substitution* (or permutation). The number of all different substitutions of a set depends only on the number n of its elements and is denoted by $n!$. Prove that $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, $6! = 720$. Find a general method of computing $n!$.

Invertible functions

12°. Determine which of the following functions are invertible:

$$f_1(x) = x^3,\ f_2(x) = x^4,$$
$$f_3(x) = x^{17},\ f_4(x) = x^{18}.$$

13. No more than two students sit at any desk in a classroom. Assign to every student the student with whom they share the same desk (those sitting alone are each assigned to themselves). What is the inverse mapping?

14. Let every English word be put in correspondence with the word written with the same letters in reverse order (a word is understood as an arbitrary sequence of letters). Is this function invertible? If so, what is its inverse?

15. A mapping of a finite set onto itself is always invertible. Give an example of a noninvertible mapping of the set of natural numbers onto itself.

16. Nine tourists must arrange themselves in three boats. In how many ways can they do this if

CONTINUED ON PAGE 41

MATH INVESTIGATIONS

Endless self-description

Finding the limits in Hilgemeier's "likeness sequence"

by George Berzsenyi

IN CHAPTER 29 OF HIS WONDERFUL BOOK *On Computers and the Imagination* (St. Martin's Press, 1989), Clifford A. Pickover treats an interesting procedure for generating sequences of numbers. I learned about it from a California high school student, Peter Wang, who was one of the USA Mathematical Talent Search participants last year, and I'd like to share it with my readers.

The idea for "Die Gleichniszahlen-Reihe" was first described in a German article (with that title) by M. Hilgemeier as follows: start with a 1 in row one; read that row and record the result ("one 1") in row two as 1, 1; create row three similarly from row two—that is, since there are two 1's in row two, row three will become 2, 1; row three is read as "one 2, one 1," hence row four will be 1, 2, 1, 1. Continuing in this manner, we create the first ten rows as follows:

1
1, 1
2, 1
1, 2, 1, 1
1, 1, 1, 2, 2, 1
3, 1, 2, 2, 1, 1
1, 3, 1, 1, 2, 2, 2, 1
1, 1, 1, 3, 2, 1, 3, 2, 1, 1
3, 1, 1, 3, 1, 2, 1, 1, 1, 3, 1, 2, 2, 1
1, 3, 2, 1, 1, 3, 1, 1, 1, 2, 3, 1, 1, 3, 1, 1, 2, 2, 1, 1

My first challenge to you is to **write a computer program to generate many more rows in the same manner**. The program should also count the number of different entries in each of the rows, as well as the frequency of certain combinations thereof, so as to allow for various conjectures. In particular, you will want to **show that only the numbers 1, 2, 3 appear in any of the rows**. This fact was established by Hilgemeier earlier, but it's still not known whether the combination 3, 3, 3 can ever occur in any row. This was the challenge that prompted Peter Wang to write to me.

More generally, suppose that one starts with a row of 1's, 2's, and 3's such that not more than three 1's or 2's and no more than two 3's appear consecutively. **Will the next row necessarily inherit this property?** One should be able to generate such starting rows randomly and detect possible additional conditions that will guarantee the "likeness" of the next row. Alternately, one should also ask: **Is it possible to go backward from any given row of an even number of entries?** You'll soon discover that the answer to this question is no, but it takes a lot more to detect the reason(s) for it.

It's not difficult to see that no row can contain more than twice as many members than the previous row. By examining the number of entries in each of the first 27 rows (in the 27th there are 1,000 ones, 636 twos, and 376 threes), Hilgemeier found that each row has about 1.3 times as many entries as the previous one. See the article by John Conway in the November/December 1990 issue of *Quantum* (pp. 31, 63) for a precise mathematical formulation of this observation. Interestingly, the number of 1's, 2's, and 3's seems to grow from row to row at the same rate.

In his book, Pickover also explores the possibility of starting with a nonnegative integer, $p \neq 1$, in row one and then generalizing the other rows, as well as the possibility of starting with two distinct nonnegative integers (say, p and q, $p \neq 1 \neq q$) in row one. Clearly, there are a number of other possible "seeds." My final challenge to my readers is: **Generalize**. Of course, the best generalizations are those that will provide additional insight into the specific case under consideration—hence, one should never lose sight of the original problem.

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

NEW OPTIONS IN MATHEMATICS

College of Arts and Sciences, Northwestern University

MENU

Mathematical
Experience for
Northwestern
Undergraduates

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

HISTORY

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

FEATURES OF MENU

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

WHY A SPECIAL PROGRAM?

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

APPLICATION PROCEDURE

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

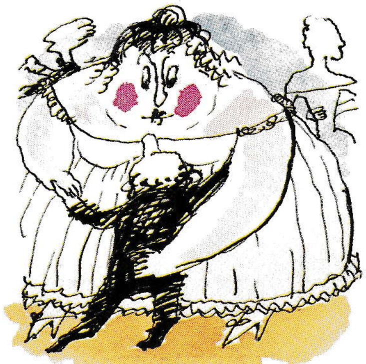
* High School students interested in entering ISP or MMSS must complete the relevant application concurrently with the Northwestern application in the last year of high school.

C A S

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

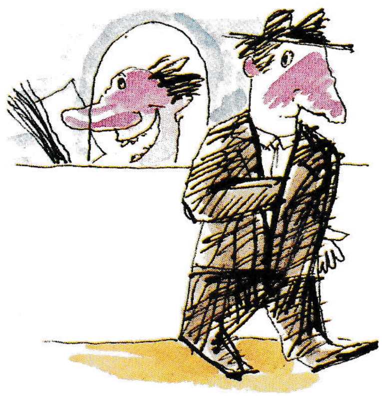
BRAINTEASERS

Just for the fun of it!



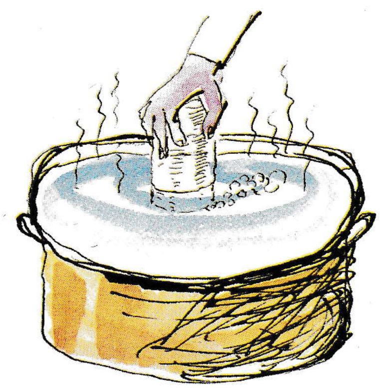
B91

Dancing regularities. At a party each boy danced with three girls, and each girl danced with three boys. Prove that the number of boys at the party was equal to the number of girls. (V. Proizvolov)



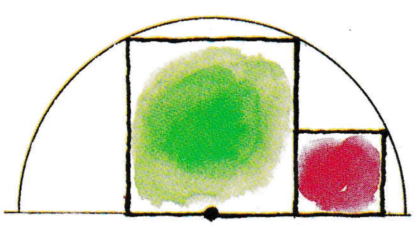
B92

Careless cashier. I went to the bank to cash a check. As the cashier gave me the money, I put it in my empty wallet without counting it. During the day I spent \$6.23. When I checked my wallet in the evening, it contained exactly twice the amount of the check I had cashed. Strange! A little calculation revealed that while making the payment, the cashier had interchanged the figures for dollars and cents. What was the amount of the check? (S. Sidhu)



B93

Bubbles in a glass. An upside-down glass was immersed in a pan filled with hot water. After a while, air bubbles started coming out of the glass. Why?



B94

Squares in a semicircle. Two squares are inscribed in a semicircle as shown in the figure at left. Prove that the area of the big square is four times that of the small one.

B95

Rock to rock. From a pile of 1,001 rocks one rock is taken away and the rest of the pile is divided into two piles. Then one rock is taken away from a pile with more than one rock, and one of the piles is divided into two, and so on. Is it possible that after a number of such operations all piles have three rocks each? (S. Rukshin, S. Genkin)

ANSWERS, HINTS & SOLUTIONS ON PAGE 58

Art by Pavel Chernusky

Cooled by the light

"Degrade first the Arts if you'd Mankind degrade,
Hire Idiots to Paint with cold light & hot shade..."
—William Blake, annotations to The Works of Sir Joshua Reynolds

by I. Vorobyov

ADMIT IT NOW—DOESN'T my title bug you just a little bit? After all, on a hot day we look for a shady spot to get away from the Sun's rays. So how can we cool something by shining light on it? To bring a body's temperature down, we have to remove energy, but light brings energy with it. And yet a "photonic refrigerator" has been created, and it cools to the remarkably low temperature of $2.4 \cdot 10^{-4}$ K.

Now that we know it's possible, let's see if we can figure out how to "make it cold" with light.

The atom team and the photon team

First of all, remember that temperature is related to the average kinetic energy of the thermal motion of atoms:

$$\frac{3}{2}kT = \frac{m\overline{v^2}}{2}. \qquad (1)$$

For this energy to decrease, we need only slow down the atoms. Well, why not use photons to slow them down? We know that a photon has a momentum $p = \varepsilon/c$, where ε is the photon's energy and c is its speed—that is, the speed of light. If a moving atom absorbs the first photon it meets, then this atom will slow down. But how can we get a photon to hit an atom?

Clearly there's no point in trying to aim at an individual atom. We need to hit a beam of atoms with a beam of photons "right between the eyes." We naturally turn to a laser as the source of light that will act as a brake. It emits a directed beam of photons of the same energy. To get a directed beam of atoms, let's do the following. Evaporate a tiny piece of a substance in a vacuum chamber (either with an electric discharge or a pulse of laser light). Two barriers with small holes separate out a thin beam from the escaping cloud of vapor. And that's how we get the "atom team" that will meet the onrushing laser beam (fig. 1).

Isn't it marvelous? The basic outlines of the refrigerator I described have already taken shape. But will the light be absorbed by the relatively rarefied beam? (As a rule, gases are rather transparent. This means that most photons pass through a gas without bumping into any of its atoms.) And if the photon *is* absorbed, will the atom necessarily slow down and lose some of its energy in this process? These two questions are closely linked.

The levels of the possible internal energy of an atom form a ladder with empty intervals between the steps (fig. 2). The first excited state is separated from the ground state (with the lowest internal energy) by the energy E_1; the second excited state is separated from the first excited state by the energy E_2; and so on. The steps themselves have a certain thickness: in a given excited state most of the atoms have energies from $E - \Gamma/2$ to $E + \Gamma/2$, where

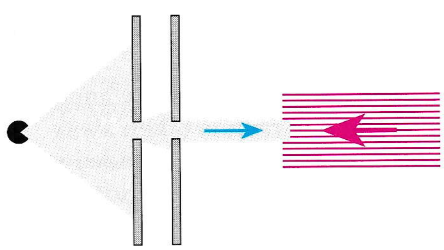


Figure 1

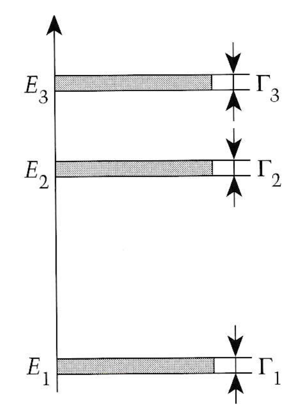
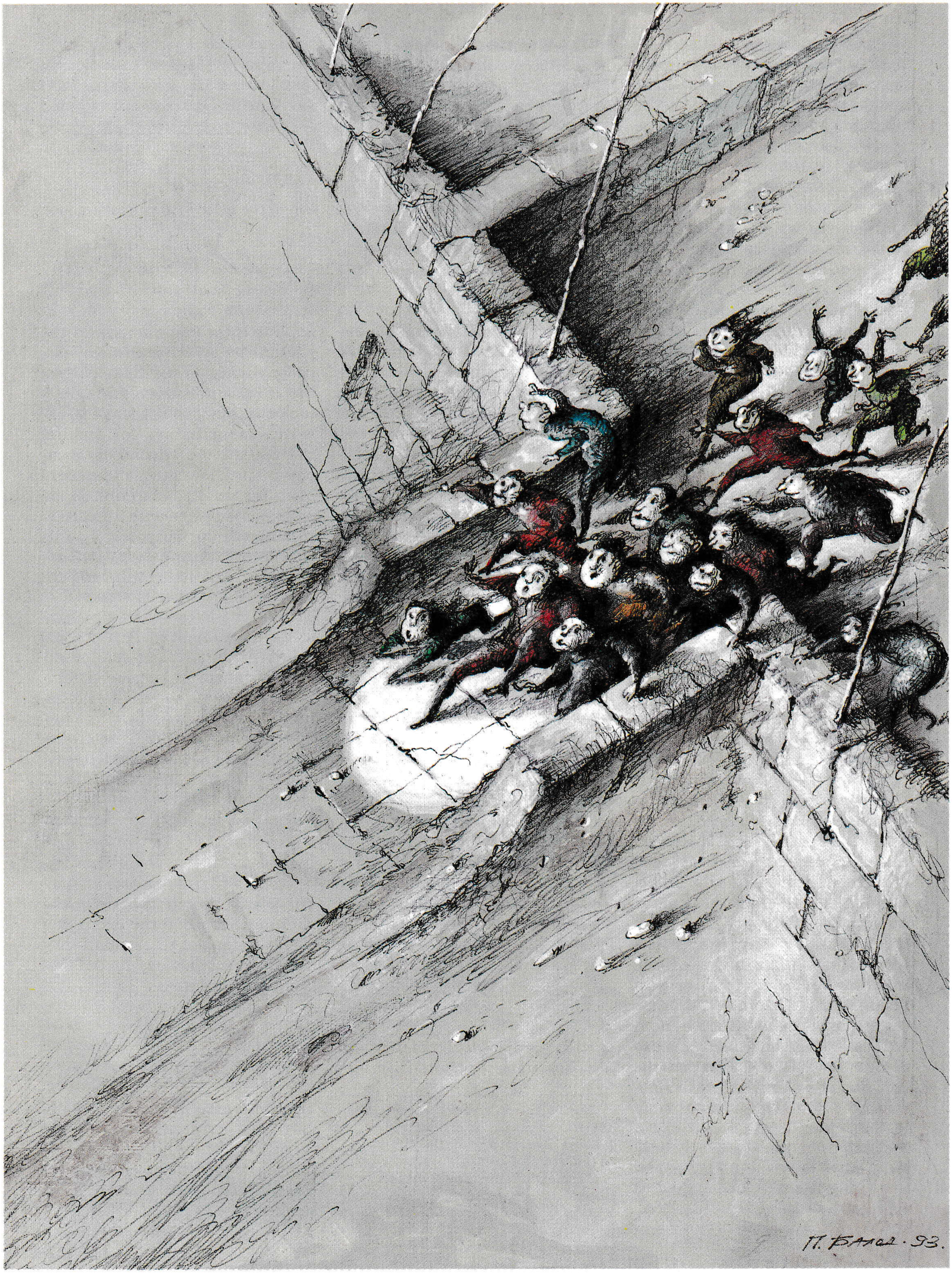


Figure 2

Art by Pavel Balod


П. Багл. 93.

Γ is the width of the level with energy E. The probability of meeting an atom outside this range is small and decreases sharply as the deviation from the middle of the step increases.

There is a definite correspondence between the energy of the atom and that of the absorbed photon: an atom can only absorb a photon whose energy corresponds to a transition from one step to another. It is this resonance in the absorption that in large part explains why gases are so transparent: if the photon energy differs from the transition energy by more than the width of the level, the probability of absorption decreases sharply to zero.

So if we select the right photon energy, absorption occurs with a transition of the atom from one state to another in which the internal energy is greater. Because of the kick received during the absorption of the photon, the atom slows down, and its kinetic energy decreases. This does not contradict energy conservation, though. In this case a portion of the atom's kinetic energy and the photon's energy increase the internal energy of the atom. An examination of the conservation of momentum and energy when the atoms and photons interact will make it possible to understand the idea of cooling by light.

Let's leave calculations aside for the moment and think about this. It's hardly likely that a single absorption can completely extinguish the speed of an atom. So its fate after the absorption is also important. An atom is stable only in the ground state, when it is in its lowest energy level. The lifetime of an atom in an excited state is limited. For the first level, a typical lifetime is of the order of 10^{-8} s. On average, in that period of time an atom returns to its ground state and in so doing emits a photon. How does the emission of a photon affect the motion of the atom? Does it compensate for the slowing down that occurred in the absorption stage? Does it cause the atom to leave the atomic and laser beams? This calls for quantitative estimates.

These emitted photons can also be absorbed by other atoms: the conditions of resonance apply to them as well. The number of these "secondary" photons is, however, much smaller than the number of photons in the intense laser beam. In addition, they are radiated in all different directions, and most of them quickly leave the atomic beam. So we can neglect the absorption of emitted photons.

Tuning and resonance

Let's look at an actual experiment performed with sodium. We'll consider a transition from the ground state to the first excited state, with an energy $E = 2.1$ eV and width $\Gamma = 4.4 \cdot 10^{-8}$ eV.[1] Such a narrow resonance needs a high accuracy in the choice of the photon energy. The atomic mass of sodium is $m = 22$ GeV/c^2, where $c = 3 \cdot 10^8$ m/s is the speed of light.[2] In converting from temperature to energy and vice versa, it's convenient to use a rounded figure for the Boltzmann constant $k \cong 10^{-4}$ eV/K (more precisely, energies of 1 eV correspond to 11,600 K). A speed for the thermal motion of sodium atoms of approximately 10^3 m/s corresponds to a temperature 10^3 K, while a speed of 1 m/s corresponds to a temperature of 10^{-3} K. The orders of magnitude are approximately the same for other atoms as well.

Consider the conditions of resonance absorption in a head-on collision of an atom with a velocity v and a photon with some energy ε (fig. 3). According to the law of conservation of momentum,

$$mv_1 = mv - \frac{\varepsilon}{c},$$

where v_1 is the velocity of the atom after absorbing the photon. According to the energy conservation law,

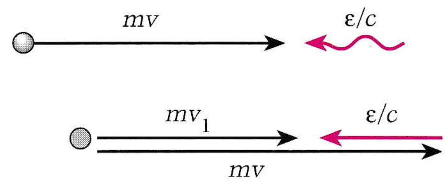


Figure 3

$$\frac{mv_1^2}{2} + E = \varepsilon + \frac{mv^2}{2}.$$

Solving these equations gives the photon energy ε that ensures resonance absorption and the velocity of the atom v_1 after absorption. At a temperature of the order of 10^3 K, which is necessary for sodium evaporation, the velocity of the atoms is much lower than the speed of light. It's reasonable to find an approximate solution with a precision suitable for our purposes. To this end, let's write the conditions for the balance of momentum and energy in a different way—

$$\varepsilon = mc(v - v_1), \qquad (2)$$

$$E - \varepsilon = \frac{m(v^2 - v_1^2)}{2} \qquad (3)$$

—and divide equation (3) by equation (2):

$$\frac{E - \varepsilon}{\varepsilon} = \frac{v + v_1}{2c}. \qquad (4)$$

The relative difference between the photon energy and the excitation energy turns out to be small. Neglecting this difference, for our first approximation we get

$$v - v_1 \cong \frac{E}{mc} \cong 3 \cdot 10^{-2} \text{ m/s}. \qquad (5)$$

This difference in velocities is obtained by substituting ε for the slightly larger magnitude E in equation (2), so that the precise value of this difference is actually a bit smaller.

[1] One electron volt (eV) is equal to the energy gained by an electron in passing through a potential difference of one volt; 1 GeV = 10^9 eV.

[2] The expression GeV/c^2 has the dimensionality of mass and serves as its unit.

In our next approximation we'll exploit the fact that the difference between the velocities is small compared to the velocities themselves. Replacing v_1 with v in equation (4), we obtain

$$\frac{E-\varepsilon}{\varepsilon}=\frac{v}{c}. \qquad (6)$$

Taking the difference in velocities into account results in an energy shift that is much smaller than the width of the level, so it's immaterial to us—it won't affect the resonance condition. For the same reason we're allowed to substitute ε for E in the denominator of equation (6), so

$$E-\varepsilon=E\frac{v}{c} \qquad (7)$$

with a precision sufficient for our purposes.

Therefore, the condition for intensive absorption of photons has been found: their energy must be lower than the energy of the level by a fraction equal to the ratio of the atom's velocity to the speed of light. But the negligible decrease in the velocity—3 cm/s—is vexing. A one-time slowing of the atom in fact gives us nothing.

We should bear in mind that the velocities of atoms in a beam are different, so only some atoms satisfy the resonance condition. In fact, according to equation (7), for the upper and lower boundaries of the level we get

$$E\pm\frac{\Gamma}{2}-\varepsilon=E\frac{V_\pm}{c},$$

where $V_\pm$ are the boundary velocities leading to resonance for a given photon energy. The width of the velocity range is

$$\Delta V=V_+-V_-=c\frac{\Gamma}{E}\cong 6\ \text{m/s}. \qquad (8)$$

Atoms whose velocities are beyond this range will not slow down.

Re-emission and resonance tuning

A one-time slowing at absorption is not enough. However, when an atom has spent a short time in the excited state, it returns to the ground state by emitting a photon. After that, we hope, it's ready for another absorption. So let's look at the radiation of photons by excited atoms when they return to their ground states.

Laser photons go where we point them, but radiated photons fly off every which way. The change in the atom's velocity when it gives up a photon depends on the angle θ between the velocity **v** of the excited atom and the direction of the photon (fig. 4). The energy of the emitted photon ε′ depends on the same angle θ. According to the conservation of energy,

$$\frac{m\left(v_1'\right)^2}{2}+\varepsilon'=\frac{mv_1^2}{2}+E, \qquad (9)$$

where v_1' is the atom's velocity after emitting the photon. The law of conservation of momentum must be written in vector form:

$$m\mathbf{v}_1'+\mathbf{p}'=m\mathbf{v}_1.$$

It's clear that when a photon is emitted "sideways," the direction of the atom's motion changes by some angle ϕ.

Let's do another approximate calculation. In our first approximation, ε′ = E. Since E/c is much smaller than mv_1, then v_1 and v_1' are again rather close and the angle ϕ is rather small. Using the radius mv_1', we'll cut off a portion of the segment mv_1 (see figure 4). The small arc of the circle can be considered a linear segment perpendicular to the segment mv. It's evident from the small right triangle that the length of this arc is $E\sin\theta/c$. Then the angle at which the atom deviates from its original direction is

$$\phi=\frac{E\sin\theta}{mvc}, \qquad (10)$$

and the change in the atom's speed is

$$v_1-v_1'=\frac{E\cos\theta}{mc}. \qquad (11)$$

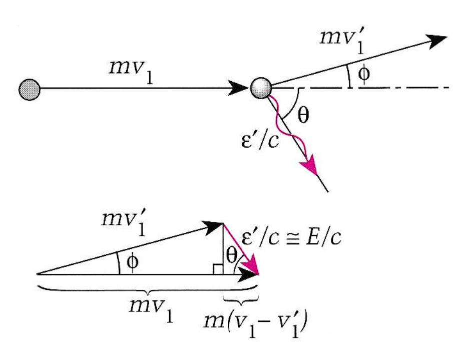


Figure 4

For our next approximation we rearrange the terms in equation (9) and factor the right-hand side:

$$\varepsilon'-E=\frac{m}{2\left(v_1'-v_1\right)\left(v_1'+v_1\right)}.$$

Substituting from equation (11) and setting $v_1'=v_1$, we get the difference between the energy of the radiated photon and the energy of the excited state:

$$\varepsilon'-E=\frac{Ev_1\cos\theta}{c}.$$

The magnitudes ε′ and ε are the same only when the photon is radiation backward—that is, at θ = π (see equation (7)). Only in this single case will the loss when the photon is emitted compensate exactly for the slowing that occurs when it is absorbed. (Convince yourself that this is an exact, not an approximate, result.)

In each individual act of radiation, the change in the atom's velocity depends on the angle of the photon emission, and according to equation (11) it falls within the range $-E/mc$ to $E/mc\cong 3$ cm/s. Due to the randomness of the emission angle, the average value of the velocity change can be regarded as zero.

According to equation (10), the largest angle of deviation of the atom's velocity is

$$\phi_{\max} = \frac{E}{mv_1c}.$$

For a velocity of the order of 10^3 m/s, this equals approximately $3 \cdot 10^{-5}$ rad, but for a velocity of 10 m/s, it's about $3 \cdot 10^{-3}$ rad.

At first, dozens and even hundreds of absorption–emission cycles won't draw an atom out of the resonance region. The changes in velocities per cycle are small, the deviations of the directions of the velocities are small and random, and the angular deviation of the beam of atoms accumulates very slowly.

But this continues only for a certain time. After the velocity decreases by $\Delta V = c\frac{\Gamma}{E} = 6$ m/s (see equation (8)), resonance stops. For the atom to be slowed further, we now need to fine-tune our system. Analyzing the resonance condition described by equation (6), we might come up with two methods of tuning.

The first method is to smoothly increase the photon energy ε in accordance with the decrease in the atom's velocity. To do this, we need a laser that radiates photons of varying energy. Such lasers exist.

The second method is to manipulate the energy itself. In particular, energy levels vary slightly in an external magnetic field. In a field that is constant in time but varies smoothly along the beam's axis, an atom with any velocity will find a place with a suitable value for the field. Both methods can actually be implemented.

Let's look at the first method in more detail. We'll choose an initial photon energy such that it corresponds to resonance at a velocity V that is greater than the velocities of, say, 90% of the atoms in the beam. We smoothly increase the photon energy to a value equal to the energy E of the level corresponding to resonance for atoms that are almost stationary.

At first, atoms in the range of velocities $V - \frac{1}{2}c\frac{\Gamma}{E}$ to $V + \frac{1}{2}c\frac{\Gamma}{E}$ slow down. Because the atoms slow down, this range gradually empties: the velocities of almost all the atoms will be lower than V.

When the photon energy increases, these slowed atoms will again enter into resonance along with other atoms that initially had lower velocities. Thus, all atoms with velocities lower than V will enter into resonance at some time and begin to slow down.

Let's estimate the slowing time from a velocity V to a velocity of almost zero. The number of absorption–emission cycles during this time is equal to $N = V/(E/mc)$, where E/mc is the decrease in velocity during the act of absorption (see equation (5)); we neglect the change in velocity when an excited atom emits a photon. The length of the cycle Δt is defined by the lifetime τ in the excited state, since absorption takes place quickly when the intensity of the laser beam is high. Thus, the total slowing time is

$$t = N \cdot \Delta t = N\tau = \frac{mcV}{E}\tau.$$

For an initial velocity $V = 3 \cdot 10^3$ m/s, $t = 10^{-3}$ s. (We recall that, for sodium, $m = 22$ GeV/c^2, $E = 2.1$ eV.)

The distance during which the slowing occurs is estimated to be $l = Vt/2 \cong 1$ m. This distance is important not only in choosing the dimensions of the vacuum chamber. The diameter of the laser beam is limited. If there is a marked angular deviation of the velocities from the axis of the beam of atoms, atoms will leave the beam (fig. 5). If the diameter of the beam is of the order of 1–2 cm, a deviation of 10^{-2} rad is sufficient over a distance of 1 m. Deviations of such size are achieved during sideways emission for atoms with velocities of 3–10 m/s. The possibilities of slowing with a single laser beam are limited by the almost complete loss of atoms. These velocities correspond to a temperature range of 0.1 K to 0.01 K. Temperatures of 0.05 K that are actually obtained are in close agreement with our rough calculations.

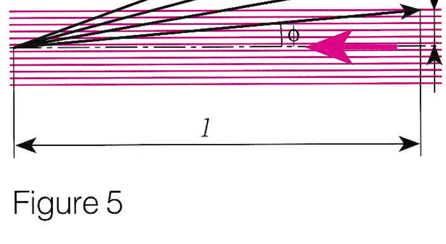


Figure 5

A trap within a trap

Even slow atoms disperse rather quickly due to variations in their velocities. "Supercold" atoms must somehow be gathered and kept in a compact group. To this end, a trap made of light was created—actually, two traps, one inside the other.

The outer, larger trap, which provides additional cooling, is a sort of elaboration of the slowing beam. It's formed by crossing six laser beams whose photon energy $\varepsilon_0 = E - \Gamma/2$ corresponds to the lower boundary of the excited state (fig. 6).

In this case the law of conservation of energy results in the following equality:

$$\frac{mv^2}{2} - \frac{\Gamma}{2} = \frac{mv_1^2}{2}.$$

If the atom's kinetic energy $mv^2/2 > \Gamma/2$, absorption of a "slowing" pho-

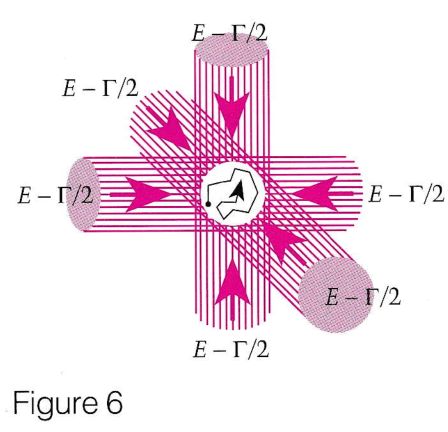


Figure 6

ton is possible, but absorption of a "dispersing" photon doesn't occur. For an atom that has already absorbed a photon to absorb one flying along behind, the second photon must have an energy a bit higher than the lower boundary of the level. So the atom will be slowed down, and the direction of its velocity due to the absorption and emission will change rather markedly. This will lead to a very tangled path and, as a result, to a rather extended period of wandering in the region where the six beams cross. In practice, for a region with a volume of several cubic centimeters, the containment time was found to be approximately 0.5 s. This is more than enough time for the kinetic energy of many atoms to decrease to the half-width of the level.

It's clear that when the atom's energy becomes lower than $\Gamma/2$, absorption will stop. So a limiting temperature is defined by the width of the level: $T_{\text{lim}} = \Gamma/2k$ (see equation (1)). The temperature actually attained, $2.4 \cdot 10^{-4}$ K, is very close to the limiting temperature.

The six-beam trap can't keep atoms there for a long time, though. They leave it during their wanderings and must certainly leave it when they've slowed to an energy of $\Gamma/2$ or to a velocity of about 0.5 m/s. So it turns out that an additional trap is needed.

It was necessary to create a small pit in the six-beam trap where the slowed atoms could go. Unfortunately, to explain this convincingly would require a huge digression. So I'll just tell you that the pit consists of a region with a volume of about 10^{-9} cm^3 near the focus of still another laser beam whose photons have an energy markedly lower than the resonance energy. By moving a lens, one can shift the focus and transport the atoms captured nearby.

The energy depth of the pit is only $5 \cdot 10^{-3}$ K in temperature units. So a collision with an atom of a residual gas at a temperature of 300 K literally kicks the sodium atom out of the pit. With a vacuum of the quality of those obtained to date, it's possible to hold atoms for about 10 s.

The story of the photonic refrigerator stretches from 1968, when the notion of a trap at the focus of a laser beam was first seriously proposed, to 1986, when it became possible to trap 500 atoms of sodium and keep them for several seconds.

Modern advances have given us "an amazing power over atoms," in the words of one of the creators of the photonic refrigerator, Steven Chu. I have a feeling that atoms and photons will offer us yet another opportunity to invent and build something just as beautiful. ◘

HOW DO YOU FIGURE?

Challenges in physics and math

Math

M91

Trapping chips. Two chips, one white and one black, are placed on the extreme squares of a board consisting of 30 squares arranged in a single row. Two players take turns moving their respective chips to one of the neighboring squares (if it's not occupied by the other chip)—a player is not allowed to skip a move. The player who can't make a move loses. Who wins the game: the player who goes first or the player who goes second? Answer the same question for a board of N squares. (A. Talalai)

M92

Legs can't be bases. Prove that it's impossible to construct two trapezoids (not parallelograms!) such that the legs of each are congruent to the bases of the other. (V. Proizvolov)

M93

Thirty thousand thirty. The sum of two integers is 30,030. Prove that their product is not divisible by 30,030. (S. Fomin)

M94

Quadrilateral faces. Does there exist a polyhedron with 1993 faces all of which are quadrilaterals? For what values of the number n of faces does such a polyhedron exist? (V. Dubrovsky)

M95

An end to mimicry? The island of Bluebrownblack is inhabited by 13 blue, 15 brown, and 17 black chameleons. If two chameleons of different colors meet, they simultaneously change to the third color (for instance, a blue and brown chameleon both turn black). Can all the chameleons turn the same color after a number of such meetings? (V. Ilyichov)

Physics

P91

Cartesian diver. A wooden ball floats in water in an enclosed vessel. How does the depth of immersion change when air is pumped into the vessel such that the air pressure in it is doubled? (S. Krotov)

P92

Ideal gasworks. A movable piston inside a horizontal cylinder is attached to the base by a spring. The spring is relaxed when the piston is at the extreme left-hand end of the cylinder (see figure 1). There is an ideal gas to the left of the piston that occupies a volume V_1 under a pressure P_1. On the other side is a vacuum. How much heat must be supplied for the gas to double in volume and pressure? Heat leakage through the walls and heating of the walls, spring, and piston are neglected. The molar heat capacity of the gas at constant volume is $C_V = \frac{3}{2}R$.

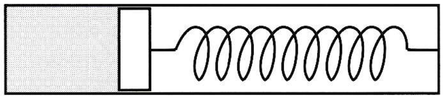


Figure 1

P93

Lightning in water. If lightning hits a body of water during a thunderstorm, a dead fish is sometimes seen afterward on the surface. Why is this? After all, the probability of lightning hitting a fish is negligibly small. (A. Buzdin)

P94

Drawing rays. Figure 2 shows the positions of two point sources A and B. Also shown are the corresponding images A' and B' formed by a lens. Find the position of the lens in each case. (L. Aslamazov)

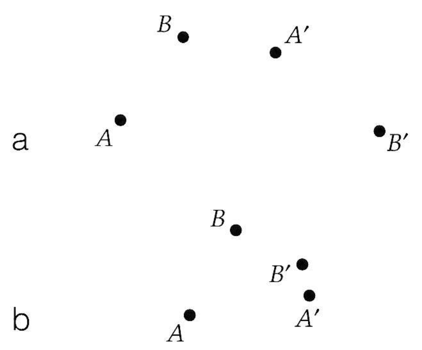


Figure 2

P95

Light pressure. A parallel beam of light strikes a ball of radius R covered with soot (fig. 3). The intensity of the light is I, where I is the energy carried by the light beam through a unit cross-sectional area per unit time. What force does the light beam exert on the ball? (V. Peterson)

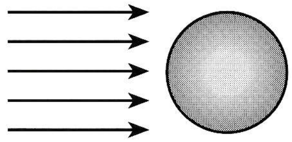


Figure 3

ANSWERS, HINTS & SOLUTIONS ON PAGE 55

IN THE LAB

Focusing on the fleet

Archimedean smoke and mirrors

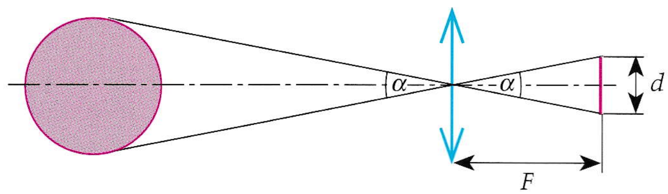
by Sergey Semenchinsky

According to legend, Archimedes used mirrors to burn an enemy fleet anchored near Syracuse. Let's take a close look at this legend from the viewpoint of modern physics and determine the conditions allowing something to be burned with a mirror or lens.

Our reasoning will be based on the fact that the energy radiated from a unit surface of a heated body per unit time is proportional to the fourth power of the body's temperature (this is known as Stefan's law). If the energy supplied to the body is greater than the radiated energy, the body will heat up. This process will go on until the energy radiated per unit time is equal to the energy supplied.[1] Therefore, to heat a body to a temperature T, the energy flux incident on a unit area must be proportional to T^4.

Let's take a lens with a surface area S and focal length F and form an image of the Sun on some surface. The entire flux of solar energy incident on the lens will be collected in the focal plane of the lens at a small spot whose diameter is

$$d = 2F\tan\frac{\alpha}{2}$$



and whose area is

$$s = \frac{4\pi F^2 \tan^2 \dfrac{\alpha}{2}}{4},$$

where α is the angular size of the Sun (see the figure above). Taking angle α to be equal to 30′, we get

$$d = \frac{\pi F^2}{4\cdot(110)^2}$$

(since $2\tan 15' \cong 1/110$). The flux of energy incident on a unit area of the Sun's image is S/s times that of the flux incident on the same area without a lens. It follows from Stefan's law that the steady-state temperature T at the focus is $(S/s)^{1/4}$ greater than the temperature T_0 to which the Sun would heat a surface without a lens. The relationship between the area of a lens (or mirror) and the temperature at its focus is

$$S = \frac{\pi F^2}{4\cdot(110)^2}\,\frac{T^4}{T_0^4}. \qquad (1)$$

It's widely known[2] that the temperature needed to ignite, say, paper is approximately 250°C (520 K), while a temperature of 500–700°C (800–1,000 K) is needed to ignite dry wood.

An experiment performed by the French naturalist Buffon showed that a piece of dry wood soaked with pitch can actually be ignited from a distance of 158 feet (about 47 m). His setup consisted of 168 mirrors, each 48 square inches (310 cm^2) in area, attached to a common frame. Clearly it's easier to ignite wood soaked with pitch than an unsoaked sample. So we can assume at least that the pitched wood would ignite at the lowest temperatures mentioned—that is, at 800 K. Then, following equation (1) and assuming that $T_0 \cong 60°\text{C} \cong 330$ K, we get $S = 4.9$ m^2, while in the Buffon experiment the total area of the mirrors was 5.2 m^2. Thus, Buffon's experiment doesn't contradict our theory.

Now let's try to calculate the area of the mirror that Archimedes

[1]We neglect the thermal flow from the body due to the thermal conductivity of air. Experiments have shown that at the temperature at which wood ignites, this thermal flow is small compared to the radiation.

[2]See, for example, Ray Bradbury's novel *Fahrenheit 451*.

Art by Leonid Tishkov


Тишков
1993

Will it ignite or not?

IN 1974 THE POLISH MAGAZINE *AROUND THE World* published a note about another verification of the legend of Archimedes and the burning ships.

This legend was subject to doubt because the technology of that time wasn't advanced enough to construct a concave mirror so big that the flux of sunlight from it could ignite ships at a considerable distance. So it was considered a legend invented by Greek storytellers.

But if it was impossible to construct one big mirror, why not use many small ones? In December 1973, the Greek physicist Ionas Sakkos verified this hypothesis experimentally, not at Syracuse, though, but in the port of Athens. He used 70 soldiers and a copy of a Roman wooden ship. Each soldier was given a polished copper sheet measuring 150 cm × 90 cm with a handle. The soldiers aimed the reflected sunlight at one spot on the ship, which was located 200 m from the shore.

After several attempts the soldiers managed to collect all their individual spots of light at one point on the ship. In two seconds the ship started to smoke, and in three seconds it burst into flame. A minute later what was left of the ship disappeared beneath the water.

And that was how a December day in 1973 produced yet another triumph for Archimedes.

—V. L. Bulat

would need. We suppose the ships were made from the same material as Buffon's piece of wood. According to the legend, the distance from the mirror to the ships was equal to the length of an arbalest shot (about 400 m).[3] In this case, $S \cong 350$ m^2. However, Archimedes not only had to ignite the ships, he had to ignite them as quickly as possible—since the ships could move (unlike Buffon's piece of wood), they wouldn't wait around to be set on fire. The calculation of the time needed for wood to be heated to its ignition temperature is rather involved—for example, one must take into account the thermal conductivity and heat capacity of wood. So let's just borrow some experimental data from a reputable book.[4] According to these data, the illumination needed to ignite oak boards in 20 seconds is 70 times the illumination of summer sunlight at normal incidence. To obtain such illumination one must enlarge the area of the mirror to 700 m^2. A mirror of this size isn't out of the question—the ancient Greeks were capable of constructing even larger objects.

But here's the rub: the rays reflected from the various parts of the mirror must be brought together at one point, which requires painstaking preparation of such a huge surface. And that's not all. It would be impossible to control the location of the spot of light from such a mirror. To turn a mirror into a controllable weapon one needs to be able to change its focal length, and the time needed for refocusing must be very short. Could such a complicated device be created in ancient times? Apparently not.

However, Archimedes had another option: he could line up several thousand soldiers on the fortress wall, each holding a mirror (1,400 men with 0.5-m^2 mirrors would be sufficient). They would all aim their reflected light at the same spot on a particular ship designated in advance. If the design of the walls at Syracuse allowed them to do this (which is quite likely, since Archimedes himself took part in their construction), and if the mirrors were prepared beforehand, everything could be done quickly enough. Of course, if the Romans knew why thousands of soldiers with mirrors had appeared on the walls of Syracuse and were aware of the danger they were facing, they could fire fusillades from their ships and scatter the unarmed throng. But they didn't know! After the first ship caught fire, the others could have fled. But most likely panic set in—after all, a "miracle" had occurred.

And that's how the people of Syracuse could set fire to several ships and retreat to the safety of their fortifications. Of course, this is just a hypothesis, but as far as the physics goes, there's nothing impossible in it.

[3]The arbalest is a type of crossbow. The length of an arbalest shot was taken from the novel *The White Detachment* by Sir Arthur Conan Doyle.

[4]D. Lawson, *The Atom Bomb and Fires*.

Calling all modem maniacs!

What did you like in this issue of *Quantum*? If you find pen-and-paper communication too old-fashioned, you can send your comments, questions, and suggestions to the managing editor by electronic mail at the following address:

72030.3162@compuserve.com

We look forward to hearing from you.

KALEIDOSCOPE

Keeping track of points

"It is not possible to observe and determine the movements of a finite body without establishing ab ovo[1] the movement of every tiny part of it."
—Leonhard Euler

WE GAIN A WONDERFUL ability to study the most varied and complex motions by reducing them to the simplest possible motion: a point moving along a line. But even this seemingly trivial motion requires a whole bag of concepts to describe it. In this Kaleidoscope you'll get a chance to work with some of these concepts: trajectories, coordinates, tracks, and displacements. Behind each of these ideas is a long history, bound up with the discovery of the laws governing the motion of bodies both on Earth and in the heavens.

Questions and problems

1. Can we predict that two balls will collide if we know that the trajectories of their centers cross?

2. A circle of radius R rolls along a circle of radius $4R$ (fig. 1). How many rotations will the smaller circle make in returning to its starting position?

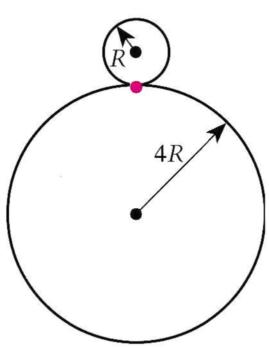


Figure 1

3. A string (assumed not to stretch) is wound on a cylinder whose circumference is l. The other end of the string is attached to a mass (fig. 2). What distance will the mass travel when the cylinder is allowed to roll through one complete turn without any slippage of the string or cylinder?

4. Three masses are attached to a string that is threaded through a system of pulleys as shown in figure 3. Find the direction and magnitude of the displacement vector for mass M_2 if mass M_1 is moved upward 5 cm and mass M_3 is moved downward 3 cm.

5. An object moves along a straight line starting from position x_0 with a velocity that changes with time as shown in figure 4. Draw graphs of the coordinate x, the displacement S_x, and the distance traveled S as functions of time.

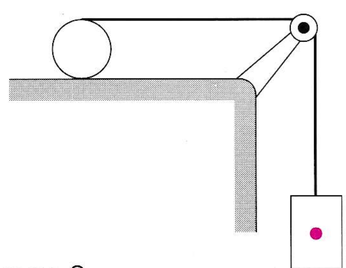


Figure 2

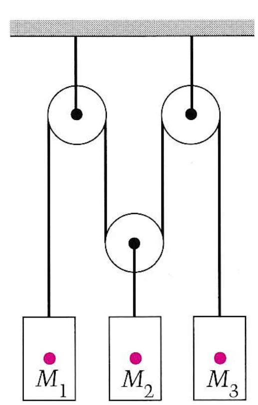


Figure 3

6. What is the trajectory (relative to the Earth) of a mass oscillating on the end of a spring hung from the ceiling of a car that moves uniformly along a straight line (fig. 5)?

7. What is the trajectory of a particle moving in a longitudinal traveling wave?

8. A child throws a ball from a train in the direction opposite to the velocity of the train. How will the ball move relative to (a) the train, (b) the railway bed?

9. What is the trajectory of a charged particle that enters a uniform electric field at an angle to the field?

10. Are there points of a moving

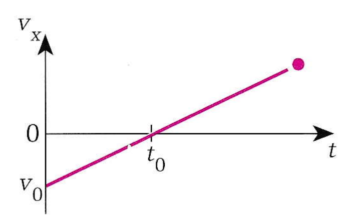


Figure 4

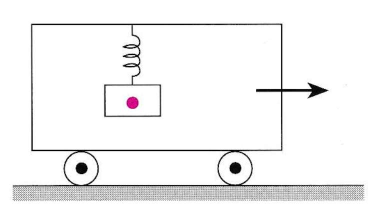


Figure 5

[1] That is, from the beginning (literally "from the egg" in Latin).—*Ed.*

Art by Sergey Barkhin

train that don't move forward but backward? If there are, what are their trajectories?

11. A piece of chalk is set in motion along the diameter of a circle (fig. 6). Relative to the Earth it travels a distance equal to the diameter by the time the circle makes one half of a rotation. What track will the chalk leave on the circle? Friction is negligible.

12. The fragments of a shell that exploded at the top of a tower flew off with the same initial speed v_0. How will these fragments be distributed spatially after the explosion? What is the trajectory of each fragment?

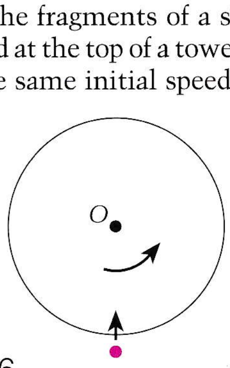


Figure 6

Microexperiment

Hang a heavy bob on a long string and move it from its equilibrium position by a small angle. Let it go (a) without an initial velocity; (b) with a velocity perpendicular to the vertical plane drawn through the support point. Along what trajectories will the bob move?

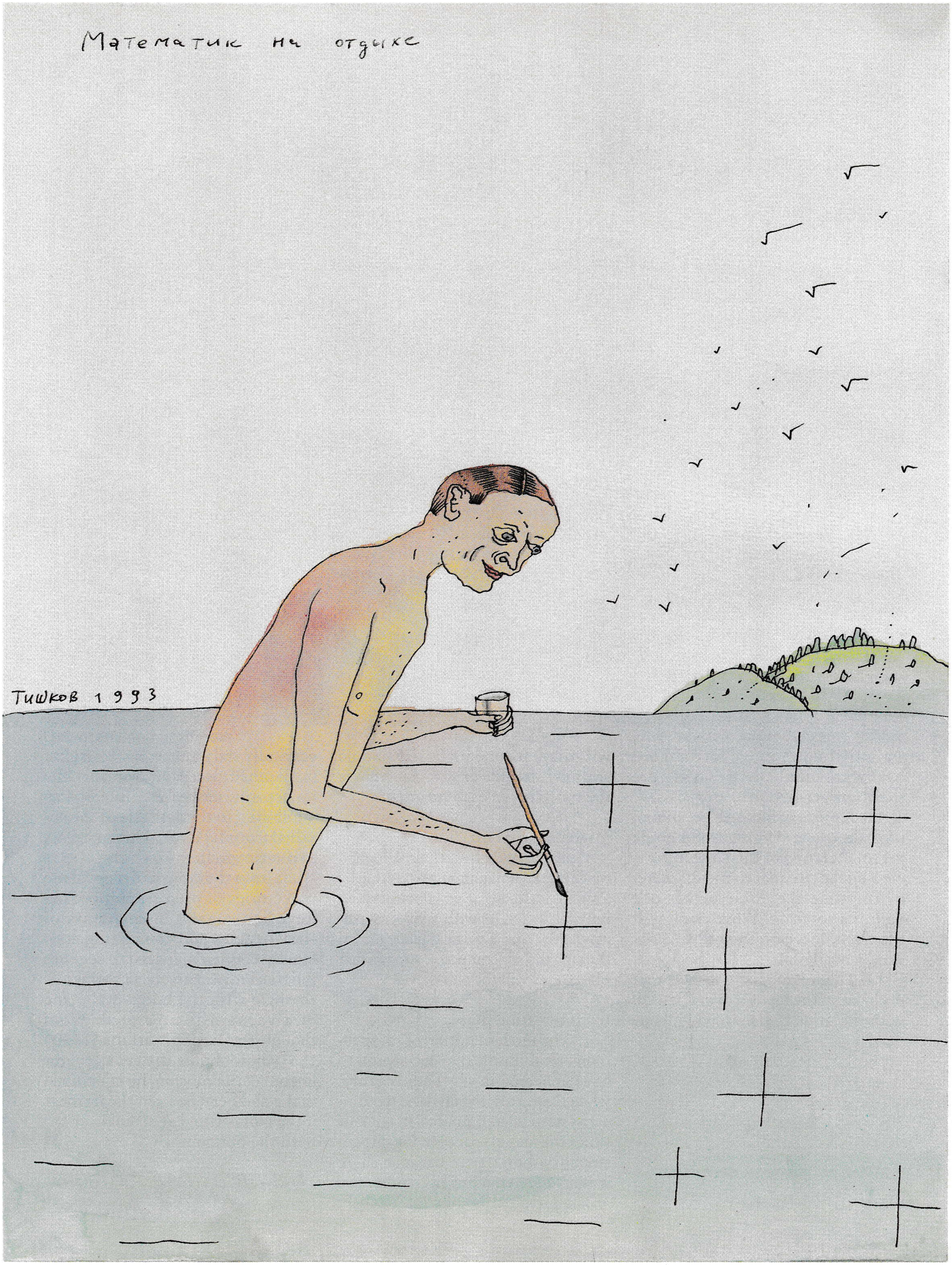
It's interesting that . . .

. . . the modern notion of three-dimensional physical space seems to have appeared in the 17th century, when Descartes invented the rectilinear system of coordinates. In ancient times the idea of space having dimensionality didn't arise because there was no understanding of coordinates.

. . . physical bodies fall strictly vertically only at the Earth's poles. In all other places on the planet the trajectories of freely falling bodies are shifted to the east due to the so-called Coriolis force, which appears in rotating systems.

. . . from the time of Aristotle the trajectories of projectiles were believed to consist of segments of straight lines and connecting arcs. Finally Galileo managed to see that the trajectory of an object thrown at an angle with the horizontal plane in a vacuum is a parabola. And though the Italian Tartalia (1500–1557) didn't know the laws governing projectile motion, he concluded that a shell can be fired the farthest if the gun is tilted at an angle of 45° with the horizontal.

ANSWERS, HINTS & SOLUTIONS ON PAGE 59


Математик на отдыхе
ТИШКОВ 1993

Some things never change

So use invariability to your advantage

by Yury Ionin and Lev Kurlyandchik

THE PROBLEMS DISCUSSED in this article are alike in that each of them involves (perhaps with an appropriate reduction) some "configurations" of numbers or other symbols and a set of operations that can be applied to these configurations. We'll ask such questions as these: Is it possible to find a sequence of operations that turns one given configuration into another? Under what conditions restricting the configurations and the sequence of operations can this be achieved? What configurations can be obtained from a given one?

We'll see that problems like these are most efficiently solved by finding unchangeable features in a changing configuration as it is subjected to the allowed transformations.

Problem 1. *Ten plus signs and fifteen minus signs are written on the blackboard. You can erase any two signs and write in their place a plus sign if they were the same and a minus if they weren't. This operation is repeated 24 times. What sign remains on the blackboard?*

Let's replace every plus sign with the number 1 and every minus sign with −1. Then the allowed operations can be described as erasing any two numbers and writing down their product. So the product of all the numbers on the blackboard remains unchanged (because multiplication is both commutative and associative). This product was originally −1, so finally it must also be equal to −1—that is, the last number will be −1, or the last sign will be a minus.

The same reasoning can be reshaped as follows. Let's replace all the pluses by zeros, minuses by ones, and note that the sum of any two numbers is of the same parity as the number written down in their place when they are erased. Since initially the sum of all the numbers was odd (it was equal to 15), the last number left on the blackboard has to be odd—that is, 1; so the sign left on the blackboard is a minus.

Finally, a third solution can be based on the observation that under every operation the number of minuses either doesn't change or decreases by two. Initially the number of minuses was odd, so a minus will be left in the end.

Now let's examine all three solutions.

The first one was based on the invariability of the product of the numbers on the blackboard, the second on the invariability of the parity of their sum, and the third on the invariability of the parity of the number of minus signs. We can say that in each of these solutions we were able to find an *invariant*: the product of ones and negative ones, the parity of the sum of zeros and ones, the parity of the number of minuses. The solutions to the problems and exercises that follow are also based on aptly selected invariants.

Exercise 1. A number of plus and minus signs are written on the blackboard. It is permitted to erase any two signs and write a plus instead if they were different and a minus if they were the same. Prove that the last sign that will be left doesn't depend on the order of erasure.

Problem 2. *Plus and minus signs are arranged in a 4 × 4 table as shown in figure 1. It is permissible to reverse all the signs in one horizontal line, one column, or along any line parallel to a diagonal of the table (in particular, in any corner unit square). Does there exist a sequence of these operations leading to a table without minus signs?*

Replace the pluses and minuses with +1's and −1's again. Multiply the numbers in the squares shaded

+	+	−	+
+	+	+	+
+	+	+	+
+	+	+	+

Figure 1

Art by Leonid Tishkov

At left: "A mathematician on vacation."

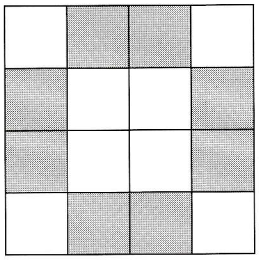


Figure 2

in figure 2. This product is an invariant, because any of our operations either leaves these numbers intact or changes exactly two of them. Originally this product is equal to –1, so it's impossible to get a table without minuses, for which this product is +1.

Exercise 2. Solve problem 2 for the tables in figures 3, 4, and 5.

Problem 3. *Written on the blackboard are a number of zeros, ones, and twos. Instead of any two different digits we can write a digit different from these two (0 and 1 can be replaced by 2, and so on). These operations are repeated until only one figure remains on the blackboard. Prove that this figure doesn't depend on the order in which these operations are performed.*

Let x_0, x_1, and x_2 be the numbers of zeros, ones, and twos, respectively. Every operation changes each of these numbers by one, and so it changes the parities of all three numbers. When there's only one figure left, two of the numbers x_0, x_1, and x_2 become equal to 0, and the third is 1. Therefore, initially two of these numbers were of one parity and the third was of the other parity. So regardless of the order of the operations, the only number among x_0, x_1, and x_2 that can be equal to 1 in the end is the one that initially differed in parity from the other two.

Our solution implies that in the case when all three numbers x_0, x_1, x_2 are of the same parity it's impossible to erase all figures but one. However, our solution does not imply that this can really be done if there are both odd and even numbers among them (and at least two of the initial numbers are nonzero). In fact, such a sequence of operations can always be found in this case—we leave this fact as a rather simple exercise for the reader.

Let's change the operations in problem 3: we'll require that four figures are erased at a time—two of one kind and two of another—and that one figure of the third kind is written down in their place. Suppose that after a number of such operations only one figure is left. Can we tell in advance, knowing the initial numbers of zeros, ones, and twos, what figure will remain on the blackboard?

The parity argument doesn't work here, because one of the numbers x_0, x_1, x_2 changes its parity under each operation, while the parities of the other two are preserved, so numbers with initially different parities can get the same parities. But if we consider the remainders of x_0, x_1, x_2 modulo 3 rather than modulo 2 (looking at parities is, after all, equivalent to reducing the numbers modulo 2), we notice that our operations leave them equal if they were equal and different whenever they were different initially. (In other words, the remainders of $x_1 - x_0$, $x_2 - x_1$, and $x_0 - x_2$ modulo 3 are invariant.) The rest of the reasoning follows that of the solution to problem 3.

Problem 4. *In each square of an 8×8 array an integer is written. We can choose an arbitrary 3×3 or 4×4 subarray and increase all the numbers in it by one. Is it always possible to obtain numbers divisible by 3 in all squares of the initial array?*

The answer is no. Let's find the sum of the numbers in the 48 squares shaded in figure 6. Since any 4×4 square contains 12 shaded unit squares and any 3×3 square contains 9 or 6 such small squares, the allowed operations do not change the remainder of this sum when divided by 3. Therefore, if this sum is not divisible by 3 initially, the shaded squares always contain numbers not divisible by 3.

Exercise 3. Given the conditions of problem 4, is it possible to obtain an array not containing even numbers from an arbitrary initial array?

Problem 5. *The numbers 1, 2, ..., n are arranged in some order. We can exchange any two adjacent numbers. Prove that an odd number of such exchanges produces an arrangement necessarily different from the initial one.*

Let $a_1, a_2, \ldots, a_n$ be the numbers 1, 2, ..., n written in the given order. Such a string of numbers is called a *permutation* of 1, 2, ..., n. The numbers a_i and a_j in this permutation are said to form an *inversion* if $i < j$ but $a_i > a_j$—that is, the greater of these two numbers precedes the smaller. Exchanging any two adjacent numbers, we reverse their order, but the order of any other pair of numbers is, of course, preserved. So the number of inversions increases or decreases by one. After an odd number of pair exchanges we change the parity of the number of inversions and, therefore, the permutation as well.

Exercise 4. Prove that the statement of problem 5 remains valid

+	+	−	+	+	+
+	+	+	+	+	+
+	+	+	+	+	+
+	+	+	+	+	+
+	+	+	+	+	+
+	+	+	+	+	+

Figure 3

+	+	+	+	+	+
+	+	+	+	+	+
+	+	−	+	+	+
+	+	+	+	+	+
+	+	+	+	+	+
+	+	+	−	+	+

Figure 4

+	+	+	+	+	+
+	+	+	+	+	+
+	+	−	−	+	+
+	+	+	+	+	+
+	+	+	+	+	+
+	+	−	−	+	+

Figure 5

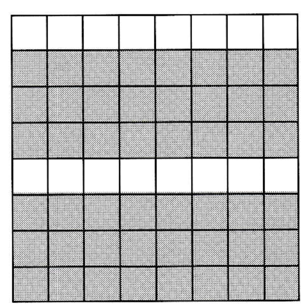


Figure 6

even if we're allowed to swap *any* two numbers in the given permutation. (Hint: show that any two numbers can be exchanged by means of an odd number of "adjacent exchanges.")

Another term for a pair exchange is *transposition*. Using this term, we can formulate the statement of exercise 4 as follows: *an odd number of transpositions changes a permutation.* The solution to problem 5, along with the fact stated in the hint to exercise 4, shows that every transposition changes the parity of the number of inversions. A permutation is called *even* or *odd* if the number of inversions in this permutation is even or odd, respectively. So we can now say that performing a transposition changes the parity of the permutation whose elements are transposed.

Problem 6. *Twenty-five cars started from different points along a closed speedway in the same direction and at the same time. According to the rules of the race, the cars can pass one another, but double passing is forbidden. The cars finished simultaneously, all at their respective starting positions. Prove that there was an even number of passes during the race.*

Imagine that one of the cars is painted yellow and number the remaining cars in their order at the start (car number one starts immediately behind the yellow car, the second one behind the first, and so on). Imagine there is a scoreboard indicating the order of the cars as they follow the yellow car. Then every time a numbered car overtakes another numbered car, two numbers on the scoreboard exchange places.

Let's see what happens when a numbered car passes the yellow one. If the order of the numbers before the pass was $a_1, a_2, \ldots, a_{24}$, then after the pass the scoreboard will read $a_2, a_3, \ldots, a_{24}, a_1$. But we can obtain the same permutation as a result of 23 transpositions:

$$a_1, a_2, a_3, \ldots, a_{24} \to a_2, a_1, a_3, \ldots, a_{24} \to a_2, a_3, a_1, \ldots, a_{24} \to \ldots \to a_2, a_3, \ldots, a_1, a_{24} \to a_2, a_3, \ldots, a_{24}, a_1.$$

If the yellow car passes another one, the permutation $a_1, a_2, \ldots, a_{24}$ turns into $a_{24}, a_1, a_2, \ldots, a_{23}$. This can also be achieved by a series of 23 transpositions (how?).

Thus, *any* pass amounts to an odd number of transpositions—that is, changes the parity of the indicated permutation. But the final permutation coincides with the initial one, so, by exercise 4, the total number of transpositions must be even (in other words, the overall permutation must be even). Therefore, the number of passes had to be even as well.

With this introduction, you're on your own. We're confident you can handle the exercises that follow.

Exercises

5. Four ones and five zeros are written around a circle in an arbitrary order. A one is inscribed between every two equal numbers and a zero between different numbers; then the initial numbers are erased. Is it possible to obtain a set of nine zeros after a series of these operations?

6. Wendy tore up a sheet of paper into 10 pieces, then tore some of these into 10 pieces, and so on. Could she obtain 1993 pieces in this way?

7. The numbers 1, 2, ..., 1993 are written on the blackboard. Two numbers are erased and replaced by the remainder of their sum when divided by 13. This operation is repeated until one number is left. What is this number?

8. Every number from 1 to 1,000,000 is replaced by the sum of its digits. The resulting numbers are repeatedly subjected to the same operation until all the numbers have one digit. Will the number of ones in the end be greater or less than the number of twos?

9. The sum of digits of a three-digit number is subtracted from the number. The same is done to this difference, and so on. What number is obtained after a hundred repetitions?

10. A circle is divided into 10 sectors and one chip is placed in each. We can move any two chips to the neighboring sectors but they must move in opposite directions. Is it possible to bring all the chips together in one sector?

11. (a) A minus sign is placed at the vertex A_{12} of a regular 12-gon $A_1A_2\ldots A_{12}$, and plus signs are placed at all other vertices. One is allowed to reverse signs at any three vertices forming an isosceles but not a right triangle. Can we get a minus sign at A_1 and plus signs at all other vertices after a number of such operations?

(b) Will the answer to part (a) remain true if we're allowed to change signs at the vertices of *any* isosceles triangle?

12. A plus or minus sign is written in every square of a 4×4 array. We can change all the signs in any line or any column. The smallest number of minus signs that can be arrived at via these operations starting with a given array is called the *character* of this array. What values can the character take?

13. Thirty chips—ten white and twenty black—are placed around a circle. Any two chips with three chips between them can be swapped. Two arrangements of chips are *equivalent* if one of them can be obtained from the other after a number of such transpositions. What is the greatest possible number of nonequivalent arrangements?

14. The numbers 1, 2, ..., 1993 are written in increasing order. Any four numbers can be rearranged in reverse order in the same places. Is it possible to obtain the reverse order 1993, 1992, ..., 2, 1 of the entire set of numbers?

ANSWERS, HINTS & SOLUTIONS ON PAGE 60

Two more problems involving invariants can be found among the math challenges in this issue. See also the Toy Store in this issue—it's devoted to mathematical puzzles that are closely related to invariants.


Tom Bunk

PHYSICS CONTEST

Thrills by design

"Centrifugal power . . . what stillnesses lie at your center resting among motion?"
—Muriel Rukeyser (1913–1980)

by Arthur Eisenkraft and Larry D. Kirkpatrick

AT THE XXIV INTERNATIONAL Physics Olympiad, which was held in Williamsburg, Virginia, during July 1993, students from the 41 participating countries spent a day investigating the physics of some of the amusement park rides at Busch Gardens. Those roller coasters and bumper cars and swings certainly move the awareness of momentum, forces, and acceleration from our brains to our guts.

One ride that appears simple enough to analyze is the rotor. As in many physics explorations, the analysis reveals a hidden effect—a treasure that you may not have previously known. The rotor is a hollow cylinder of radius 2.5 m. Riders stand inside the cylinder with their backs against the wall. As the rotor spins, they feel as if they are being pushed against the wall. When the maximum speed is reached, the floor drops out! As shouts emerge, the riders don't fall. The friction between the wall and the riders keeps them from slipping down.

Let's analyze this ride, keeping our minds ready to discover that extra treasure. From the perspective of a person above the ride, a person is kept from flying through the wall by the push of the wall. This normal force $\mathbf{F}_N$ is directed toward the center of the cylinder. The force of gravity $\mathbf{F}_g$ pulls the person vertically down. The frictional force $\mathbf{F}_f$ is directed vertically up and must be equal to $\mathbf{F}_g$ or the person will slide down.

The horizontal normal force $\mathbf{F}_N$ must supply the required centripetal force to keep the rider moving in a circle:

$$F_N = \frac{mv^2}{R},$$

where m is the mass of the rider, v is the velocity of the rider, and R is the radius of the rotor. The equality of $\mathbf{F}_g$ and $\mathbf{F}_f$ yields the following relationship:

$$mg = \mu F_N,$$

where μ is the coefficient of friction. Substituting for F_N, we can find the required coefficient of friction:

$$\mu = \frac{gR}{v^2}.$$

The coefficient of friction determines the required minimal speed for any rotor! Next time you get to watch or ride on a rotor, take a look at how the ride's designers have increased the coefficient of friction—did they add carpeting to the walls, or did they use rough paint?

Some of you are probably still searching for the surprise discoveries. One minor surprise is that the ride works just as well irrespective of the mass of the rider. The more interesting surprise is that, from the reference frame of the rider, the question "Which way is up?" takes on new meaning. The rider feels gravity pulling one way and a centrifugal force pulling outward. The combination of the two defines a "new gravity" in the rider's balance system. Riders think that they are lying at an angle and are no longer vertical. Next time you're on the rotor, try to be aware of this effect. Estimate the angle of apparent tilt, and check to see if it's consistent with your estimates of $\mathbf{F}_N$ and $\mathbf{F}_g$.

One of the challenges for the engineers working for amusement parks is to develop new, exciting, and safe alternatives to the tried-and-true classics. We thought that *Quantum* readers might enjoy such a design challenge. We'll describe a traditional physics problem—one that was used in the International Physics Olympiad in Budapest, Hungary, in 1976. Your challenge is to understand the physics of the problem and then to incorporate the design into an amusement park ride.

A hollow sphere of radius $R = 0.5$ m rotates about its vertical diameter with an angular velocity $\omega = 5$ s^{-1}. Inside the sphere at the height $R/2$ a small block revolves together with the sphere. (Use $g = 10$ m/s^2.)

Art by Tomas Bunk

A. What is the coefficient of friction required for the block to continue to revolve at this height?

B. What is the coefficient of friction required when $\omega = 8\ s^{-1}$?

C. Investigate the problem of stability (1) for small variations in the position of the block and (2) for small variations in the angular velocity of the sphere.

D. Can the block be replaced by a person as a design for an amusement park ride? Are there any inherent problems with such a ride? Would the public enjoy such a ride? How would you get on and off such a ride?

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

How about a date?

In the March/April issue we asked you to explore the problems created in radiocarbon dating when the atmospheric concentration of carbon 14 varies with time. An excellent solution was submitted by Ben Davenport from the North Carolina School of Science and Mathematics. Ben was a semifinalist for this year's US Physics Team that competed in the International Physics Olympiad in Williamsburg.

A. We begin by calculating the value of the decay constant:

$$\lambda = \frac{\ln 2}{T_{1/2}} = 1.21 \cdot 10^{-4}\ \text{year}^{-1} = 3.83 \cdot 10^{-12}\ \text{s}^{-1}.$$

We can calculate the number N_0 of ^{14}C atoms from the ratio of the ^{14}C to ^{12}C atoms and the fact that 12 g of carbon contain Avogadro's number of atoms:

$$N_0 = \frac{(6.02 \cdot 10^{23}\ \text{atoms})(1.30 \cdot 10^{-12})}{12} = 6.52 \cdot 10^{10}\ \text{atoms}.$$

Therefore, the decay rate for the 1-g sample of carbon shortly after the animal died was

$$R_0 = \lambda N_0 = 0.250\ \text{decays/s} = 15\ \text{decays/min}.$$

B. Solving the equation for the change in the decay rate for t, we can obtain the time since the animal died:

$$t = \frac{\ln(R/R_0)}{-\lambda} = \frac{\ln(1/15)}{-\lambda} = 7.07 \cdot 10^{11}\ \text{s} = 22{,}400\ \text{years},$$

where R is the current decay rate of 1 decay/min.

C. Unfortunately there was an error in the statement of the problem that caused this part of the problem to be harder than was intended. The rate of decrease was meant to be 0.1% rather than the stated 1%. Let's solve the problem with both rates of decrease.

With a decrease of 0.1% per century, we can use an iterative technique to find the age. As a first approximation to the decay rate R_0' at the time of the animal's death, let's use our age from part B. Then

$$R_0' = R_0\left(1 - \frac{0.001t}{100\ \text{years}}\right) = 0.776(15\ \text{decays/min}) = 11.6\ \text{decays/min}$$

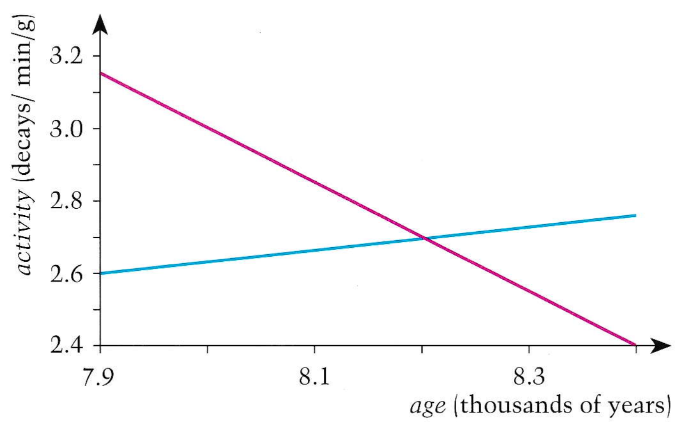


Figure 1
The age assuming a 1% decrease per century.

and

$$t = \frac{\ln\left(R/R_0'\right)}{-\lambda} = \frac{\ln(1/11.6)}{-\lambda} = 6.40 \cdot 10^{11}\ \text{s} = 20{,}300\ \text{years}.$$

If we iterate this procedure two more times, we settle in on an age of 20,500 years.

This technique will not work for the stated decrease of 1%, because the ratio drops to zero in 10,000 years. However, we can use a graphical technique to solve for the answer. If we let t represent time in the past, then the decay rate at any time in the past must be given by

$$R_0' = Re^{\lambda t}.$$

The historical decay rate of a new 1-g sample of carbon is given by

$$R_0' = R_0\left(1 - 0.01\frac{t}{100\ \text{years}}\right).$$

The graph of these two functions is shown in figure 1. Notice that the two functions cross at an age of 8,200 years.

D. This same graphical technique works very well for the sinusoidal dependence, where the historical decay rate is given by

$$R_0' = R_0\left[1 \pm 0.05\sin\left(\frac{2\pi t}{628\ \text{years}}\right)\right],$$

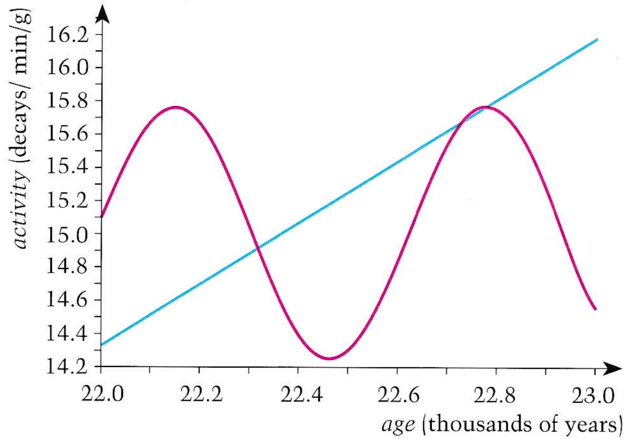


Figure 2
The ages assuming an increasing ratio at present.

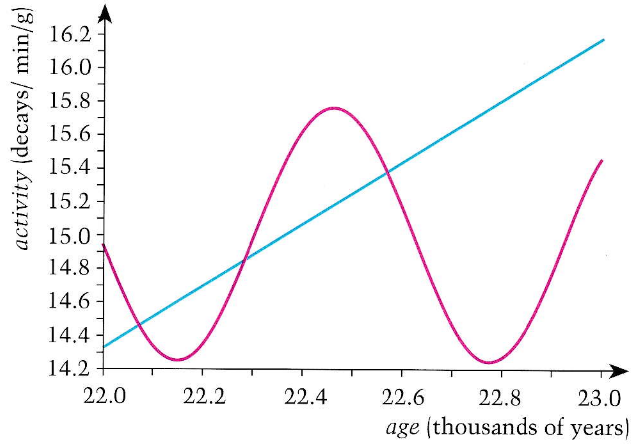


Figure 3
The ages assuming a decreasing ratio at present.

and we have used $\cos(t - \pi/2) = \sin t$ to make R_0' have the current value when $t = 0$—that is, at the current time. We get two different results (figures 2 and 3) depending on whether we assume that the ratio of the two carbon isotopes is increasing or decreasing at the current time. We see that in both cases we obtain three possible ages for the sample. About the best that we can say is that the age is 22,500 ± 300 years if the ratio is increasing and 22,300 ± 300 years if the ratio is decreasing. Ben observed that our solution to part A is equivalent to using the average atmospheric concentration and lies within both of these ranges. ◘

"HOME ON THE RANGE" CONTINUED FROM PAGE 16

(a) there must be three persons in each boat; (b) there must be no more than four and no less than two persons in each boat; (c) there must be at least one person in each boat? (The boats are numbered 1 through 3.)

17*. If the hosts have enough chairs, it's customary not to seat more than one guest on a chair: the set of guests is invertibly mapped into the set of chairs. If there are six chairs in the room, how many ways are there to seat (a) one, (b) two, (c) three, (d) four, (e) five, (f) six guests?

18*. Invertible mappings of a finite set M into another finite set N are called *permutations* in combinatorics. The number of such mappings depends only on the numbers m and n of elements in M and N, respectively, and is denoted by A_n^m.[6] Show that

$$A_1^1 = 1,\ A_2^1 = A_2^2 = 2,\ A_3^1 = 3,$$
$$A_3^2 = A_3^3 = 6,\ A_{10}^2 = 90,$$

and establish a general rule for calculating A_n^m. Prove that $A_n^{n-1} = A_n^n$ for all n.

19*. Problem 16(c) can be formulated in the abstract: how many mappings of a set of 9 elements onto a set of 3 elements are there? Denote by D_n^m the number of mappings of a set of n elements onto a set of m elements. Verify that

$$D_3^2 = 6,\ D_4^2 = 12,\ D_4^3 = 36,\ D_n^n = n!$$

and try to state the general rule for computing D_n^m (this is a somewhat more difficult problem than problems 8, 11, and 18).

20*. Find the number of functions defined on a set of 28 elements and taking the four values P, K, S, and V seven times each. (This is a problem about the number of ways to distribute housekeeping chores among Petya, Kolya, Sasha, and Volodya—see example 3.) ◘

[6]The elements of M can be thought of as the numbers 1, 2, ..., m; then any mapping in question assigns a certain order to their m images in N: the image of 1 can be considered the first of these m images, the image of 2 the second, and so on. Thus, a permutation of m elements chosen from n is defined. The notation A_n^m stems from the French word *arrangements*.—*Ed.*

AT THE BLACKBOARD

An ideal gas gets real

And relativity pays a call on electromagnetic induction

ASK ANYONE WHAT HAPPENS to the temperature of an ideal gas expanding in a closed vessel without heat exchange with the environment. Almost everyone is sure to answer that the gas will get cooler. Don't you believe it! It ain't necessarily so.

Let's imagine a little experiment. Let one part of the heat-isolated vessel be occupied by an ideal gas with a pressure p_1 and temperature T_1, and let the other part be empty (fig. 1). At some point we remove the barrier between the parts of the vessel. Naturally the gas will expand, making its way into the vacuum. After its molecules have collided repeatedly with the walls and with one another, an equilibrium state is established. It's clear that now the volume of gas has doubled: $V_2 = 2V_1$. What are its pressure p_2 and temperature T_2?

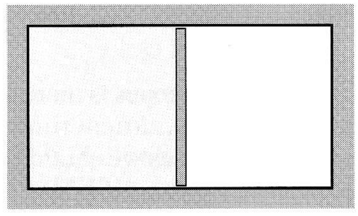


Figure 1

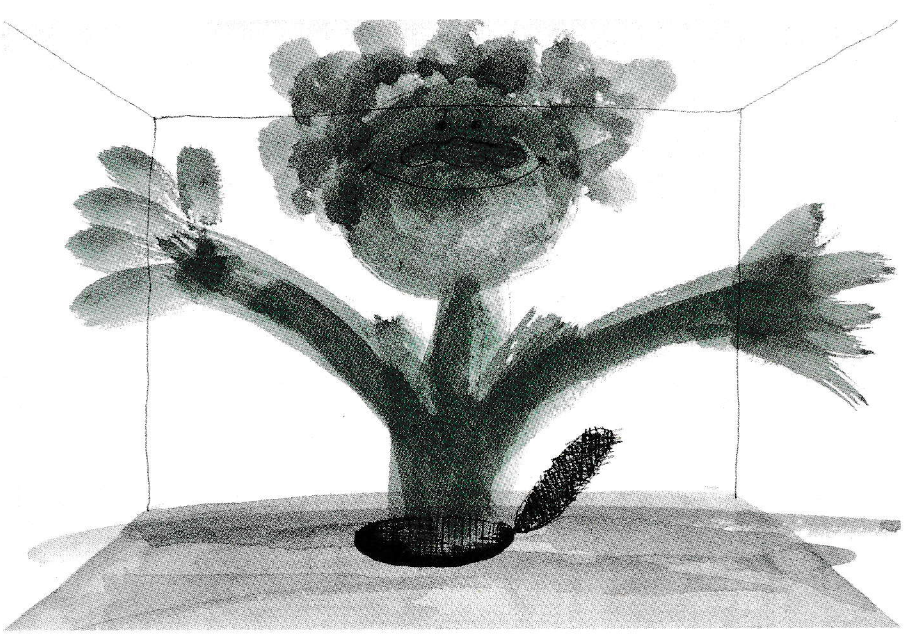
On the one hand, since the process is adiabatic, the points corresponding to the initial and final states of the gas are on the adiabatic curve 1–2′ (fig. 2). This curve, as you probably know, decreases more abruptly than the isotherm, so the temperature of the gas must decrease: $T_2' < T_1$.

On the other hand, let's see what the first law of thermodynamics tells us. A quantity of heat Q added to a gas increases its internal energy U and allows the gas to do the work W associated with the expansion:

$$Q = U + W.$$

In our case $Q = 0$ (because we've made the process adiabatic). What kind of work is performed by the gas? No work at all, since it expands into a vacuum. The gas meets no resistance from that empty space. So both the force and the work are equal to zero: $W = 0$. Thus, the change in the internal energy is also equal to zero: $U = 0$. However, since in the case of an ideal gas the internal energy depends on temperature only, the temperature doesn't change: $T_2 = T_1$; the pressure is equal to $p_2 = p_1/2$. This means that the points corresponding to the initial and final states are on the isotherm 1–2.

What happens between these two states? Unfortunately, the thermodynamics you learn in school tells you nothing about that. Why? Because it holds only for very slow (quasi-static) processes that occur at rates much slower than the thermal velocity of molecular motion. In our case, as soon as we remove the barrier, the gas will simply rush into the vacuum with a velocity of the



Art by Pavel Chernusky

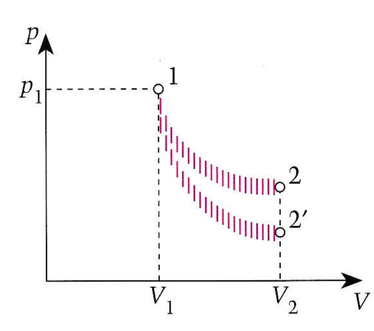


Figure 2

order of the thermal velocity of the molecules or even faster, since there are individual molecules in the gas whose velocity is much higher than the thermal velocity. Here thermodynamics is just plain wrong. That's why in figure 2 we depicted this unfamiliar process with slashes instead of solid lines.

All our reasoning holds for an ideal gas. But suppose the gas isn't ideal. Then its molecules interact with one another, and the internal energy of the gas consists of the kinetic energy of its molecules and the potential energy of their interaction.

Figure 3 describes the dependence of the potential energy U for the interaction of two molecules on the distance r between them. Where the potential energy is minimal (the point r_0), the substance condenses—that is, turns into a liquid. According to our conditions we have a gas at the outset, so the average distance between molecules corresponds to the point $r_1 >> r_0$. After the volume doubles, the average distance between molecules is equal to $r_2 = r_1\sqrt[3]{2} > r_1$. During expansion the gas has been pulled slightly upward along the slope of the potential well. Who worked to increase the potential energy by ΔU? No one did. And neither did the gas. So we're forced to admit that the increase in the potential energy is due to a decrease in the kinetic energy of the moving molecules. This means that the temperature—which is a measure of the average kinetic energy of gas molecules—decreases slightly as a result of the expansion. But this holds only for a real gas.

—Albert Stasenko

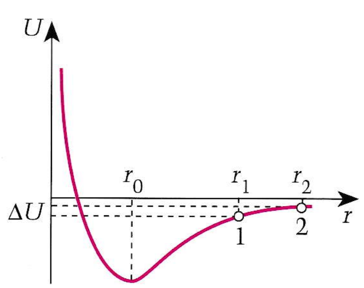


Figure 3

Michael, meet Albert

In 1831 Michael Faraday discovered the phenomenon of electromagnetic induction. He found that a change in the magnetic flux through any surface bounded by a closed loop causes an electric current to arise there. Faraday's experiments proved convincingly that the strength of the induced current, or the electromotive force (emf), doesn't depend on what caused the magnetic flux to change. We can change the external magnetic field, leaving the circuit stationary—here we must either move the source of the field (the coil or magnet) or change the current in the coil creating the field (for instance, by opening and closing a switch, as Faraday did). But we can take a completely different approach: without changing the magnetic field, we can achieve a change in the magnetic flux by moving the loop itself or by changing its shape (which is what happens, for instance, in a generator, where the induced emf arises in a wire frame as it rotates in a constant magnetic field). In either case, the induced emf is proportional to the rate of change in the magnetic flux (Faraday's law), while its direction is defined by Lenz's law.

Faraday himself thought it quite natural that both variations are described by the same law. However, careful analysis shows that this situation is far from obvious. Let's take a closer look at this question.

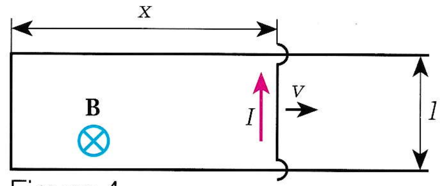
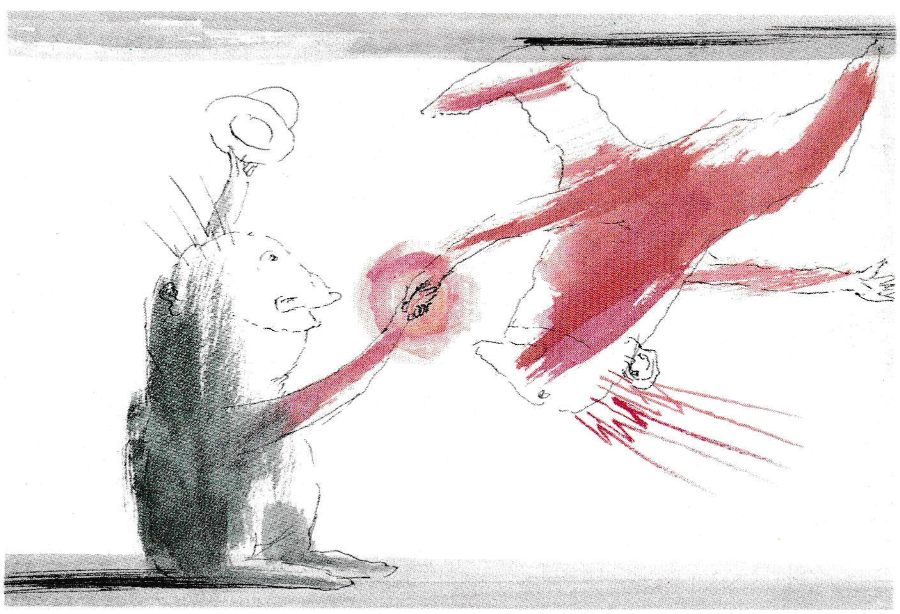


Figure 4

When the loop is moved in a magnetic field that does not change in time, the Lorentz force, which acts on any moving charge, plays the role of an extraneous force creating current in the circuit. Here's an example you may recall from your textbook. Let a rectangular loop be placed in a homogeneous magnetic field **B** with its plane perpendicular to the field (fig. 4). One of the sides of the circuit is a wire with sliding contacts. If we move the wire with a speed v, we induce an induction $\mathcal{E} = Bvl$ (where l is the length of the wire). Indeed, the Lorentz force $F = qvB$ (directed as shown by the arrow) acts on the free charge q in the moving wire. The emf corresponding to this magnetic force is

$$\mathcal{E} = \frac{W}{q} = \frac{Fl}{q} = \frac{qvBl}{q} = Bvl,$$

where W is the work of the force over the length of the wire. Compare this result with the rate of change in the magnetic flux:



$$\left|\frac{\Delta\Phi}{\Delta t}\right| = \left|\frac{\Delta(BS)}{\Delta t}\right| = \left|\frac{\Delta(Bxl)}{\Delta t}\right| = Bvl.$$

We can easily verify that the sign of the emf obeys Lenz's law.

The example above shows that when conductors move in a constant magnetic field, the generation of the induced current is not a fundamentally new physical phenomenon.

We get a completely different picture with a stationary loop placed in a magnetic field that changes with time. Since free charges in the conductor are initially stationary (of course, we don't take random thermal motion into account), a magnetic field does not act on them and, thus, cannot cause them to move in a certain direction. As a result, the induced current can arise only under the influence of an electric field. How does the electric field arise and what properties does it have?

It's clear that this electric field is very different from the well-known electrostatic field. For example, it creates an emf in a closed loop. This means that the work it performs in moving charges along the closed path is not equal to zero. It is a vortex field—that is, its lines of force take the form of closed lines. There are other differences as well.

We see that an analysis of the situation arising with a stationary loop in a variable magnetic field results in a whole range of new physical phenomena that show a direct interrelationship between electric and magnetic fields. In 1860 James Clerk Maxwell came to the conclusion that a magnetic field that varies with time always generates an electric field. Following his inner sense of the symmetry of physical laws, Maxwell further postulated a fact that wasn't supported at the time by any experiments: an electric field that varies in time always generates, in turn, a magnetic field. Giving these statements a mathematically symmetrical form (now known as Maxwell's equations), he completed his construction of a unified theory of the electromagnetic field.

After all these statements it might appear that the law of electromagnetic induction describes two quite different physical phenomena. In a constant magnetic field, the induced current in a moving loop is caused by the magnetic field itself. A variable magnetic field generates an electric field that causes charges to move along a stationary loop. Then why is the law of electromagnetic induction the same for these two cases? Is it just a striking coincidence? Actually, this "coincidence" points up the profound link between the theory of the electromagnetic field and the special theory of relativity. This theory is based on the principle of relativity developed by Albert Einstein, who formulated the special theory of relativity in 1905. According to this principle *all* phenomena in nature must occur the same way in all inertial reference systems. An important consequence of this principle is the fact that it is impossible unambiguously, outside of a dependence on the reference system, to say what fields exist in the surrounding space.

By way of example, let's look at the interaction of two electrons from the viewpoint of two observers (fig. 5). Observer A asserts that the moving charges create around themselves both an electric and a magnetic field and that, in addition to the Coulomb repulsion, there is a magnetic attraction between them (as between parallel currents). Observer B doesn't agree with observer A and asserts that there is no magnetic field, since the charges are at rest, and only the Coulomb repulsion operates between the electrons. The principle of relativity, however, soon brings the observers into agreement, asserting that they're both right, since the notions of the electric and magnetic fields are relative—they depend on the reference system. Both these fields are parts of a single whole—the electromagnetic field.

Now let's get back to Faraday's law and imagine the following experiment. Let's bring a permanent magnet near a closed conducting loop (fig. 6). A galvanometer connected to the loop shows the induced current. A deflection of the galvanometer arrow can be seen by both observer A on the magnet and observer B on the loop. "It's clear as crystal," observer A says. "The loop is moving in the constant field of my magnet, the Lorentz force operates on the charges in the loop, and that's why there's a current!" Observer B says the opposite: "The magnet is approaching my stationary loop. A variable magnetic field generates an electric field, and it's this field that generates current in the loop!" According to Einstein, however, both observers are right. Both see the same deflection of the galvanometer needle. This simply means that Faraday's law must be identical for both cases.

—Alexey Chernoutsan

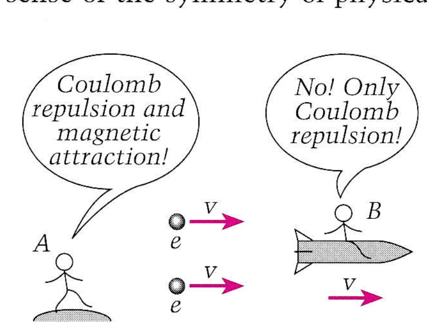


Figure 5

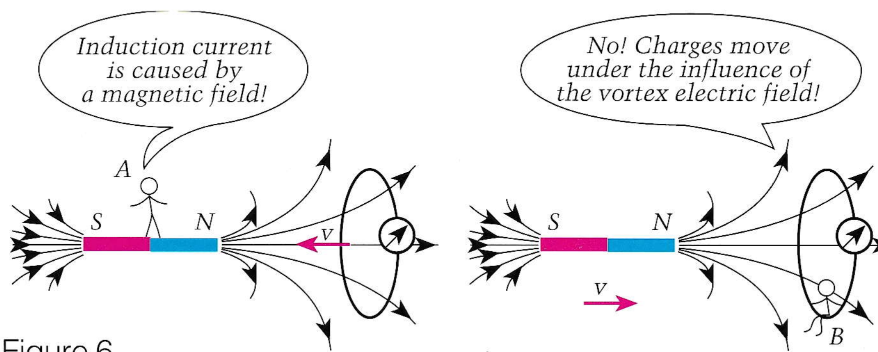


Figure 6

FOLLOW-UP

Martin Gardner's "Royal Problem"

The playing field is level, but is it also square?

by Jesse Chan, Peter Laffin, and Da Li

THIS IS A GENERALIZATION of a problem originally posed by Martin Gardner (see "A Royal Problem," coauthored by Andy Liu, in the last issue of *Quantum*). The problem goes like this: "On an m by n chessboard, $m \geq n \geq 3$, a White Queen is on square $(1, 1)$, while a Red King is on square (m, n). The Queen moves first unless $m = n$, in which case the King must move out of check. Thereafter, moves alternate. The Queen wins if and only if the King is forced to her initial square $(1, 1)$ in a finite number of moves. With perfect play, which royalty wins?"

We claim that the King wins if and only if $m = n$.

The reader doesn't have to know chess beyond the moves of the Queen and King. If the King can move to where the Queen is, he can capture her and win the game. The King is said to be in check if the Queen can capture him on her turn. The King is not allowed to move into or remain in check.

If the King on his turn has no legal move but is not in check, the game ends in what is called a stalemate. In our game, this can only happen if he is at a corner other than $(1, 1)$, and it is a victory for the King.

In our analysis, all positions are considered at the moment when it is the King's turn to move. We first prove that the King wins if $m = n \geq 3$. We consider the game from his point of view and define a "forbidden zone" into which he must not move. This zone consists of all squares (i, j) where $i + j \leq n - 1$. It always contains the forbidden square $(1, 1)$. For $n = 3$, it consists only of this square. The case $n = 4$ is shown in figure 1, with the forbidden zone shaded.

We'll prove that not only can the King avoid going to $(1, 1)$, he can't even be forced into the forbidden zone by the Queen. For this to happen, the King must be on one of the following types of squares:

A. $(n - 1, 1)$ or $(1, n - 1)$;
B. (i, j), where $i + j = n$, with $i > 1$ and $j > 1$;
C. (i, j), where $i + j = n + 1$, with $i > 1$ and $j > 1$.

For $n = 4$, the relevant squares are marked accordingly in figure 1.

Consider case A. Suppose the King is on $(n - 1, 1)$ as shown in figure 2. He can move to $(n - 2, 2)$, $(n - 1, 2)$, $(n, 1)$, or $(n, 2)$, none of which is in the forbidden zone. These moves are marked with x's in figure 2.

The only squares from which the Queen can control all four squares are $(n - 1, 1)$, $(n - 1, 2)$, and $(n, 2)$. The first is already occupied by the King. If the Queen is on either of the other two squares, she will be captured. So the King can't be forced into the forbidden zone when he is on $(n - 1, 1)$ or, by symmetry, on $(1, n - 1)$.

Cases B and C can be handled similarly, the King having even more options. This completes the proof that the King wins if $m = n \geq 3$.

While this proof is very simple, one may well ask how we came to think of the forbidden zone in the first place. Our initial approach is by mathematical induction on n. It's not difficult to see that the King wins if $n = 3$.

For $n = 4$, we mark off two overlapping 3×3 boards on the 4×4

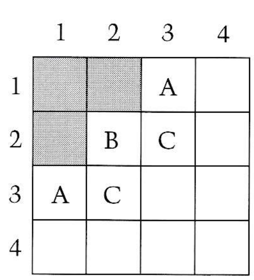


Figure 1

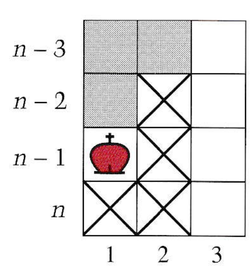


Figure 2

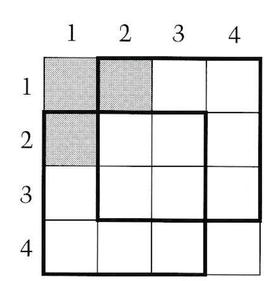


Figure 3

board, as shown in figure 3. Each smaller board has its own forbidden square, and the two join up with the actual forbidden square to form the forbidden zone in figure 1.

We now consider two cases. If the King and Queen are on the same 3×3 board, we already know that the King has a safe square within the same board. If the King and Queen aren't on the same 3×3 board, the Queen is too far away to restrict the King's movement effectively.

It's easy to see how the general inductive step goes. We omit the details because our simplified proof makes mathematical induction unnecessary here.

To complete the justification of our claim, we give a winning algorithm for the Queen if $m > n \geq 3$. We consider the game now from her point of view. She will win if she can achieve the position in figure 4, with the King on (i, j), provided that $i + j \leq n$.

From this position, all possible moves by the King are indicated by arrows. The Queen's responses are shown by arrows with matching labels.

Note that after each move, the position is again that in figure 4. The King's column number never increases, and it can't remain constant forever. Thus, the King will be driven to column 1 eventually. It's now a simple matter for the Queen to march him up column 1 to $(1, 1)$.

We've already proved that the Queen can't win on a square board. This is a good place to pause and see why the squareness of the board makes such a big difference. It's certainly possible for the Queen to get the King into the position in figure 4, with the King on (i, j), where $i + j \leq n$. It's also possible for the King to

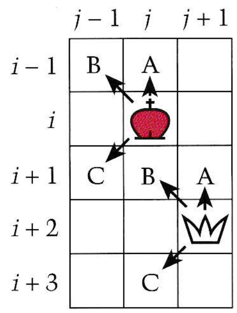


Figure 4

keep $i + j = n$. By choosing option C every time, he will reach column 1 on $(n - 1, 1)$. The Queen must now move to $(n + 1, 2)$, but this is possible if and only if $m > n$.

We now prove that the Queen can win, with or without getting the King into the position in figure 4. Her initial objective is to achieve any of the three positions shown in figure 5. If $n \geq 5$, this is easily accomplished by the Queen giving check on $(m, 1)$. The King can move to either $(m - 1, n - 1)$ or $(m - 1, n)$. The Queen then goes to $(m, n - 3)$ or $(m, n - 2)$ accordingly.

For $n = 4$, the Queen first moves to $(m - 1, 1)$. For $n = 3$, the Queen first gives check on $(m - 2, 1)$. In each of these two special cases, at most two more moves will lead to a desired position. The reader can work out the details.

From the positions in figure 5, all possible moves by the King are indicated by arrows. The Queen's responses are indicated by arrows with matching labels. If the King happens to be in column 1 or n, option A or D for the Queen in figure 5b isn't possible. She should make the alternate response, indicated by the gray arrows with the matching labels.

Note that after each move, the position is again one of the three in figure 5. The King's row number never increases, and it can't remain constant forever. So the King will be driven to row 1 eventually.

When the King reaches row 1 at $(1, j)$, the Queen abandons the strategy indicated in figure 5. Instead she gives check at $(3, j)$, which she can always do. We now consider three cases.

Case 1: $j < n$ and the King moves to $(1, j - 1)$. The Queen moves to $(2, j + 1)$ and marches the King along row 1 to $(1, 1)$.

Case 2: $j < n$ and the King moves to $(1, j + 1)$. The Queen continues to check along row 3. If the King moves back toward $(1, 1)$ before reaching $(1, n)$, the Queen can convert the situation to that in case 1. If the King goes to $(1, n)$, the Queen makes an unexpected move, from $(3, n - 1)$ to $(4, n - 1)$. The King's moves are now forced: King to $(2, n)$, Queen to $(3, n - 2)$, King to $(1, n - 1)$, and Queen to $(3, n)$. She has now achieved the winning position in figure 4, since $1 + (n - 1) = n$.

Case 3: $j = n$. The King must move to $(1, n - 1)$. The Queen gives check at $(3, n - 1)$. The King's response will lead to either case 1 or case 2.

This completes the proof that the Queen wins if $m > n = 3$.

Jesse Chan, **Peter Laffin**, *and* **Da Li** *are high school students in and around Edmonton, Alberta, Canada. They are members of a Saturday mathematics club under the direction of Andy Liu and produced this article as a team project. The club also competes regularly in the Tournament of Towns (see the Happenings department in the January 1990 and November/December 1990 issues).*

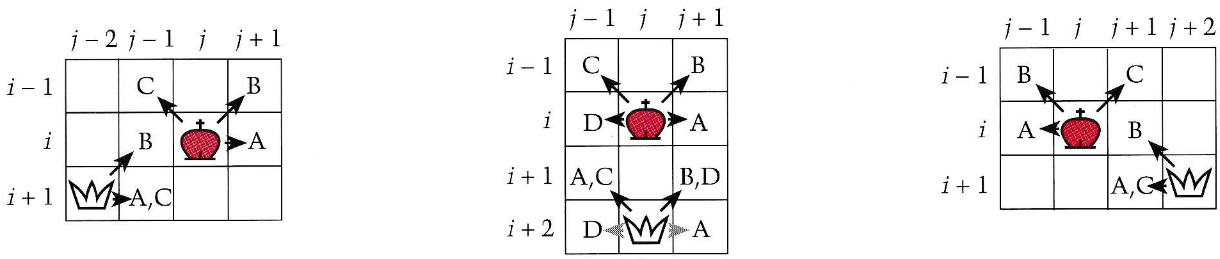


Figure 5 a b c

The problem book of history

Bringing a mathematical turn of mind to a study of the past

by Yuly Danilov

IT WOULDN'T BE AN EXAGgeration to say that the human race—or at least the community of scientists—can be divided into "problemists," who know the point of a difficult and beautiful problem, can properly appreciate an elegant shortcut, and know the joy of a sudden insight, and "nonproblemists" (including those extremists, the "antiproblemists"), who take no pleasure in any of the above. Historians, biologists, philologists, philosophers, and the like often belong to the latter group. It is with a feeling close to sympathy that mathematicians look upon their unreasonable (so they think) fellow scientists who don't know the charm of the problemistic paradise. They try their best to bring their prodigal colleagues into that joyful region.

And that is the motive force behind a strange new kind of problem book that has appeared on the scene. It's as if it was written by a two-headed person who combines two specialties in one body: a mathematician (the *donor*), eager to share the experience of problem solving, and a historian, biologist, or philologist (the *recipient*), who is a kind of receptor for a new problemistic culture.

The mathematician and historian Sergey Smirnov has written just such a book of problems. Smirnov is convinced that a problem book of history can and must be as fascinating as commonly known olympiad problem books of mathematics, physics, and chemistry. A good problem book (like a good textbook) can't be compiled in the silence of one's study, at a solitary desk. The author needs feedback. That's why since 1987 the students in one Moscow school have been studying history not from a textbook (or at least not *only* from a textbook) but from a problem book. Since then the problem book of history grew considerably and now includes ancient history, the Middle Ages, and Russia up to the time of Peter the Great.

The history problems enjoyed great success and managed to penetrate the annual Lomonosov Tournament, which is an interdisciplinary olympiad in astronomy, biology, and linguistics held every autumn in Moscow.[1]

One's appetite is whetted by eating, so when the tournament laureates return home they feel an insistent hunger for a good nonstandard problem. They wander the halls at school with signs pinned to their backs: "Suffer from insomnia! Will take any problem!"

Sergey Smirnov thinks these people need help. Their teachers and parents need help even more. A new problem book (in three parts) is currently being prepared for publication, but excerpts have been published in the Russian journal *Znanie—sila* ("Knowledge Is Power").

All the problems are printed with solutions. Almost all of these problems have been posed at the Lomonosov Tournament and will certainly test the knowledge and imagination of our readers. Go ahead, put yourself in the shoes of tournament participants!

Some of the problems are formulated as explicit questions that test your erudition and wit. Other problems are hidden in literary or historical texts. These texts are thickly strewn with historical mistakes—try to find them! Something described there didn't happen that way, or happened somewhere else, or some time else, or didn't happen at all. Each error you find will give you the chance to feel smarter than the author—don't pass up this opportunity!

Maybe other problem books like this one will soon appear—one on biology, another on works of fiction, maybe even one about structural linguistics . . . After all, problem solving is the most important activity for a specialist in any field, and the more problems we solve, the clearer it becomes that we encounter such problems at every step we take.

[1]Mikhail Lomonosov (1711–1765), a multitalented scholar, was one of the founders of Moscow State University—see the Publisher's Page in the January 1990 issue of *Quantum*.—*Ed.*

ОШИБКИ: ...ДРЕВНЕЕГИПЕТСКОЕ ИЗОБРАЖЕНИЕ СВЯЩЕННОГО БЫКА АПИСА...

ОТВЕТ: ЭТО НЕ МОЖЕТ БЫТЬ ДРЕВНЕЕГИПЕТСКИМ ИЗОБРАЖЕНИЕМ СВЯЩЕННОГО БЫКА АПИСА, ТАК КАК ОНО ВЫПОЛНЕНО НА ОБОЯХ ИЗГОТОВЛЕННЫХ МОСКОВСКОЙ ОБОЙНОЙ ФАБРИКОЙ В 1992 ГОДУ ХУДОЖНИКОМ Ю. ВАЩЕНКО.

Walk like an Egyptian

Problems

1. Why does flooding on the Nile take place in the summer (not in the spring), and why does it last three months?

2. Why did so many ancient Egyptian inscriptions survive to the present day even though there were few literate people in Egypt?

3. Why didn't any other people adopt the Egyptian way of writing?

4. Could Khufu (Cheops) have carried out the same religious reform as Ikhnaton?

Fractured history

The young Pharaoh Thutmose III mounted his horse and galloped off to inspect the construction of his pyramid. His attendants rode camels. The sovereign of Egypt was accompanied by ambassadors from Hammurabi, king of Babylonia; from the country of Urartu; and from the state of the Hittites. The steel swords of the Greeks—the hired bodyguards of the pharaoh—glittered brightly in the sunlight.

The supervisor in charge of construction, Ptahhotep, who was also the high priest of the god Aton, gave a progress report to the pharaoh and complained about the lack of slaves for baking bricks—otherwise the construction would proceed more quickly. The pharaoh answered that soon he would go to war against the Assyrians and bring back many new slaves. Then the building would quickly be completed, and Thutmose's tomb would be taller than the tombs of his great ancestors Djoser and Khufu. The foundation of the pyramid was made enough deep this time and would not sink in the sand, as happened with the tomb of the unfortunate Pharaoh Ikhnaton.

Thutmose III praised his humble servant for a job well done. He promised Ptahhotep that after the successful completion of the great building, the high priest would be allowed to wear a silk skirt instead of the usual cotton one and to erect his own tomb near the pharaoh's at the state's expense.

"Glory to the Great House!" cried the architect, overjoyed. "May it live, healthy and strong! May the benevolence of the goddesses Maat and Ishtar never desert their favorite—the sovereign of the Two Kingdoms!"

Can you spot the errors in this narrative?

At left: "Errors: . . . An ancient Egyptian depiction of the sacred bull Apis; . . . Answer: This cannot be an ancient Egyptian depiction of the sacred bull Apis, since it was painted on wallpaper manufactured by the Moscow Wallpaper Factory in 1992 by the artist Y. Vashchenko."

Art by Yury Vashchenko

The glory that was Greece

Problems

1. List all the towns you know in Phoenicia and Greece. Why were there more towns in Greece?

2. They say that in Athens one could attend the theater free of charge. Not only that, everyone was obliged to attend. Why?

3. All the philosophers of ancient Greece that we know about lived after Homer. Why is that?

4. Sparta was founded much later than Athens, yet the state structure of Sparta seems more primitive. Why?

Fractured history

This is a day of celebration: exactly thirty years ago, on the broad seaside plain near Thermopylae, an Athenian phalanx crushed the numerous but disordered Persian horde. The king of kings Artaxerxes III escaped disguised in a woman's chiton, while the cavalry of "immortals" covered the general retreat. On that very day the Olympic archery champion, the young Pythagoras, added to his golden wreath a more valuable trophy: the iron crown of the Sāsānians, which was left in the tent by the cowardly king.

A year later a granite statue of the young hero was erected in the forum in front of the Parthenon. Pericles himself delivered a solemn speech and made a sacrifice to Asclepius on behalf of his best warrior, and old Euripides composed an ode in honor of this child of fortune, whose arrows struck down the enemies of his homeland as assuredly as they struck the wooden target.

Is it possible to surpass such glory? Would it not be better to perish at the height of one's powers, leaping off the Tarpeian Rock as the Spartans do, than to live out one's years bragging in the bazaar about past exploits, not noticing the sneers of the youth people? These are the questions Pythagoras put to himself time and again, and could find no answer. At last Socrates, his old comrade-in-arms, gave him some advice (Socrates was no less brave than Pythagoras but was far less lucky).

"Pythagoras! Both of us need to change professions immediately and seek new happiness in life. I intend to study with Paracelsus and become a physician, and you must go to old Euclid and prove to him that it is not only timid milksops who can scale the heights of geometry!"

This was his wise friend's advice. Pythagoras took it, and once again put in hard years at the gymnasium . . .

Euclid was a severe teacher. Often repeating that "there are no royal or olympian roads in geometry," he recognized Pythagoras as a worthy student only after Pythagoras had read all eight books of the *Elements* and compiled a problem book to go with it. All the problems were quickly solved by Pythagoras and his new friends: Diophantos, Plato, and Aristotle. That is, all but one: how to measure a diagonal in a square. It turned out that no one in the world—not even the wise Egyptians, not even the Teacher himself—could do it! On that very day Pythagoras made a decision: "Here is the new aim in my life! If I reach it, then I shall sacrifice to Hecate an unprecedented offering: one hundred black horses!"

Much time has passed since old Euclid died. His students dispersed all over Europe. Plato sailed to Rome, Aristotle went south, all the way to Macedonia, to teach a local prince mathematics. He didn't do a bad job! Alexander did what no Athenian could do—he drove the Persians into the heart of Asia and reached the end of the world: the fabled land of China. From there the Greek king sent marvelous plunder to Athens: talking parrots, sweet white pebbles, and unusual manuscripts on the giant leaves of some strange tree. Nobody could read them, but in one manuscript there was a small drawing that upset Pythagoras for some reason. It depicted an ordinary chessboard, but why was it divided into triangles?

Pythagoras spent an entire year trying to understand this. Then the solution came to him in a flash: waking up three days ago, he understood at once that an unknown Chinese mathematician was able to calculate the diagonal of a square! How, Pythagoras couldn't say, but the answer was clear from the drawing. If Pythagoras couldn't arrive at the same result on his own, then all his years of study under Euclid were wasted!

And here we are: the long-awaited proof is ready—a very strange drawing similar to Persian sharovars[2] . . . Could it be that without Alexander and his Persian campaign, and without his old friend Aristotle, who taught Alexander, Pythagoras would never have been able to add the last theorem to Euclid's great treatise, the famous Pythagorean trousers?[3]

Only the gods know the answer. But they don't talk to mortals . . . Anyway, old Pythagoras can be satisfied with what he accomplished. The second half of this life wasn't spent in vain. Today he will honor the gods with his long-promised offering!

Can you spot the errors in this narrative?

COMMENTARY ON PAGE 59

---

[2]Loose trousers gathered at the ankles.—*Ed.*

[3]Famous in Russia, anyway, because of a student rhyme: Пифагоровы штаны во все стороны равны (Pythagorean trousers are equal in all directions).—*Ed.*

HAPPENINGS

World-class physics in colonial Williamsburg

Highlights from the XXIV International Physics Olympiad

by Larry D. Kirkpatrick

THE US PHYSICS TEAM WON four medals and an honorable mention at the XXIV International Physics Olympiad that was held in July on the campus of the College of William and Mary in Williamsburg, Virginia. This was the first time in the eight-year history of the Team that it has been awarded four medals and only the second time that all five team members have received an award.

The US Physics Team was led by gold-medal winner Dean Jens from Ankeny, Iowa, who tied for 12th among the 196 high school competitors from 41 nations with a score of 36.4 compared to the top score of 40.65 out of a maximum of 50. Dean placed third on the theoretical portion of the exam, receiving a score of 24.9 points out of a possible 30; the top theoretical score was 25.65. Daniel Schepler from Beavercreek, Ohio, was awarded a silver medal and placed 32nd in the competition with a score of 31.55. He was followed by the two bronze medal winners: Hal Burch from Ponca City, Oklahoma, and Dmitri Linde from Stanford, California, who received the highest experimental score among the US Team members. Chang Shih Chan from Philadelphia, Pennsylvania, was awarded an honorable mention. All five team members placed in the top 40% of the brightest physics students in the world.

The 1993 US Physics Olympiad Team, coaches, and sponsors at the National Academy of Sciences in Washington.

There were 16 gold medals (G), 17 silver medals (S), 32 bronze medals (B), and 38 honorable mentions (H) awarded at the closing ceremonies by Nobel laureates Leon Lederman, Jerome Friedman, and Val Fitch. There was a tie for the top award between Harald Pfeiffer of Germany and Junan Zhang of China. The youngest competitor was Akshay Venkatesh of Australia, who is 11 years old and a bronze medal winner. Three countries received five medals each: China (GGSSB), Germany (GGSBB), and Russia (GGGSS). Seven other countries received five awards each: Bulgaria (SBHHH), the Czech Republic (GGSSH), Great Britain (GSSBH), Hungary (GGGBH), Romania (GGBBH), Turkey (GBHHH), and the USA (GSBBH). Australia (BHHH) and Canada (BBBH) were among the countries winning four awards each.

The competition at the International Physics Olympiad consists of two five-hour examinations that are taken individually. The theoretical portion consists of three problems. In the first problem students were told that there was a downward electric field at the Earth's surface and asked to determine the surface charge density, the total charge on the Earth, and the net volume charge density near the ground. They were then asked to analyze a hypothetical device designed to measure this field. In the second problem students were asked to analyze the forces exerted by a laser beam passing through a triangular prism. The final problem involved the deflection of an electron beam by a charged wire and the resulting interference pattern produced on a distant wall.

The first experimental problem was in stark contrast to the hot, humid East Coast weather. The students were asked to measure the latent heat of vaporization of liquid nitrogen using an aluminum block and then using an electrical resistor. The second experimental problem was a very difficult exercise in determining the magnetic moment of a dipole magnet and the axial dependence of the magnetic field of an unknown magnet. The top score on the experimental exam was a remarkable 19.5 out of a possible 20 by Gabor Veres of Hungary.

"Drachen Fire," soccer, & more

There is much more to the International Physics Olympiad than the examinations. According to the students the most valuable aspect is the opportunity to meet and interact with students from around the world who share their love of physics and mathematics.

The students were treated to a tour of Colonial Williamsburg, an afternoon at Water Country USA, a physics demonstration night, a day at the Busch Gardens amusement park, tours of the Continuous Beam Electron Accelerator Facility and NASA–Langley, a street party, a computer workshop, a Paper Olympics, a day at Virginia Beach with a stopover at a shopping mall, an American Music Fest, and a talent night after the closing banquet. As you can see, the visiting students were given a very good introduction to American student life.

The US students were also successful in the informal competition that took place during the trip to Busch Gardens. Hal Burch teamed with competitors from Australia, Canada, Mexico, and Poland to win first place in the analysis of the physics of the "Big Bad Wolf," a suspension-type roller coaster. Dean Jens's team was composed of students from Greece, Italy, the Netherlands, and Spain and won second place in the analysis of "Drachen Fire," a roller coaster with many twists and loops. In addition, three members of the US Physics Team managed to tie students from the Czech Republic in a soccer match.

The US Physics Team was coached by the author, a professor of physics at Montana State University and *Quantum*'s field editor for physics, and Theodore Vittitoe, a physics teacher at Libertyville (Illinois) High School, with the assistance of P. Wilson Bascom from Wootton High School in Rockville, Maryland. The selection and training of the US Physics Team is the responsibility of the American Association of Physics Teachers under the direction of its Executive Officer, Bernard Khoury. Fund raising is directed by the American Institute of Physics on behalf of its member societies. The International Physics Olympiad was hosted by AAPT with the assistance of AIP under the direction of *Quantum* contributor Arthur Eisenkraft. The examination committee was headed by Anthony French of MIT and the local organizing committee by Hans von Baeyer.

The selection of the US Physics Team began last fall with the solicitation of nominations from high school physics teachers. Two written exams narrowed the field to the top 20 students, who attended a weeklong training camp held at the University of Maryland during the last week in May. The top five students were chosen to represent the Team and sent home with a hundred practice problems to further hone their skills. These students gathered once again at the University of Maryland for three days of training in laboratory skills just before traveling to Williamsburg for the International Physics Olympiad.

Students wishing to try out for positions on next year's US Physics Team, which will compete in Beijing, China, should contact Bernard V. Khoury, AAPT, 5112 Berwyn Road, College Park, MD 20740-4100 for application materials.

The 1993 US Physics Team

In the list below, team members who represented the US in Williamsburg are marked by an asterisk; each student's school and physics teacher are noted in parentheses.

James Ayers, Houston, Texas (Langham Creek High School, Bob Menius)

Adrian Banard, Alexandria, Virginia (Thomas Jefferson High School for Science and Technology, John Dell)

Val Fitch awarding a gold medal to Dean Jens.

*Hal Burch, *bronze medal*, Ponca City, Oklahoma (Oklahoma School of Science and Mathematics, Wei Chen)

Mrunil Champaneri, Des Plaines, Illinois (Maine Township High School East, Scott Welty)

***Chang Shih Chan**, *honorable mention*, Philadelphia, Pennsylvania (Northeast High School, Raj Rajan)

Joey Chang, Cherry Hill, New Jersey (Cherry Hill High School West, Hirendra Chatterjee)

Kurt Franke, San Antonio, Texas (John Marshall High School, Jeff Szabo)

***Dean Jens**, *gold medal*, Ankeny, Iowa (Ankeny High School, Donald Savage)

Jack Langsdorf (alternate), Watertown, Connecticut (The Taft School, Jim Mooney)

***Dmitri Linde**, *bronze medal*, Stanford, California (Gunn High School, Art Farmer)

Alexander Morcos, Greensboro, North Carolina (North Carolina School of Science and Mathematics, John Kolena)

Edward Otte, Tampa, Florida (Berkeley Preparatory School, Walter Murfee)

Andrew Poynot, Slidell, Louisiana (Louisiana School for Mathematics, Science and Arts, Ali Rahbar)

Dmitriy Rogozhnikov, Bronx, New York (John F. Kennedy High School, Patrick Clark)

***Daniel Schepler**, *silver medal*, Beavercreek, Ohio (Beavercreek High School, Margo Debrosse)

Robert Schneck, Charlotte, North Carolina (North Carolina School of Science and Mathematics, John Kolena)

Carl Streeter, Oshkosh, Wisconsin (Oshkosh North High School, Jeffrey Elmer)

Christopher Tom, Honolulu, Hawaii (Iolani, Carey Inouye)

Robert Wagner, Cherry Hill, New Jersey (Cherry Hill High School West, Hirendra Chatterjee)

Jonathan Weinstein, Lexington, Massachusetts (Lexington High School, Mark Rodriguez)

Bulletin Board

Guest account on supercomputer

Lawrence Livermore National Laboratory now has a "guest account" that students and teachers across the nation can use to access the National Education Supercomputer (NES) without attending a workshop or obtaining training outside the classroom.

Operated by the Laboratory, the supercomputer is a Cray X-MP donated by Cray Research in 1990 exclusively for educational use. It is used to run large-scale simulations and models through a "point and click" interface on the microcomputer. The simulations currently available include ray tracing and climate modeling.

The guest account is actually a part of the National Education Bulletin Board System (NEBBS), a Sun workstation connected to the NES through special networking software, said Brian Lindow, technical director of the NES Program.

The guest account gives teachers and students the ability to run simulations and models on the NES. Schools with network or modem access can start using the applications on the NES and the microcomputer immediately by calling a local Tymnet number to gain access to the network. The simulations are run on the NES remotely through NEBBS.

The menu-driven software on NEBBS allows the guest to upload input files and automatically moves the files to the NES and starts the simulation. When the simulation is finished, the "movie file" is moved back to the NEBBS for downloading.

Schools wishing to access the guest account need a color Macintosh or an IBM-compatible computer with VGA card and mouse, a telephone line, and a modem. For additional information, contact Brian Lindow, c/o Lawrence Livermore National Laboratory, L-561, PO Box 808, Livermore, CA 94551, or call 510 294-5464.

Duracell Scholarship Competition

The 12th Annual Duracell/NSTA Scholarship Competition will award 100 prizes to inventive students (up from 41 prizes last year). The awards will now be in the form of US Savings Bonds, ranging from $20,000 to $100. Winners are still encouraged to use the proceeds from the bonds to further their education.

The program is open to students nationwide in grades 9 through 12. To enter the competition students must create and build a working battery-operated device that performs a practical function and is designed to educate, entertain, or make life easier in some way. Each entry must be designed and built by the entrant.

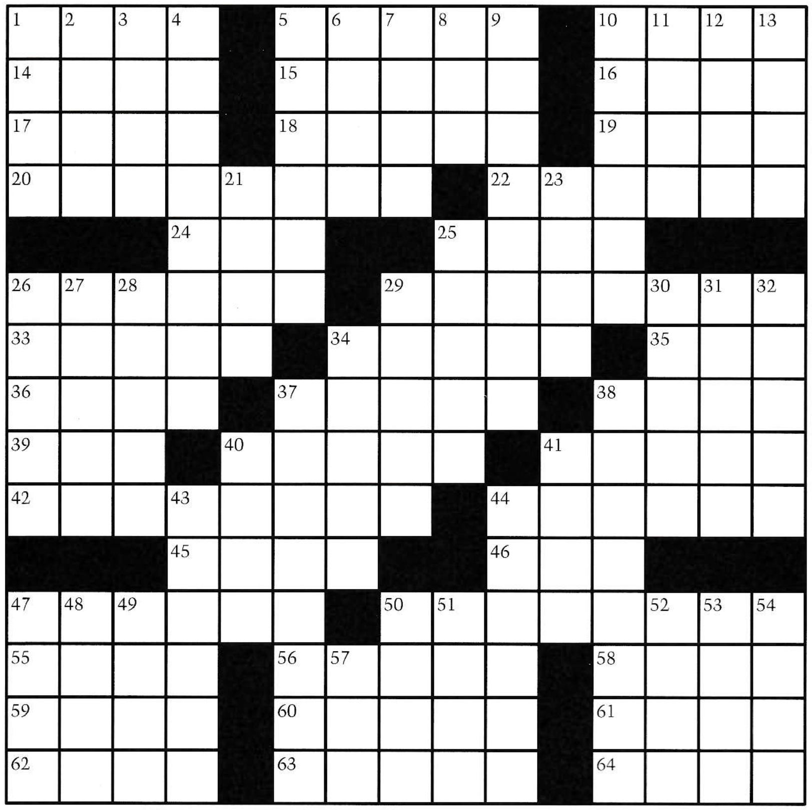
Winners in previous years have invented some particularly useful devices. One of last year's second place winners was Annette Plepenhagen, who attended Thomas Jefferson High School in Alexandria, Virginia. Her invention was an educational device that helps teach people how to write Katakana, the Japanese writing system used for words of foreign origin. Another prize winner, Justin Chester, who attended Coronado High School in Colorado Springs, Colorado, invented a device designed to halt carjackings by allowing a car owner to turn off a vehicle using remote control.

Entrants must have a teacher/sponsor. In addition to the scholarship awards, first and second place winners will receive an all-expense-paid trip to Anaheim, California, for the NSTA National Convention, March 30–April 2, 1994. Parents and teacher/sponsors will accompany the winners, all expenses paid.

Entry kits will be mailed to teachers this fall. For more information or additional kits, write to Duracell/NSTA Competition, 1840 Wilson Boulevard, Arlington, VA 22201-3000. All entries must be received by January 21, 1994.

Criss × cross science

by David R. Martin



Across

1 Former
5 Rental agreement
10 Store
14 Width times depth
15 Artist's tool
16 Healing plant
17 Legal claim
18 Plant stem
19 Aquatic animal's organ
20 Decimal number part
22 Earth's middle layer
24 Exist
25 ___ Descartes
26 Former USSR leader
29 Gregor Mendel's field
33 Organic soil
34 Acid and base compounds
35 A cheer
36 First man
37 Greek letters
38 Appointment
39 Curl
40 ___ decimal system
41 Joined
42 Humans and monkeys
44 Units of length
45 Moroccan tree
46 Greek letter
47 Counterbalance
50 Conic section
55 Naked
56 Moist
58 Snake, e.g.
59 Bakery worker
60 Of sheep
61 Shopper's delight
62 Golfing needs
63 Belgian province
64 Building wings

Down

1 Hand part
2 Opera highlight
3 Spotted
4 Element 73
5 Diminish
6 Devours
7 A continent
8 Fall month: abbr.
9 Periodic table members
10 Electro or ferro follower
11 Landed
12 Move on wheels
13 Distant: pref.
21 Eye part
23 Once (dialect.)
25 Common electric circuit element
26 Keen
27 Ruling Englishman
28 Violin maker
29 AND and NAND, e.g.
30 Angry
31 Provide food
32 Storage structures
34 Waste water conduit
37 Particle accelerator
38 Computer accessed information
40 Challenge
41 Of a benzene ring position
43 Microwave sources
44 Pennsylvanian town
47 Death notice
48 Front
49 Gibbs ___ energy
50 Demure
51 Japanese aborigine
52 Hydrous silica
53 Lounge
54 Summer drinks
57 Eggs

SOLUTION IN THE NEXT ISSUE

SOLUTION TO THE JULY/AUGUST PUZZLE

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

ANSWERS, HINTS & SOLUTIONS

Math

M91

The game in question is "uninteresting" in a certain sense, because its result does not depend on the players' strategies and is completely determined by the initial position of the chips—more exactly, by the distance between them (except that the game can be endless—if, say, the players alternately move their chips back and forth).

Clearly, every half-move (a move of one player) changes the parity of the distance between the chips—that is, of the number of squares between them. A chip can be trapped only in one of the extreme squares—if the other chip is in the nearby square. And for $N = 30$, or any other *even* N, this can happen only after an even number of moves (because the distance between the chips in the final position is zero). So in this case *only the second player can win*—and will win by moving his or her chip always in the same direction.

Similarly, for odd N only the first player can win.

This solution is based on the fact that the process under consideration has an *invariant*: the parity of the distance between the chips. This kind of invariant is rather typical for many problems like this. More involved applications of the method of invariants can be found in the article "Some Things Never Change" in this issue; see also the solution to M95. (V. Vasilyev)

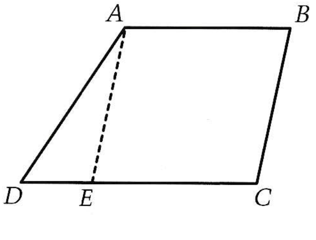


Figure 1

M92

Drawing a line from the vertex of the shorter base of a trapezoid parallel to the leg from the other vertex of this base (see figure 1), we construct a triangle ADE two of whose sides are equal in length to the trapezoid's legs, and the third side is the difference of the bases. Now the Triangle Inequality for ADE ensures that the difference of the legs is always less than the difference of the bases (in our figure, $|AD - BC| = |AD - AE| < DE = CD - AB$). So only one of the two pairs of opposite sides can serve as the bases of a trapezoid. (V. Proizvolov)

M93

Let one of the two given integers be x; then the other one is $a - x$, where $a = 30{,}030$, and the divisibility of their product by a can be written as $x(a - x) = ak$ for some integer k, or $x^2 = a(x - k)$. So x^2 must be divisible by a, which means that x is divisible by every prime factor of a and, therefore, by their product. (Strictly speaking, this follows from the uniqueness of factorization of an integer into primes—see, for instance "Divisive Divisors" in the September/October 1991 issue of *Quantum*.) The number $a = 30{,}030$ is factored as

$$a = 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13.$$

So x must be divisible by this product, and, consequently, $x \geq a$ or $a - x \leq 0$. But $a - x$ is a positive integer. This contradiction proves the statement of the problem.

Of course, the same argument works for any number a that is "square-free"—that is, a number whose factorization contains only the first powers of primes. For other numbers a, the statement of the problem is wrong: if an integer a is divisible by $p^2 > 1$, we can take $x = a/p$, then $x(a - x) = a(a/p - a/p^2)$ is divisible by a. (V. Vasilyev)

M94

The number n can take any value not less than 6 except $n = 7$. Examples of the polyhedrons in question for all these values of n can be constructed by truncating a "bipyramid"—the solid formed by two congruent pyramids put together base to base. Let the common base of these pyramids be a k-gon $A_1A_2...A_k$ ($k \geq 3$). When we truncate the corner A_1 of the bipyramid with a plane drawn perpendicular to the plane $A_1A_2...A_l$ through the midpoints of the edges A_1A_2 and A_1A_k, the four triangular faces meeting at A_1 turn into quadrilaterals, and a new quadrilateral face is formed in the truncating plane. If we now cut off the corner A_2 in the same manner, all the quadrilateral faces created before remain quadrilaterals (though two of them will be truncated still more), but three new quadrilateral faces will appear, and so on.

Cutting off all the corners A_1, ..., A_k, we get a $3k$-hedron with quadrilateral faces (for instance, the 9-hedron in figure 2); we can also cut

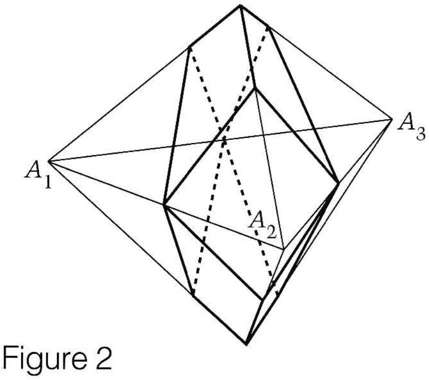


Figure 2

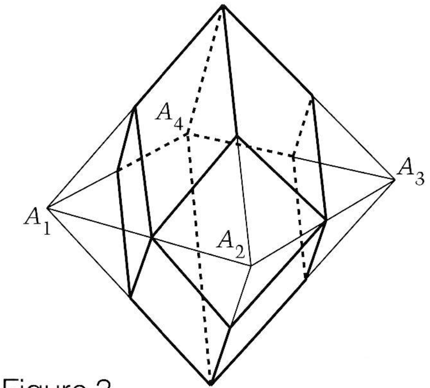


Figure 3

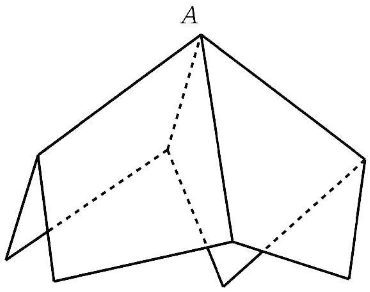


Figure 4

off only the corners $A_1, \ldots, A_{k-1}$, leaving A_k as it is, and thus create a $(3k-1)$-hedron, whose faces will be quadrilaterals all the same (see figure 3 for $k = 4$, $n = 3k - 1 = 11$). And, finally, for $k \geq 4$ we can just as well leave two nonadjacent corners—say, A_2 and A_k—intact. This yields a $(3k-2)$-hedron. Note that any number $n \geq 6$, $n \neq 7$, is representable as $3k$ or $3k - 1$ (for $k \geq 3$), or as $3k - 2$ (for $k \geq 4$). In particular, $1993 = 3 \cdot 495 - 2$, so in the case specified in the problem we must apply our construction for $k = 495$ and cut off all the vertices A_i except, say, A_2 and A_k.

Now let's show that the number n cannot be less than 6 nor equal to 7. First, n cannot be less than 5, because any face of a polyhedron satisfying the condition borders on four other faces. Second, the fact that $n \neq 5$ can be proven by searching through a number of possibilities. But it's neater to make use of Euler's formula $F - E + V = 2$, where F, E, and V are the numbers of faces, edges, and vertices, respectively, of a polyhedron.[1] If $F = 5$, then $2E = 20$ (because in the total number $4 \cdot 5 = 20$ of the sides of all faces, each edge is counted twice), so $V = E + 2 - F = 10 + 2 - 5 = 7$. But this leads to a contradiction: since there are at least three edges issuing from every vertex, and each edge joins two vertices, $3V \leq 2E$, or, in our case, $3 \cdot 7 = 21 \leq 20$. Finally, if $n = F = 7$, we find, just as above, that $2E = 4F = 28$, and $3V \leq 28$. This means that $V \leq 9$. In addition, some vertex A must be a meeting of least four faces (because we have $3V < 4F$). The number of vertices of the polyhedron belonging to all the faces meeting at A, not counting A, is clearly not less than 8 (see figure 4), so it's exactly 8—there are no other vertices—and these faces are arranged exactly as shown in figure 4. You can verify on your own that such an arrangement can't be completed with a polyhedron having quadrilateral faces without adding new vertices.

M95

The answer is no. The coloring of the chameleons in the problem can be described by a triple of nonnegative integers (x, y, z) equal to the numbers of blue, brown, and black chameleons, respectively. Then we have to prove that the initial triple $(13, 15, 17)$ can't be transformed into any of the triples $(45, 0, 0)$, $(0, 45, 0)$, $(0, 0, 45)$ by way of the operations described in the condition—that is, by adding 2 to one of the numbers of a triple and simultaneously subtracting 1 from the other numbers. This is a typical setup in which it's reasonable to try to find an *invariant*—a function of our triples that is preserved by the given transformations and takes different values for the initial triple and for the three possible final ones.

Such a function exists: we can take, for instance, the remainder r of the difference $x - y$ when divided by 3.

Indeed, our operations replace the first two numbers (x, y) of our triples by either $(x - 1, y - 1)$, $(x + 2, y - 1)$, or $(x - 1, y + 2)$. The difference $x - y$ doesn't change in the first case and changes by 3 in the other two cases, so the remainder r doesn't change. At the same time $r(13, 15, 17) = 1$, while $r(45, 0, 0) = r(0, 45, 0) = r(0, 0, 45) = 0$.

Figure 5 provides a nice geometric interpretation of the problem, enabling us to answer a general question: under what conditions can one triple of integers be obtained from another triple using our operations?

Let $h = x + y + z$ be the sum of the numbers of a triple; of course h is an invariant, too (there is "conservation of chameleons"!). Consider an equilateral triangle with height h divided into h^2 equilateral triangles with height 1, as shown in figure 5 for $h = 9$. It's easy to see that the distances x, y, and z from any node of the triangular grid thus obtained to the sides of the big triangle are integers, and that $x + y + z = h$ (fig. 6). So the nodes represent all the possible triples considered in the problem. The "recoloring of chameleons" can be represented now by arrows joining every node (x, y, z) to the nodes (x', y', z') obtained from (x, y, z) under the corresponding transformations of triples. Take any node, draw the arrows issuing

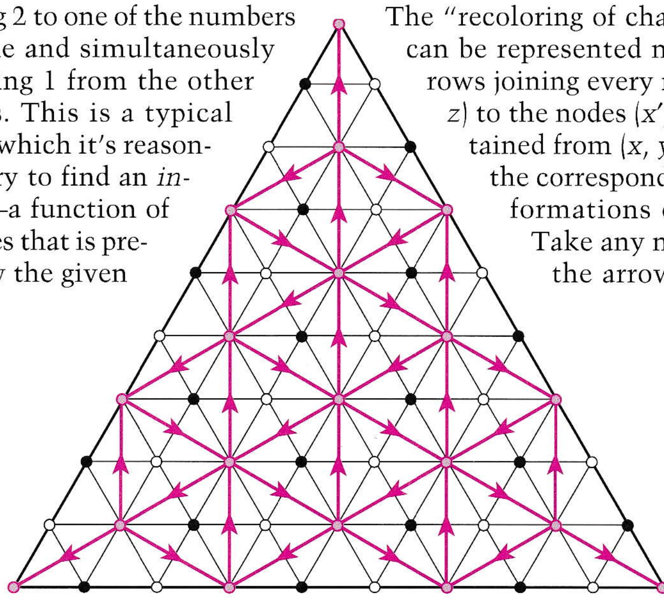


Figure 5

[1]See "Topology and the Lay of the Land" in the September/October 1992 issue.—*Ed.*

$$x+y+z=\frac{2\cdot \text{area}(ABP)}{AB}+\frac{2\cdot \text{area}(BCP)}{BC}+\frac{2\cdot \text{area}(CAP)}{CA}$$

$$=\frac{2}{AB}\left[\text{area}(ABP)+\text{area}(BCP)+\text{area}(CAP)\right]$$

$$=\frac{2\cdot \text{area}(ABC)}{AB}$$

$$=h$$

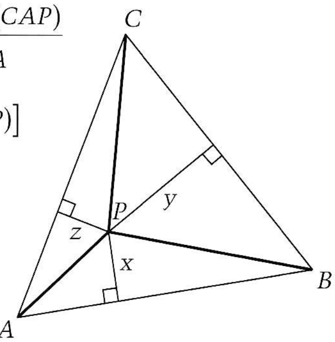


Figure 6

from it, then the arrows from their endpoints, and so on. The endpoints of all these arrows represent the triples that can be obtained from the initial triple by "recoloring." As figure 5 shows, every red node except the vertices of the big triangle is joined by a chain of arrows to any other red node. Drawing the arrows for the remaining nodes, we'll obtain two other connected sets of nodes (colored black and white in the figure; the arrows between them are not shown). Clearly, each of the three classes of nodes is characterized by the value of r: for the red nodes, $r = 0$; for the black nodes, $r = 1$; and for the white nodes, $r = 2$.

Thus, the triple (x, y, z) can be transformed into (x', y', z') if and only if the invariants r and s are the same for both triples and at least two of the numbers x, y, z are positive. In particular, if s is divisible by 3, then for all the vertices of the triangle ($(s, 0, 0)$, $(0, s, 0)$, and $(0, 0, s)$), $r = 0$; so they are accessible only from the nodes with $r = 0$ (as in figure 5). If, however, s is not divisible by 3, it's not hard to show that the invariant r will take three different values, 0, 1, and 2, at the vertices, so one (and only one) of them will be accessible from any other node. (V. Ilyichov, V. Vasilyev)

Physics

P91

At equilibrium the total force acting on the ball from all kinds of pressure is determined by the ball's weight and by the weight of water displaced. If we neglect the compressibility of water and the ball, we find that the depth of immersion does not change with an increase in air pressure. If the compressibility of the floating body is much greater than that of water, the body will sink when the external pressure increases.

This is how the toy known as the "Cartesian diver" works. An upside-down test tube floats in a beaker half-filled with water and sealed at the top with a thin rubber sheet. By pressing the sheet downward, one creates an increase in pressure in the beaker, thus compressing the air inside the test tube (which keeps it afloat). The weight of the test tube becomes more than that of the water displaced by the test tube and the air inside it, so it begins to sink. One can make the "diver" come to the surface again by removing one's hand from the rubber sheet. Try making this simple and amusing toy yourself.

P92

The heat goes to heating the gas and to producing the mechanical work of moving the piston:

$$Q=nC_v\left(T_2-T_1\right)+\frac{k}{2}\left(x_2^2-x_1^2\right),$$

where $x_1 = V_1/S$, $x_2 = V_2/S$, and n is the number of moles of gas. It follows from the ideal gas equation that $P_1V_1 = nRT_1$ and $P_2V_2 = nRT_2$. Recalling that $V_2 = 2V_1$ and $P_2 = 2P_1$, we get $T_2 = 4T_1$. Thus,

$$Q=nC_v\cdot 3T_1+\frac{k}{2}\cdot 3\left(\frac{V_1}{S}\right)^2.$$

From the equation $F = kx$ for the force due to the spring, we get $P_1S = kV_1/S$. Taking into account that $C_v = \frac{3}{2}R$, we obtain

$$Q=n\frac{3}{2}R\cdot 3T_1+\frac{3}{2}V_1P_1,$$

or

$$Q = 6P_1V_1.$$

P93

The fish is killed, of course, not by a direct hit but by the electric current flowing through the water during the discharge. The current density decreases with distance from the spot where the lightning bolt hits (fig. 7). The voltage drop across the fish's body is proportional to the current flowing through the fish and is determined by the current density in the adjacent water. Only fish that are in the region where the current density exceeds the critical value are killed. The characteristic size of this "danger zone" ranges from a few meters to several dozen meters. Obviously several fish can be in this danger zone at the same time.

A similar situation occurs when a high-voltage wire snaps and falls to the ground. In this case the density of the current flowing in the ground decreases as the distance to the contact point increases. The level of danger for people is characterized by the so-called step voltage—that is, the voltage that arises between one's feet

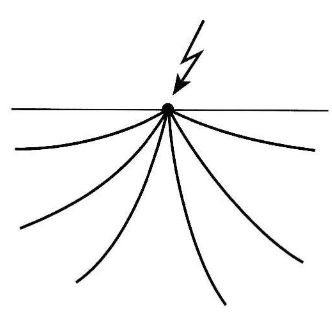


Figure 7

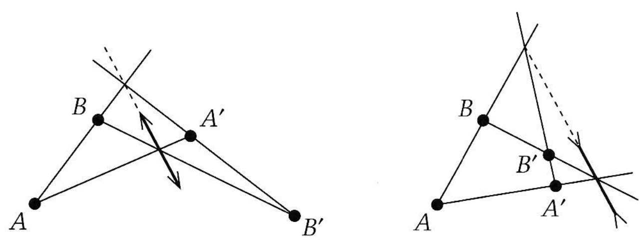


Figure 8

when they touch the ground. It's clear that small steps result in a smaller step voltage, so one must take "baby steps" to get away from the spot where a high-voltage wire has fallen to the ground.

P94

The center of each lens (one converging, one diverging) lies at the intersection of the lines AA' and BB' (fig. 8). The plane of the lens passes through the intersection of the lines AB and $A'B'$.

P95

The energy E hitting the ball during a time interval Δt is given by

$$E = IA\Delta t,$$

where $A = \pi R^2$ is the cross-sectional area of the ball. Newton's second law tells us that

$$F = \frac{\Delta p}{\Delta t}.$$

For an absolutely black ball, each photon is completely absorbed and the change in momentum Δp is equal to the momentum p. Therefore,

$$F = \frac{np}{\Delta t},$$

where n is the number of photons.

Since the momentum of a photon is equal to its energy divided by the speed of light c, $np = E/c$. Therefore,

$$F = \frac{E}{c\Delta t} = \frac{I\pi R^2}{c}.$$

Brainteasers

B91

Let N be the number of different pairs that danced at least once. Then the number of boys, as well as the number of girls, is equal to $N/3$.

B92

Your first try probably went something like this: the check was for x dollars and y cents; the cashier paid y dollars and x cents; since I was left with $2x$ dollars and $2y$ cents after spending \$6.23, the equations to determine x and y can be written as

$$x - 23 = 2y,$$
$$y - 6 = 2x.$$

But the solutions to these equations are not positive integers, so another approach is needed. You have to consider the possibility that x is less than 23 so that, to subtract 23, you have to borrow 100 cents from the dollars figure. This will modify the equations as follows:

$$(x + 100) - 23 = 2y,$$
$$x + 77 = 2y,$$

and

$$(y - 1) - 6 = 2x,$$
$$y = 2x + 7.$$

The solution to this pair ($x = 21$, $y = 49$) has the right property. So the check was for \$21.49; the cashier paid \$49.21. I was left with \$42.98 (\$49.21 – 6.23), which is exactly double the amount of the check.

You can make up a whole class of such problems, each of which differs only in the amount that was spent during the day. You get a new problem by increasing or decreasing the amount by a multiple of 3 cents. The check amount would go down by \$1.02 for each three-cent increase in the amount spent; it would go up the same amount if the spending figure goes down by 3 cents. Thus, if the amount spent were \$6.26, the check amount would be \$20.47. If the amount spent were \$6.20, the check would be \$22.51. There are limits, though—a spending figure of \$6.68 gives a check amount of 6.19. I could not have expected to spend \$6.68 if I believed I had only \$6.19 in my wallet!

Also, some check amounts are more plausible candidates for the kind of mix-up that happened. A spending amount of \$5.72 gives a check figure of \$38.83. But you can't solve *that* problem by setting up the equations given above. You'll have to use insight into other possibilities of borrowing and carrying to get the right equations.

B93

The air in the glass was warmed by the hot water. It expanded, occupying more volume, so some air escaped around the rim of the glass.

B94

It's obvious that in figure 9, obtained by 90° rotation of the given big square about the circle's center, $OC = OA = AB/2 = CD/2$. So $AF = AE$ ($= AB/2$), and $AEDF$ is the given small square, whose side lengths are 1/2 those of the big one.

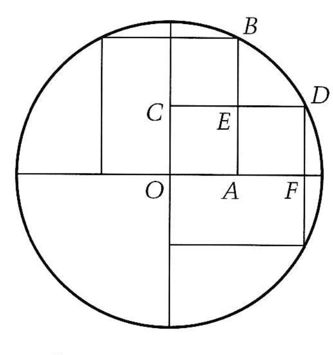


Figure 9

Strictly speaking, this argument is incomplete: it must be proven that there is *only one* "small square" satisfying the requirements of the problem (then it will have to coincide with $AEDF$ in figure 9). The uniqueness follows, for instance, from the fact that the vertex of the small square opposite A must lie at the intersection of the given circumference and the bisector of the angle EAF, which consists of a single point.

B95

After every rearrangement of the rocks their number decreases by one, while the number of piles increases by one. So, if the number of operations was n, in the end we must have $n + 1$ piles of three rocks each and n rocks taken away. This means that the total number of rocks is $3(n + 1) + n = 4n + 3$. But the number 1,001 yields a remainder of 1, not 3, when divided by 4. So the answer to the question is no.

Editor's note: This same problem was posed in the March/April 1992 issue of Quantum (p. 57). The answer that appeared in the next issue (p. 62) used the idea of an invariant, which has figured prominently in this issue.

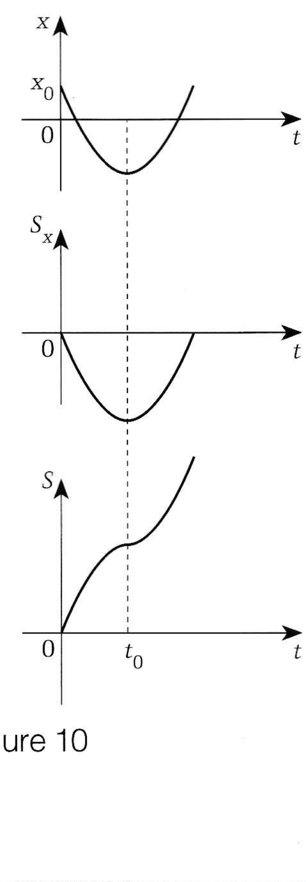


Figure 10

Kaleidoscope

1. There is no unambiguous answer if the sizes of the balls and the times of their arrival at the intersection of the trajectories are unknown.
2. The small circle makes five turns.
3. The mass will move a distance of $2l$.
4. The displacement vector is directed downward, and its magnitude is 1 cm.
5. See figure 10.
6. The trajectory is sinusoidal.
7. The particle moves along a straight line in the direction of wave propagation.
8. The ball will move (a) along a parabola; (b) along a vertical line, if the velocity of the ball is equal in magnitude to that of the car relative to the railway bed; otherwise, it will move along a parabola.
9. The particle will move along a parabola.
10. These points exist on the outer edges of the wheels. The trajectory of one such point is shown in figure 11—it's called a cycloid.
11. See figure 12.
12. The fragments are distributed on the surface of a sphere that expands with a speed v_0. The sphere's center drops with an acceleration g. Each fragment follows its own parabola as it descends.

Microexperiment. The bob will move (a) along a circular segment in the vertical plane; (b) along a horizontal circle (making it a conic pendulum).

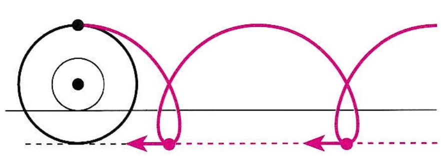


Figure 11

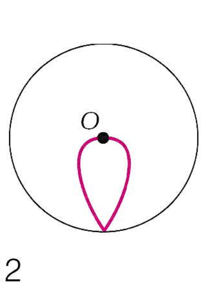


Figure 12

History

Commentary on Egypt's fractured history:

1. At the time of Thutmose III (15th century B.C.) the Egyptians didn't ride horses—they harnessed them to chariots.
2. Thutmose III didn't have a pyramid—he was buried in an underground tomb.
3. There were no camels in Egypt at that time—they didn't appear until about 1000 B.C.
4. Hammurabi ruled in Babylonia three centuries earlier than Thutmose III.
5. The "country of Urartu" didn't exist yet at the time of Thutmose III—this state arose in the 9th century B.C.
6. The Hittite state existed in Asia Minor at the time of Thutmose III, but the Egyptians didn't come into contact with it.
7. In the 15th century B.C. no one in the world could manufacture steel swords.
8. Hired bodyguards of Greek nationality appeared in Egypt after Ramses II (250 years after Thutmose III).
9. The god Aton (the Sun) wasn't worshipped in Egypt until the reign of Ikhnaton, one century after Thutmose III.
10. Egyptians almost completely avoided using clay bricks in building—they had cheap stone (limestone) in abundance.
11. Thutmose III didn't make war on the Assyrians and probably didn't know of their existence. Their remote kingdom in the Upper Tigris was rather weak at that time.
12. It makes no sense to compare the height of Thutmose III's underground tomb and the pyramids of Djoser and Khufu.
13. Djoser and Khufu weren't ancestors of Thutmose III, nor were they relatives of each other—they belonged to different dynasties.
14. Pyramids don't have founda-

tions—they're built not on sand but on bedrock.

15. Ikhnaton also had an underground tomb, not a pyramid.

16. The ancient Egyptians knew nothing about silk or cotton—they used only linen.

17. Ishtar wasn't an Egyptian goddess—she was Babylonian.

18. The protector-goddesses of both lands—Upper and Lower Egypt—were Nekhbet (depicted as a vulture) and Buto (depicted as a cobra). Maat was the Egyptian goddess of truth.

Commentary on Greece's fractured history:

1. At Thermopylae the Athenians and other Greeks were defeated by the Persians.

2. During the second Greco-Persian war the king of Persia was Xerxes.

3. The "immortal" guard was unmounted—they weren't horsemen.

4. "Woman's chiton" is a senseless turn of phrase—the chiton is Greek clothing for males only.

5. Archery was not included among the Olympic games—the Greeks considered this kind of sport "barbaric."

6. According to legend, Pythagoras was an Olympic boxing champion.

7. "Persian horde" is a senseless turn of phrase—the word "horde" is of Turkic origin and didn't appear in Europe until the time of the Huns.

8. Iron crowns appeared in Persia during the reign of the Sāsānian dynasty (3rd century A.D.)—Xerxes was of the Achaemenian dynasty and wore a golden crown.

9. The forum was in Rome—in Athens there was an agora.

10. Pericles lived considerably later than the Greco-Persian wars.

11. Euripides lived later than Pericles and considerably later than Pythagoras.

12. During the Greco-Persian wars, there were no bodyguards in Greece—they appeared during the military democracy (at the time of Homer and earlier).

13. Asclepius was the god of medicine—no one made sacrificial offerings to him in honor of military deeds.

14. The Tarpeian Rock is in Rome.

15. "Bazaar" is a word of Turkic origin—it was unknown in Greece.

16. Pythagoras died of old age on the eve of the Greco-Persian wars.

17. Socrates lived one century later than Pythagoras.

18. Paracelsus was a medieval physician (16th century).

19. Euclid lived two centuries later than Pythagoras.

20. Greek geometers did not compile problem books.

21. Plato and Aristotle lived one and a half centuries later than Pythagoras; Diophantos lived even later.

22. The Greeks sacrificed not horses but oxen and sheep.

23. Hecate was the goddess of the Moon and of witchcraft. A sacrifice in thanksgiving for a scientific discovery would have been offered to Athena, the goddess of wisdom.

24. Plato spent all his life in Athens and was never in Rome.

25. Macedonia is situated to the north, not to the south, of Greece.

26. Alexander reached India, but he probably only heard about China. The Persians apparently called this country something else—"Sin" or "Ser."

27. Alexander of Macedonia did not have the title "King of Greece," despite his control over this country.

28. Cane sugar wasn't white in Alexander's time, but yellow or even brown—no one knew how to purify it.

29. Palm-leaf manuscripts were common in both China and India.

30. The "Indian" proof of the Pythagorean theorem (dividing a larger square into smaller squares and right triangles) became known in Greece much later than Pythagoras discovered his proof.

31. The turn of phrase "Pythagorean trousers" emerged only in modern times. It could not have been used in ancient Greece—the Greeks despised such "barbaric" clothing.

32. The word "sharovars" is of Turkic origin and appeared only in the Middle Ages.

Some things

1. Replacing every plus sign in this exercise by a minus sign, and every minus sign by a plus sign, we turn it into a simple generalization of problem 1 discussed in the article.

2. Write +1 instead of "+" and –1 instead of "–." If we choose any 4×4 subsquare of our diagram and create within it the pattern of shaded squares illustrated in figure 2, the product of the eight numbers thus chosen will be invariant under the operations allowed (you can check this yourself). With this observation, we can apply the method of problem 3 in the text: the product of the shaded numbers must be positive for any position of the 4×4 subsquare or we will not be able to achieve a table consisting entirely of +1's. For the distributions of the signs given in figures 3 and 5 (p. 36), this product equals –1 if the "template" of shaded squares is placed in the top left position on the given 6×6 square; for figure 4 it's –1 for the bottom right position. This means that the answer to all three questions is no.

3. Shade the first, second, fourth, fifth, seventh, and eighth rows of the table considered in the problem. We can see that the parity of the sum of the numbers in the shaded squares is not changed by the permitted operations. For a table without even numbers, this sum is even (48 odd numbers) initially. If this sum is initially odd, then we can never get all the numbers to be even with the given operations.

4. If (i, j) denotes the exchange of the numbers standing in the ith and jth places in a permutation, then the exchange of two nonadjacent numbers—say, in the first and kth places—can be organized as a series of $(2k - 3)$ neighbor exchanges: $(1, 2)$, $(2, 3)$, …, $(k - 1, k)$, $(k - 2, k - 1)$, $(k - 3, k - 2)$, …, $(1, 2)$. A worked example will make this solution clear.

5. Let's use –1's instead of zeros. If $x_1, \ldots, x_9$ are the current numbers in their order around the circle, then the operation we consider simply replaces them with the products $x_1x_2, x_2x_3, \ldots, x_8x_9, x_9x_1$ (if $x_1 = x_2$, then $x_1x_2 = 1$; otherwise $x_1x_2 = -1$). So the product of the numbers after the operation is $(x_1x_2\ldots x_9)^2 = 1$ no matter what $x_1, \ldots, x_9$ were. But the product of nine –1's is –1, and so it cannot appear after our operation.

6. Every time Wendy tears, the number of pieces increases by 9, so its remainder modulo 9 is invariant. But $1 \not\equiv 1993 \pmod 9$. She can never get 1993 pieces.

7. The answer is 10 (the remainder of $1 + 2 + \ldots + 1993$ when divided by 13).

8. There will be more ones than twos. The operation in question preserves the remainders modulo 9, so after a certain number of steps each number turns into its remainder modulo 9. Since $1{,}000{,}000 = 9 \cdot 111{,}111 + 1$, it follows that 111,112 of these remainders are ones and 111,111 are twos.

9. Let x_1 be the given number and $x_2, x_3, \ldots, x_{100}$ the successive numbers obtained by repeated subtractions of their sums of digits. Then all these numbers (except x_1) and all their respective sums of digits are divisible by 9. If $x_{100} > 0$, then $x_{100} \ge 9$, $x_{99} \ge 9 + 9$, $x_{98} \ge 9 \cdot 3$, and $x_{21} \ge 9 \cdot 80 = 720$. Since there are only three 3-digit numbers greater than 720 whose sum of digits is equal to 9 (801, 810, and 900), at least 16 of the 19 numbers $x_{20}, x_{19}, \ldots, x_2$ have a sum of digits not less then 18. But that would mean that $x_1 > x_2 \ge x_{21} + 16 \cdot 18 + 3 \cdot 9 \ge 720 + 315 > 1{,}000$, which is not a 3-digit number. Therefore, $x_{100} = 0$.

10. Label the sectors counterclockwise by the numbers 0, 1, ..., 9 and define the *number of a chip* (for some given arrangement of the chips) as the number of the sector where it belongs. When a chip makes a counterclockwise move (from sector i to sector $i + 1$), its number increases by one except for the move from sector 9 to sector 0, which decreases the number by 9. But in either case we can say that the chip's number *increases* by one modulo 10. Similarly, a clockwise move *decreases* the chip's number by one modulo 10. So, when two chips are moved in opposite directions, the *sum S of the numbers of all chips modulo 10* (that is, the remainder of the sum of all the numbers when divided by 10) is preserved. In the case when all chips are in one sector n, $S = 0$, since $10n \equiv 0 \pmod{10}$; if there is one chip in each sector, $S \equiv 0 + 1 + 2 + \ldots + 9 = 45 \equiv 5 \pmod{10}$. So we can't gather all the chips in one sector.

We can say a bit more about the situation given in this problem. Suppose the chips are distributed haphazardly among the ten sectors (and not one to each sector). Then we can always gather any 9 chips of the 10 in, say, the zero sector. One chip keeps traveling clockwise, enabling each of the others to move counterclockwise, step by step, until the other chips each reach the targetted sector. Then the tenth chip ends up in sector S, where S is the value of our invariant for the initial arrangement. It follows that all arrangements fall into 10 classes corresponding to the 10 values of S such that two arrangements can be transformed into each other if and only if they are in the same class (have the same value of S).

11. (a) Replace the signs "+" and "–" with the numbers +1 and –1, respectively. Consider three squares $A_1A_4A_7A_{10}$, $A_2A_5A_8A_{11}$, and $A_3A_6A_9A_{12}$ (fig. 13). Any isosceles triangle inscribed in the given 12-gon is either a right triangle inscribed in one of the squares or has exactly one common vertex with each square. So the products of the numbers at the vertices of each square change their signs under any allowed operations simultaneously. But in passing from one of the given arrangements to another, only two of the three products should be changed, so this transition is impossible.

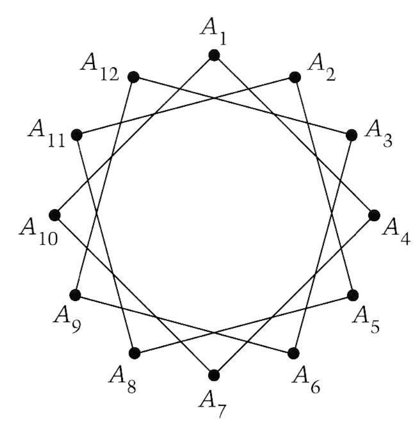


Figure 13

(b) Changing the signs at the vertices of three isosceles right triangles $A_{12}A_3A_6$, $A_6A_9A_{12}$, and $A_9A_{12}A_3$ results in changing only the sign at A_{12} and yields plus signs everywhere. After that, we similarly change the sign at A_1 only, thus creating the required arrangement.

12. The character can take any value from 0 to 4. Changing signs along columns that have a minus sign in the top place, we get plus signs in the top row. Then we change signs in the rows that have more than two minus signs. Now minus signs occur only in the three bottom rows, no more than two minus signs in each. If the total number of minus signs is 5 or 6, two of these rows have exactly two minus signs and we can change signs in both, one, or none of them to get a column having three minus signs without changing the total number of minus signs. Finally, we change signs in this column, decreasing the total number of minus signs by two, arriving at no more than 4 minus signs. And this number, in general, can't be diminished. Among other possibilities, the reader may examine the situation starting with four minus signs along a diagonal of the array and plus signs everywhere else.

13. Number the places occupied by the chips around the circle from 1 to 30. Any allowed transposition swaps the chips in places with numbers of the same parity. Consider the 15 "odd" places. Join place 1 to 5, 5 to 9, ..., 25 to 29, 29 to 3, ..., 27 to 1—that is, those pairs of places where "transposable" pairs of chips are positioned. We get a closed 15-sided (self-intersecting) polygonal curve. And two chips adjacent along this curve can be swapped. So any two chips in odd places can be

swapped (see the solution to exercise 4). It follows that any permutation of these chips is feasible. Likewise, any permutation of the chips in even places is feasible as well. This means that two arrangements of chips are equivalent if and only if they have the same number of white chips in the odd places (and therefore in the even ones, too). This number can take 11 values (0, 1, ..., 10), so 11 is the number in question.

14. The operations in the problem consist of two transpositions each (a transposition of the two extreme numbers of a quadruple and a transposition of the two inner numbers). So they can generate only even permutations. But the permutation we need to create is odd, because it consists of an odd number (997) of transpositions: (1, 1993), (2, 1992), ..., (997, 998). In general, the answer will be no for any number of the form $4k + 2$ (as well as $4k + 3$) instead of 1994.

For numbers of the form $4k$ or $4k + 1$, the inverse order is an even permutation, and the consideration of parities is not sufficient to answer the question. In fact, it's not difficult to show that here the answer is yes. A more involved analysis proves that our operations generate any even permutation.

Toy store

1. The number of arrangements for a given $(n_1, \ldots, n_{12})$ is equal to $24!/(n_1!n_2!\ldots n_{12}!)$. This also accounts for the various locations of empty space. The number of possible sets $(n_1, \ldots, n_{12})$ is $34!/(11!23!)$.

2. Let's put a chip numbered 0 on the empty space. Then every move becomes an exchange of the zero chip with some other chip. Notice that during such an exchange the orientation of the grid's triangle where the zero chip is located turns to the opposite direction (if it was pointed north, it will point south, and vice versa). Since the zero chip ends up at its starting location, the orientations of the triangles that it passed must have changed an even number of times, so the number of moves was even. But evenly many pair exchanges make an even permutation.

Conversely, the operation described in figure 5b in the article, and its modifications, allow us to create any desirable arrangement of all chips, with the possible exception of the two located at the opposite vertices of the central hexagon of the network in figure 4 (they correspond to chips a and b in figure 5b). These last two chips may or may not end up being swapped in the end. Suppose they are. This would mean that we have gotten an even permutation (the empty space came back to its starting position!), which differs from the required (also even) permutation in exactly one pair exchange—that of the last two chips. But this is impossible, because a pair exchange alters the parity of a permutation.

Corrections

Vol. 3, No. 6:

p. 6, col. 1, l. 11: *for* $1 - 1/t$ *read* $1 - 1/e$.

p. 7, col. 1, l. 22: *for* $i = n$ *read* $i = k$.

p. 7, col. 2, ll. 1–4: the form of the equation that "telescopes" is

$$S_k - S_{k-1} = 3/2 - t_{k,\,k-1}.$$

p. 56, table in M86: *for* D_2 *read* D^2; *for* D_3 *read* D^3.

Readers write ...

The article "Superheated by Equations" in the last issue of *Quantum* was read with pleasure by Dr. Richard Grant of the Albert Einstein Medical Center in Philadelphia. In fact, it reminded him of a natural phenomenon. He writes: "The teaser below the cover art on the contents page suggests that Dmitry Fomin's technique of heat exchange may be impractical. As a matter of fact it is used all the time in nature by having veins and arteries entwined in 'retia mirabiles.' These allow heat exchange exactly as described by Fomin (though of course not 100% efficient)." He adds that the mechanism works for diffusible substances (for example, oxygen or urea) in addition to heat.

Dr. Grant offers a citation for further reading on the subject: "The Wonderful Net" by P. F. Scholander in the April 1957 issue of *Scientific American*.

Index of Advertisers

TOY STORE

Topsy-turvy pyramids

And fishy chips

by Vladimir Dubrovsky

THE RENOWNED "15" PUZZLE created by Sam Loyd—which is said to have caused a craze that can be compared only to the worldwide success of Rubik's cube—gave rise to countless modifications.[1] As you probably know, Loyd's original invention is just a flat square box measuring 4 × 4 with 15 numbered unit-square pieces in it. The remaining empty unit square is used to slide the pieces one by one all around the box. Thus, the initial order of the pieces is completely destroyed, and the task is to slide them back to their initial positions. More generally, you may be asked to transform one given order into another—which, by the way, is not always possible.

The idea of "unscrambling a scrambled order of like pieces" that lies at the core of Loyd's toy has been continually used and developed ever since it was first introduced by the famous puzzle master. And a new dimension is added to it (literally!) by reshaping the flat pieces into blocks that can be rolled in the box rather than slid. The state of the puzzle then becomes dependent not only on the location of the pieces but on their orientations in their respective "cells" of the box (just as with the small blocks constituting Rubik's cube). The two most natural shapes for the blocks are cubes and pyramids (more exactly, regular tetrahedrons). Surprisingly, these two offspring of the slide-block family of puzzles turn out to be completely different in their properties and solutions. This article is devoted to rolling pyramids, which seem to be more distant cousins of the square ancestor . . . but let's not put the cart before the horse.

The version of the pyramids seen in figure 1 was invented by two designers from Krivoy Rog, a city in Ukraine. In fact, they have created a bunch of rolling-block puzzles and other kinds as well and won first prize in a puzzle contest sponsored by a popular Soviet newspaper in the mid-eighties. This beautiful toy consists of a hexagonal box with a grid of 24 triangular cells and 23 pyramids that fit exactly inside the cells; one empty cell is left for rolling the pyramids around. All the pyramids are identical, and each is colored "vertexwise": one of four different colors is used for each vertex. The figure illustrates their initial "regular" arrangement as proposed by the authors. As is typical of its country of origin, this toy (like countless other inventions of much greater import) was never put into production, and I would guess it exists in only two or three copies made by the designers. (It's quite easy to make yourself—if you have the enthusiasm, time, and patience.)

I had a chance to play with one of these rarities and was astonished at a number of mathematical subtleties concealed in it. The size and shape of the box, the coloring of the pyramids, and their initial order—all these things matter and were chosen by the inventors very aptly, whether by serendipity or as the result of a thorough search, or both.

Actually, I didn't really *play* with the pyramids, because after rolling them around for a while I clearly understood that without a carefully thought-out, long-range plan, every move you make in trying to approach the regular position only takes you further from it. This puzzle simply forces you to treat it as a mathematical problem, so let's do a bit of research together. And let's begin, as usual, with an experiment.

Take, or make, or just imagine one of the 23 pyramids on an infinite "box" with a triangular grid. Remember its initial position (location and orientation) and roll it along an arbitrary closed path. Compare its position when it comes back to the starting point with the initial one, and repeat this for several different routes. The result is rather unexpected: the pyramid always ends up in the initial position! This fact isn't self-evident and isn't valid for other regular polyhedrons. But it

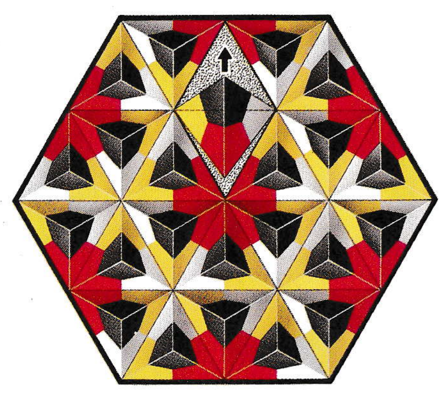


Figure 1

[1]See "A Portrait of Three Puzzle Graces" in the November/December 1991 issue.

becomes pretty clear if the plane is colored as in figure 2. Indeed, if you put our colored pyramid on any of the triangles of the colored plane so that their colorings match each other and roll it over, the colors on its new base will coincide with the colors on the plane. This means that the colors will coincide after any number of any rollings. But the colors on the pyramid's base uniquely determine its overall orientation. It follows that *the initial position (location and orientation) of a pyramid uniquely prescribes its orientation at any other location after it rolls there, no matter what its route.* (In particular, every time the pyramid returns to the starting triangle, it resumes the initial position.)

Notice that the colored triangles in figure 2 fall into eight classes such that any two triangles from one class can be obtained from each other by a parallel translation. And every triple of the four possible colors occurs in exactly two classes—with triangles pointing "north" in one of them and pointing "south" in the other. (Pairs of these classes are marked in the figure with white and black circles.) Accordingly, a pyramid rolling over the plane belongs permanently to one of these eight classes.

Now imagine a second pyramid whose colors do not coincide with those of the grid's triangle on which it initially rests. Then it will be impossible to make the colors coincide by any sequence of rollings (otherwise we would have had a permanent coincidence of colors). So we'll be always able to tell which of the two pyramids is the "first" and which is the "second," although they look exactly the same when removed from the plane. Let's call two such pyramids on the grid *distinguishable*. If the grid is not colored we can roll two pyramids, one after the other, onto the same triangle of the grid and compare their orientations. Clearly, the pyramids are distinguishable if and only if their orientations are different when they are sitting on the same triangle. Since a pyramid can be set on a given triangle in $4 \cdot 3 = 12$ ways (any of its four faces can be its base, and it can be turned in three ways on the base), the maximum number of distinguishable pyramids is also 12.

Now we turn back to our puzzle to inspect the orientations of the pyramids in the regular position (fig. 1) and find which of them are indistinguishable. All of them rest on their white–red–yellow faces. On the other hand, in figure 2 the white–red–yellow triangles constitute two (of the eight) classes, so all the pyramids in the box that are not distinguishable from a given one can have only two orientations—exactly the same as the given pyramid or rotated by 180°. All such pyramids make a pattern like the one shown in figure 3. If we lay this pattern over figure 2, we'll see that it always covers exactly one triangle of each of the eight classes! (This is one of the reasons I think the shape of the box and the regular arrangement of pyramids are so neatly chosen.) So each pyramid in the box can have only one indistinguishable counterpart, which is, in fact, symmetric to it about the center of the box (in the initial position). Of course, one pyramid—symmetric to the empty space—doesn't have a counterpart at all.

At this point we can dramatically simplify our puzzle to get rid of orientations—and even the pyramids! Let's number the pyramids in the initial position so that each indistinguishable pair gets its own number from 1 to 12 (one of the numbers will have to be given to a single pyramid), and assign the same numbers to the respective cells of the box—say, as in figure 4. Note that all twelve possible orientations are found in the regular arrangement—another indication of good design. These numbers help greatly in solving the puzzle, because they show where every pyramid must be rolled; and when it comes to the place with its number, it will *automatically* turn the right way. Then what do we need all this rolling for? We can replace the pyramids with numbered

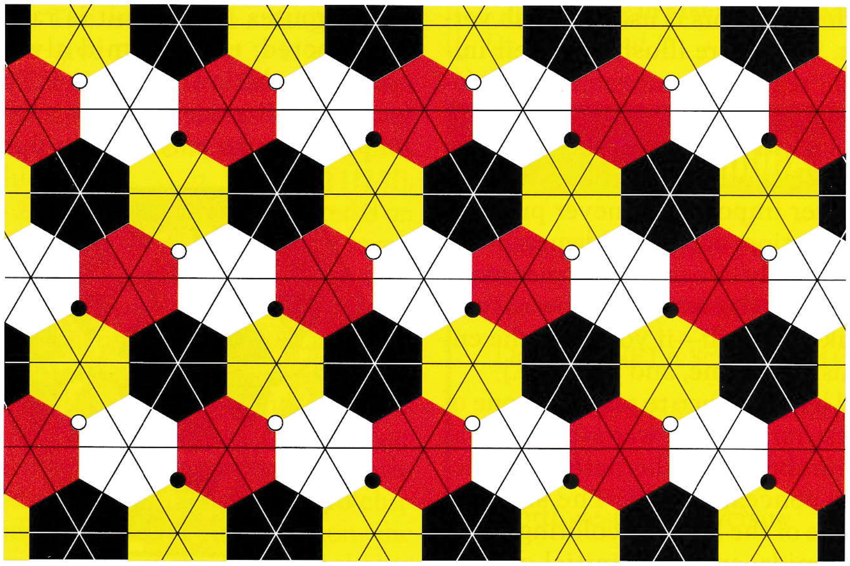


Figure 2

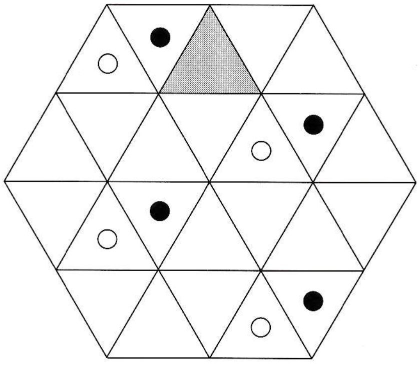


Figure 3

The pyramids with the same orientation in the regular arrangement of figure 1 form a pattern like the one shown with black circles; white circles mark the locations of pyramids whose orientation is obtained by a 180° turn from the black-circled one.

chips, and, instead of rolling, slide them from triangle to triangle (or along the lines forming the hexagonal network in figure 4, which is the same, of course).

So our puzzle is finally reduced to something very similar to the 15 puzzle. However, if you make it—which is much easier (draw the network and use a set of checkers as chips)—and play with it, you'll see it's more interesting than its prototype. At first the chips will slide to their destinations without any noticeable resistance. But gradually you'll feel that it gets harder and harder to drive them where you want without serious damage to the order already achieved. And you may get stuck at the very end. Figure 5 provides a universal recipe for constructing any permutation of chips you might need. Imagine that the left hexagon in figure 5 is the central hexagon of our network (in figure 4). Suppose also that the chips a and b have already been installed correctly. Then their exchange makes no difference, since these chips would have the same numbers. So the operation illustrated in figure 5b appears to exchange only c and d. To exchange two other adjacent elements—say, e and f—we must first shift them to places c and d (using a cyclic move like the one in figure 5a), then perform our operation, and move them back. Thus, we can exchange any two adjacent chips in the right hexagon and, therefore, any two chips along this hexagon (see "Some Things Never Change" in this issue, where it is shown that any permutation can be achieved by repeated exchanges of adjacent elements). Any permutation is a combination of a number of pair exchanges.

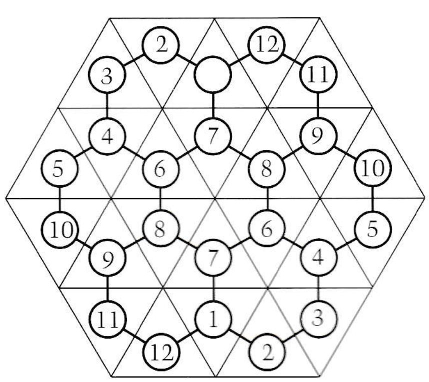


Figure 4

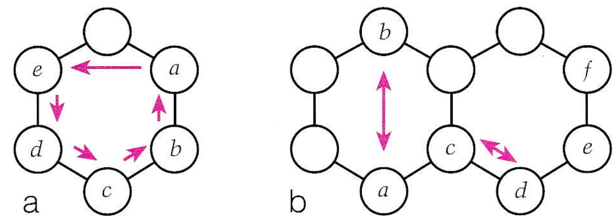


Figure 5

(a) Successively sliding chips a, b, c, d, e, a, we obtain their cyclic permutation, shown by the red arrows; (b) let L and R be the counterclockwise cyclic permutations of figure (a) for the left and right hexagons, respectively, L^{-1} *and* R^{-1} *the inverse (clockwise) cycles,* $L^2 = LL$ *and* $L^{-2} = L^{-1}L^{-1}$ *the cycles L and* L^{-1} *repeated twice. Then the operation* $R^{-1}LRL^{-2}R^{-1}L^2RL^{-1}$ *yields two simultaneous pair exchanges* $a \leftrightarrow b$ *and* $c \leftrightarrow d$.

Conluding our investigation of the chip (and also cheap) version of the puzzle, we can say that any arrangement of the chips numbered 1 to 12 such that each number except one occurs on two chips can be reordered in a standard way (like, say, figure 4) by a certain sequence of moves and, therefore, can be obtained from the standard arrangement (by reversing this sequence). It follows that any two arrangements can be transformed into each other (via the standard one, for instance)—that is to say, any two arrangements are *equivalent*.[2] But this is not true for arrangements of pyramids! Indeed, when replacing pyramids with numbered chips, we must take into account which of them are distinguishable (getting different numbers) and which are not (getting the same numbers). Assign some number from 1 to 12 to each of the 12 possible orientations of a pyramid on a certain triangle of the grid. This will uniquely define the concordant (with respect to distinguishability) numbering of orientations at any other triangle. Then each arrangement defines a set $(n_1, n_2, \ldots, n_{12})$ of 12 numbers, where n_k is the number of pyramids in this arrangement that have the orientation number k at their respective locations.

Problem 1. Show that two arrangements of pyramids are equivalent if and only if they have the same sets $(n_1, \ldots, n_{12})$. Find the number of arrangements with a given set (it depends on the values of n_i) and the number of all possible sets. [Hint: two *Quantum* articles will help you count these numbers: "Summertime, and the Choosin' Ain't Easy" (July/August 1992) and "Combinatorics–polynomials–probability" (March/April 1993)].

Problem 2. Suppose the chips in figure 4 were all numbered differently—from 1 to 23. Prove that any sequence of moves that returns the empty space to its initial location produces an even permutation of the chips (that is, a permutation generated by an even number of pair exchanges—see "Some Things Never Change" for more details). Conversely, any even permutation can be obtained in this way.[3]

You may have noticed that I've repeatedly referred to "Some Things Never Change." This is no accident. The rolling pyramids, and many related puzzles, fit perfectly the general idea of the problems considered in that article: given a set of some "configurations" of an arbitrary nature that can be transformed according to some fixed rules, find out when and how one configuration can be turned into another. Almost all the problems in that article can be viewed as "transformational puzzles" as well.

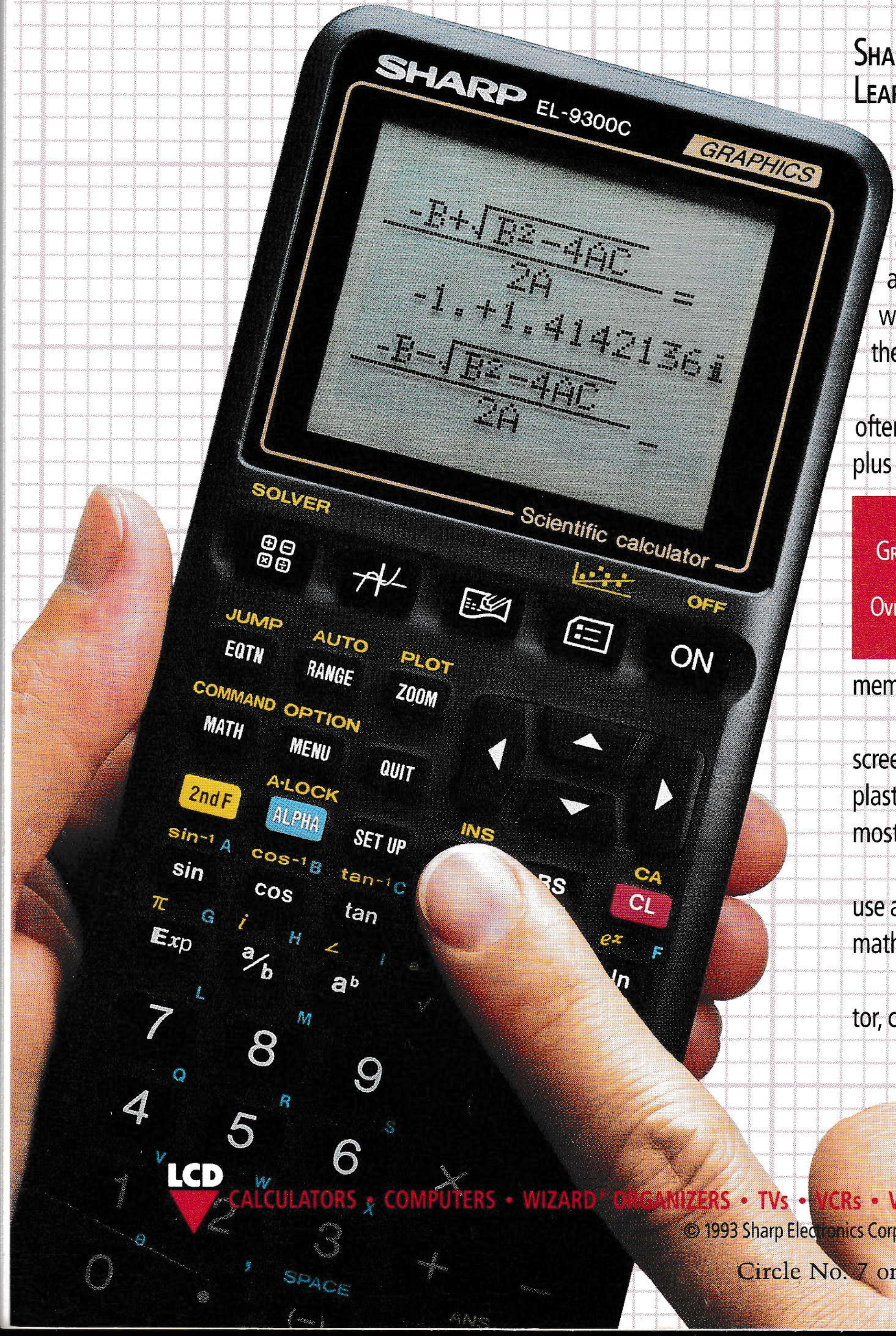
By the way—based on your reading of both articles, can you tell what the invariants of the rolling pyramids are?

ANSWERS, HINTS & SOLUTIONS ON PAGE 62

[2] Compare this with problem 13 in "Some Things Never Change."

[3] This is also true for the 15 puzzle (see "A Portrait of Three Puzzle Graces").

NOTHING MAKES PYTHAGORAS LESS GREEK TO YOUR STUDENTS.



SHARP'S GRAPHING CALCULATORS MAKE TEACHING AND LEARNING MATH EASIER THAN EVER. Whether it's the Pythagorean theorem or advanced calculus, you need a graphing calculator that doesn't create problems. Only Sharp's EL-9200C and EL-9300C graphing calculators have features that make it so easy to learn and understand math. Like the unique Equation Editor which lets users enter, edit and view equations just as they appear in text books.

Direct Access Keys allow one-touch access to the most often used functions. And the unique two-level menu system plus pop-up windows and error messages easily guide the user through operations.

NOW BUY 60 GRAPHING CALCULATORS AND GET A FREE OVERHEAD PROJECTOR KIT (A $399 VALUE)*

What's more, the EL-9300C has a Solver function that lets you solve for any variable in any equation. It even has a 32K RAM memory, communications port and back-up battery.

Both models offer incomparable high-contrast LCD screens for the brightest, clearest displays. Plus high-impact plastic cases with slide-on hard covers making them the most rugged calculators in their class.

Sharp's graphing calculators. When it comes to ease of use and durability, you won't find a better solution to your math teaching needs.

For a free brochure or your nearest educational distributor, call: **1-800-BE-SHARP.**

SHARP®
FROM SHARP MINDS COME SHARP PRODUCTS™

CALCULATORS • COMPUTERS • WIZARD® ORGANIZERS • TVs • VCRs • VIEWCAM™ CAMCORDERS • LCD PROJECTION SYSTEMS

 *Call your distributor for program details. Actual Screen Image

Circle No. 7 on Reader Service Card