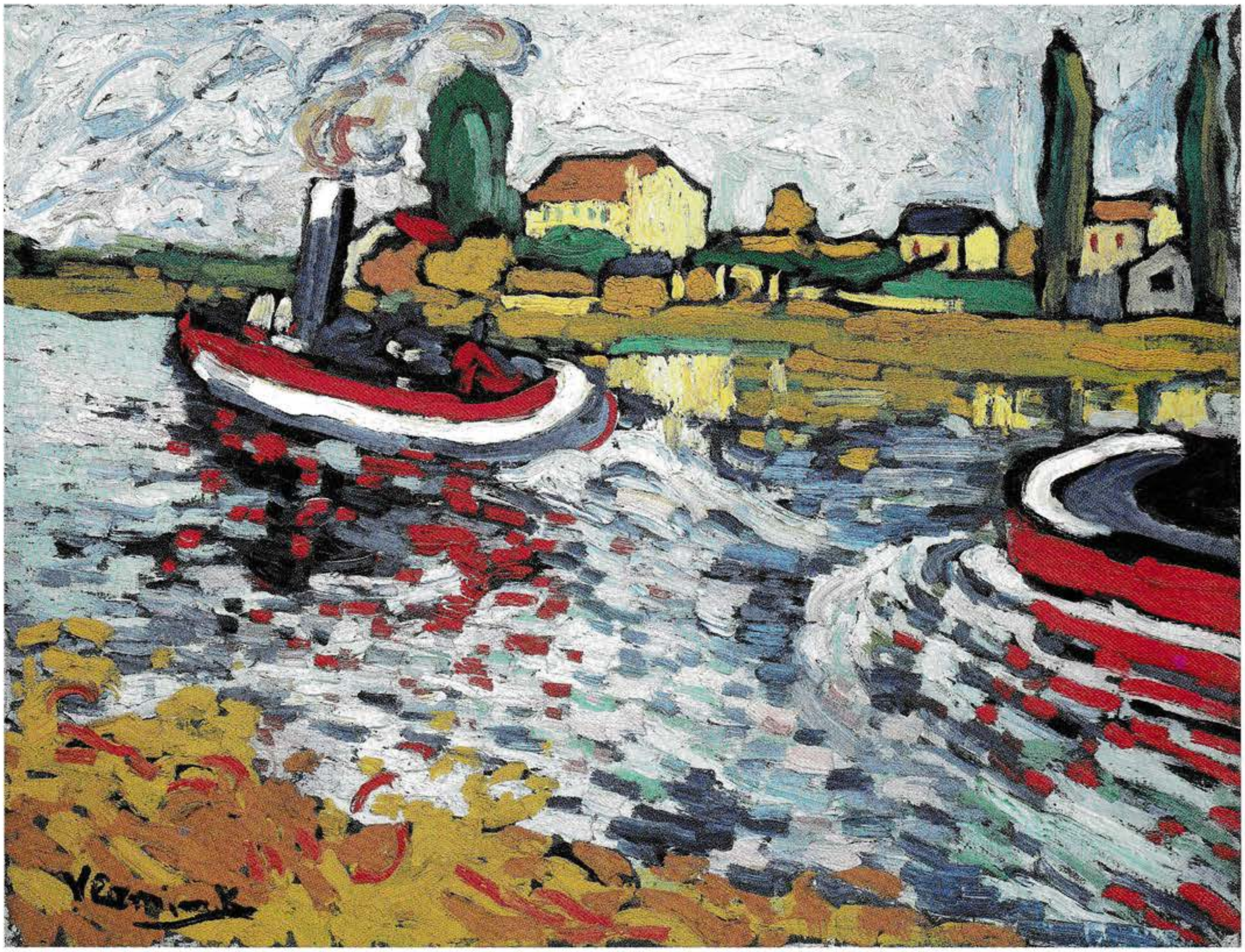
QUANTUM

JANUARY/FEBRUARY 2001

US $7.50/CAN $10.50



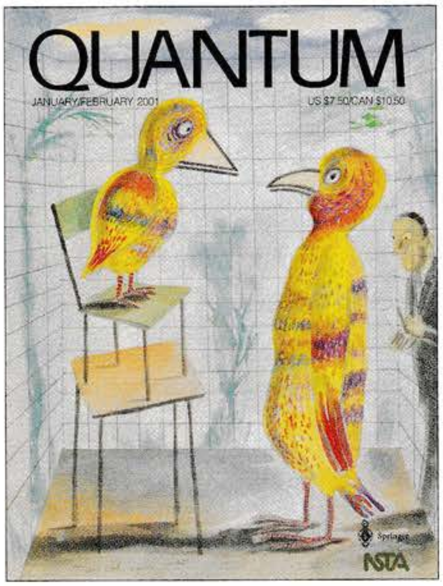
Oil on canvas, 19 3/4 × 25 5/8, Collection of Mr. and Mrs. John Hay Whitney, ©2000 Board of Trustees, National Gallery of Art, Washington, D.C.

Tugboat on the Seine, Chatou (1906) by Maurice de Vlaminck

THE "WILD BEASTS" OF THE FAUVIST MOVEMENT sought to capture the interplay of light and shadow with their rough-hewn brushstrokes. Whether it be the wake of a tugboat or the ripples stirred up by a sudden breeze, the fauvists were able to freeze the moment on canvas. As dedicated observers of the natural world, they may have wondered about a curious phenomenon while making their studies—the appearance of seemingly random patches of stillness on the rippling surface of a body of water. To find out the cause of these "islands of calm amidst stormy seas," turn to page 46.

QUANTUM

JANUARY/FEBRUARY 2001 VOLUME 11, NUMBER 3



Cover art by Leonid Tishkov

Birds of a feather flock together. Three articles in this issue—"Calculus and Inequalities," "Cauchy and Induction," and "The Sum of Minima and the Minima of Sums"—flutter about the common theme of inequality. Special attention is paid to a well-known formula proved by the French mathematician Augustin Louis Cauchy.

Indexed in *Magazine Article Summaries, Academic Abstracts, Academic Search, Vocational Search, MasterFILE,* and *General Science Source*. Available in microform, electronic, or paper format from University Microfilms International.

FEATURES

6 Electromagnetic Transformations
Flux and fixity
by V. Novikov

12 Extreme Science
The birth of low-temperature physics
by A. Buzdin and V. Tugushev

18 Create a Problem
Where do problems come from?
by I. Sharygin

DEPARTMENTS

3 Front Matter

5 Brainteasers

11 How Do You Figure?

28 Kaleidoscope
Dielectrical materialism

30 Physics Contest
A good theory

34 At the Blackboard I
The sum of minima and the minima of sums

37 At the Blackboard II
Cauchy and induction

38 At the Blackboard III
Calculus and inequalities

40 Looking Back
Ptolemy's trigonometry

42 Problem Primer
Problems teach us how to think

46 In the Lab
What happens at the boundary

49 Answers, Hints & Solutions

54 Crisscross Science

55 Informatics
IOI 2000

NSTA *Quantum* (ISSN 1048-8820) is published bimonthly by the National Science Teachers Association in cooperation with Springer-Verlag New York, Inc. Volume 11 (6 issues) will be published in 2000–2001. *Quantum* contains authorized English-language translations from *Kvant*, a physics and mathematics magazine published by Quantum Bureau (Moscow, Russia), as well as original material in English. **Editorial offices:** NSTA, 1840 Wilson Boulevard, Arlington VA 22201-3000; telephone: (703) 243-7100; e-mail: quantum@nsta.org. **Production offices:** Springer-Verlag New York, Inc., 175 Fifth Avenue, New York NY 10010-7858.

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

Subscription Information:
North America: Student rate: \$18; Personal rate (nonstudent): \$25. This rate is available to individual subscribers for personal use only from Springer-Verlag New York, Inc., when paid by personal check or charge. Subscriptions are entered with prepayment only. Institutional rate: \$56. Single Issue Price: \$7.50. 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. Mail order and payment to: Springer-Verlag New York, Inc., Journal Fulfillment Services Department, PO Box 2485, Secaucus, NJ 07094-2485, USA. Telephone: 1(800) SPRINGER; fax: (201) 348-4505; e-mail: custserv@springer-ny.com.

Outside North America: Personal rate: Please contact Springer-Verlag Berlin at subscriptions@springer.de. Institutional rate is US\$57; airmail delivery is US\$18 additional (all rates calculated in DM at the exchange rate current at the time of purchase). SAL (Surface Airmail Listed) is mandatory for Japan, India, Australia, and New Zealand. Customers should ask for the appropriate price list. Orders may be placed through your bookseller or directly through Springer-Verlag, Postfach 31 13 40, D-10643 Berlin, Germany.

Advertising:
Representatives: (Washington) Paul Kuntzler (703) 243-7100; (New York) M.J. Mrvica Associates, Inc., 2 WestTaunton Avenue, Berlin, NJ 08009. Telephone (856)768-9360, fax (856)753-0064; and G. Probst, Springer-Verlag GmbH & Co. KG, D-14191 Berlin, Germany, telephone 49 (0) 30-827 87-0, telex 185 411.

Printed on acid-free paper.

Visit us on the internet at www.nsta.org/quantum/.

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 mission of the National Science Teachers Association is to promote excellence and innovation in science teaching and learning for all.

Not a silver bullet—a golden opportunity

How technology can improve science education

THE RISE OF ONLINE TECHnologies is providing boundless opportunities for students and teachers of math and science in ways that were inconceivable 50 years ago. Thanks to the Internet, student learning is no longer confined by the walls of the classroom. Students in different classrooms are working together on hands-on experiments, are conversing real-time with experts from NASA, and are following adventurers on their scientific treks through the jungle and into the deepest reaches of the ocean.

One exciting example—and a website that should be bookmarked by all students and teachers—is *www.TryScience.org*, created through a unique IBM partnership with the New York Hall of Science and the Association of Science and Technology Centers (ASTC). TryScience provides dynamic on- and off-line science experiences for young people, offering instant access to the expertise and resources of more than 450 of the world's leading science centers. Through interactive exhibits, experiments, virtual field trips, and scientific research, this continuously updated site challenges and engages students in ways that spark imaginations and stir brain cells.

Still, IBM knows that, while unprecedented learning experiences are under way in many classrooms throughout the nation, new technologies alone won't be enough to push math and science learning forward for *all* students. As promising as technology is, it is not a silver bullet. A computer, for instance, can never replace a good science teacher, and Internet access and computer labs alone cannot improve mathematics instruction.

As we all are too well aware, students in the United States are not performing nearly as well as their international peers in a range of mathematic and scientific areas. Most recently, in the Third International Mathematics and Science Study (TIMSS), U.S. eighth graders performed in the bottom half in mathematics and only average in science compared to their peers in other developed countries. U.S. twelfth graders performed even more dismally, scoring third from the bottom in mathematics literacy, second to last in advanced mathematics, and at the very bottom in physics. These results are alarming, given research showing that children who take and pass algebra and other academic math and science courses experience greater success later on in school. To turn this situation around, we need to address weaknesses in our curriculum, which TIMSS researchers described as "a mile wide and an inch deep," and make a fundamental commitment to teacher quality in ways that integrate the best that technology has to offer.

As part of our deep commitment to education, IBM has been working hard to improve math and science education in a host of ways. Through our award-winning Reinventing Education grant program, we are contributing over $40 million to school partners in the United States and in seven countries throughout the world to develop and implement innovative technology solutions designed to drive higher student achievement and enhanced academic productivity. Each of our projects is working to overcome a specific barrier to school reform, with a number of our sites specifically addressing math and science education.

For example, working with teachers in Rochester, Minnesota, IBM has created Visual Venture, software that integrates math and science curriculum to help students in the middle grades gain a true command and appreciation of mathematical models by using them to investigate the world around them. The software features innovative image processing technology that simulates real laboratory experiments, emphasizing problem solving, inquiry-based learning, and hands-on experiences.

Through Reinventing Education, IBM also has worked with West Virginia educators statewide to use resources from the Internet to design exemplary online instructional activities for high school students. The activities focus on four core subjects, beginning with mathematics and expanding to science, language arts, and social studies, with concentrated attention given to the academic areas in which students have shown they need the most help and practice. To ensure high quality, each activity is subject to a rigorous, online jurying process before being posted on a special site on the World Wide Web for use by teachers and students throughout the state.

An independent evaluation from the Center for Children and Technology found that students who frequently use these lesson plans *scored significantly higher* on state math tests than students who used them only sometimes or not at all. This technology is now being used by over 40,000 teachers around the world to bring curriculum online in exciting ways that continue to boost student achievement.

While new curriculum is critical, IBM knows that it is perhaps even more important to give teachers the training and support they need to help students achieve at higher levels. Through a Reinventing Education partnership with the Chicago Public Schools, IBM has created Learning Village, technology which is designed to raise the quality of mathematics instruction. Learning Village provides seventh and eighth-grade teachers with instant access to new, hands-on algebra, geometry, and data analysis curricula developed by IBM curriculum specialists, as well as to off- and online resources, including a "math dictionary" that is being incorporated into lessons. Included within Learning Village is a password-only site that serves as the district's newest vehicle for professional development. The site provides online discussion sites that enable staff to collaborate on new teaching techniques and strategies, discuss mathematical content and pedagogy, and share information on problems and best practices.

IBM's leadership in math and science education extends on a policy level through Achieve, Inc., the nonprofit that resulted from IBM's leadership on the 1996 and 1999 National Education Summits and that IBM Chairman and CEO Louis V. Gerstner Jr. co-chairs. Achieve is involved in a major initiative, with the help of The College Board, to develop and administer a new standards-based math test for eighth-grade students that will compare results school by school and state by state—and will include curriculum and professional development assistance for teachers. Eleven states have agreed to participate, and IBM continues efforts to get other states on board.

IBM's leadership in information technology depends on this work, as well as our other efforts to raise the quality of math and science education and of public education more broadly. For IBM to remain at the forefront of innovation and discovery, we know that we must provide cutting-edge technologies to schools that open new doors to learning, in tandem with initiatives that get at the heart of necessary changes in math and science education. Only in this way can our schools and teachers prepare our young people to become the next generation of scientists and researchers, responsible citizens, and visionary leaders.

To learn more about IBM's efforts in education, please visit *www.ibm.com/ibm/ibmgives*.

—Stanley Litow

Stanley Litow *is Vice President, Corporate Community Relations, and President, IBM International Foundation.*

HAPPENINGS

CyberTeaser winners

YOU DIDN'T NEED ANY HIGH-tech tools to determine which musketeer placed where in the big fencing competition. Maybe a pencil and paper were enough. Maybe you even did it in your head. But to win a prize, you needed to visit our website and submit a correct answer. Did you? If not, you have another chance: the next CyberTeaser awaits at *www.nsta.org/quantum*.

Here are the ten who slashed their way to victory:

Jerold Lewandowski (Troy, New York)

Jason Langley (Lawrenceville, Georgia)

Oleg Ivrii (Toronto, Ontario)

Aris Noutsos (Tessaloniki, Greece)

Bruno Konder (Rio de Janeiro, Brazil)

Theo Koupelis (Wausau, Wisconsin)

Gottfried Perz (Graz, Austria)

Andrey Meshkov (Columbus, Ohio)

John Beam (Bellaire, Texas)

Po-Ling Loh (Madison, Wisconsin)

Congratulations! Each of you will receive a copy of this issue of *Quantum* and the classic *Quantum* button. In addition, one swimmer in the pool of successful respondents receives a special flotation device: a copy of *Quantum Quandaries*, the reading material of choice for ferry rides and slow moments anytime, anywhere.

BRAINTEASERS

Just for the fun of it!

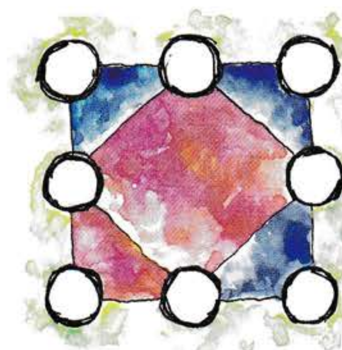
B311

Plumbing their passion. The founder of the Czech state, princess Libusha, was being wooed by three princes. She invited them to solve the following problem: "If I gave half the plums in this basket plus one more plum to one prince, half of the remainder plus one more plum to the second prince, and half of what remains plus three more plums to the third, the basket would be empty. How many plums are in the basket?"

B312

Musketeering swordplay. At the royal fencing competition in France, the first four places were taken by Athos, Porthos, Aramis, and D'Artagnan. The sum of the places taken by Athos, Porthos, and D'Artagnan was 6; the sum of the places taken by Porthos and Aramis was also 6. What was the place taken by each of the musketeers if Porthos ranked higher than Athos?

B313



A dozen and almost. Put the numbers 1 through 8 in the circles such that the sum of the numbers at the vertices of each blue triangle is 12 and the sum at the vertices of the red triangle and the red square is 11.

B314

Bir-to-bir correspondence. Solve the following number rebus. Identical letters correspond to identical digits, while different letters correspond to different digits.

(This puzzle comes to us from Azerbaijan, one of the Caucasian republics. In their language, the word BIR means one and DÖRD means four.)

B315

Spoon shadow. A sunbeam is reflected on the wall by a cup of tea. If we lower a teaspoon toward the center of the surface of the tea, what shape will the shadow take in the reflected spot on the wall?

ANSWERS, HINTS & SOLUTIONS ON PAGE 52

Art by Pavel Chernusky

Flux and fixity

What is the energy stored in a magnetic field?

by V. Novikov

THE FACT THAT A MAGnetic field generated by an electric current has energy can be seen from an analysis of electromagnetic induction. Let's carry out an experiment according to the diagram shown in figure 1. If we move

Figure 1

the contact of the potentiometer to decrease its resistance, the current in the circuit will increase. The higher the current, the larger the magnetic field generated by this current. An increase in the magnetic field means an increase in the magnetic flux passing through the surface encircled by the circuit. According to Faraday's law, the change in magnetic flux generates a self-induced emf in the circuit, which is equal to

$$\mathcal{E} = -\frac{\Delta\Phi}{\Delta t}, \qquad (1)$$

where $\Delta\Phi$ is the change in the magnetic flux over a small time interval Δt. This emf induces current, which in turn generates a magnetic field directed counter to the original magnetic field. Thus the magnetic flux of the new field "tries" to compensate for the change in the primary field; or, in other words, it does its best to maintain the original current at a constant level. To overcome this effect, one needs to perform some work against the self-induced emf. Let's calculate this work.

If the increase in current from I to $I + \Delta I$ is accompanied by the appearance of a self-induced emf $\mathcal{E}$, the energy source (battery) performs some work against this emf:

$$\Delta W = -\mathcal{E} I \Delta t.$$

Plugging equation (1) for the self-induced emf into this equation yields

$$\Delta W = \frac{\Delta\Phi}{\Delta t} I \Delta t = I\Delta\Phi. \qquad (2)$$

Now recall that every circuit can be characterized by a certain constant value L measured in henries (the self-inductance). This is the coefficient of proportionality in the equation coupling the magnetic flux Φ with the current I that induced this flux:

$$\Phi = LI. \qquad (3)$$

So if the current increases by ΔI, the magnetic flux will increase by

$$\Delta\Phi = L\Delta I. \qquad (4)$$

Therefore, equation (2) for the work can be written as

$$\Delta W = LI\Delta I = \Phi\Delta I. \qquad (5)$$

Figure 2 shows the dependence of the magnetic flux Φ on the current I. The work ΔW is equal to the area of trapezoid $ABCD$.

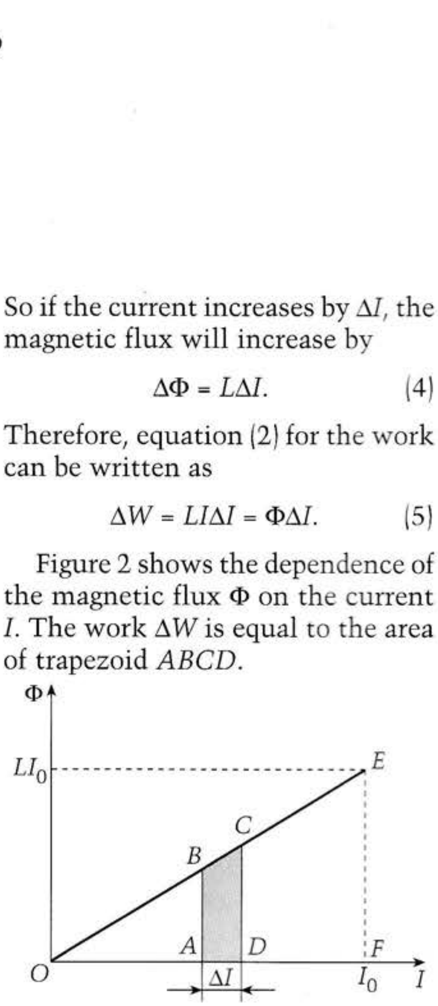
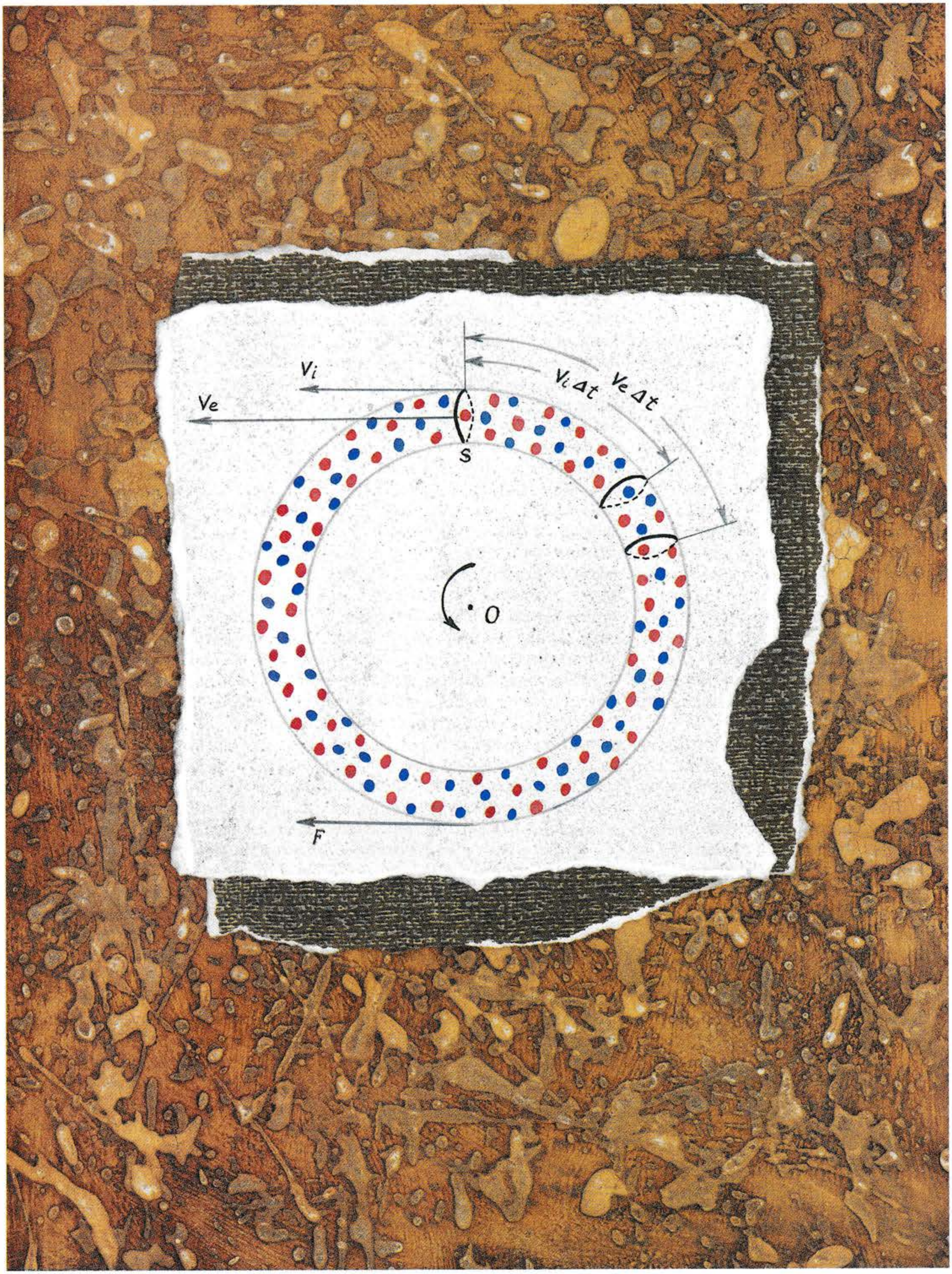


Figure 2

If the current in the circuit changes from $I = 0$ to some value I_0, the work performed by the electric battery against the self-induced emf can be obtained by summing all the ΔW. This total work is equal to the area of the triangle OEF (figure 2):

$$W = \frac{LI_0^2}{2}. \qquad (6)$$

Art by Vasily Vlasov


v_i
v_e
$v_i \Delta t$
$v_e \Delta t$
S
O
F

Now we should check whether any extra heat will be dissipated in the circuit due to the appearance of the self-induced emf.

Imagine that, before the beginning of our experiment, we deformed the circuit and reduced its area to zero (figure 3). In this simple

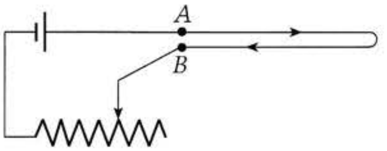


Figure 3

case no self-induced emf appears when the current increases from zero to I_0. The resistance of the deformed circuit remaines the same, so the same amount of heat is dissipated in the deformed circuit as in the original one. Of course, this is true if the rate of increase of the current is identical in both cases. Recall that in the first experiment we increased the current (by moving the contact to decrease the resistance) rather slowly, so for any value of R_t set by the potentiometer, the current had enough time to settle to the value determined by Ohm's law (we assume that the internal resistance of the source is negligibly small).

So the release of heat in the circuit doesn't depend on whether the source performs the additional work reflected in equation (6).

Everybody knows that energy doesn't disappear or appear out of nowhere. It can only be transformed from one type to another. Since the increase in the energy of a system is equal to the work performed, we might ask: What was the extra work of the electric battery turned into? The reason this work had to be performed is the ability of the circuit to "resist" any attempts to change (in our case, to increase) the magnetic flux through the area delimited by the circuit. Therefore, it's natural to call the energy into which the performed work was transformed *the energy of the magnetic field of the current-carrying circuit*. To convince ourselves of this, we might conduct an experiment that would demonstrate how the stored magnetic energy can be transformed into other forms of energy (say, into heat).

Let's consider one such experiment. Assume that points A and B of our circuits were instantaneously shorted. The voltage drop between points A and B created by the external current source will drop to zero. In the deformed circuit, the current stops flowing immediately. However, in the original, non-deformed circuit the current cannot disappear immediately because, according to the law of electromagnetic induction, any change in the current (in this case, a decrease) generates a self-induced emf that "tries" to maintain the previous level of the current. As a result, the decrease in the current in the original circuit from I_0 to zero will occur during a certain time interval. In this period some amount of heat will be dissipated in the circuit at the expense of the magnetic energy "stored" in the space around this circuit. (Try to show on your own that the amount of heat released after shorting the circuit between points A and B equals the magnetic energy $LI_0^2/2$ accumulated by this circuit.)

So we can conclude that the current I flowing in the circuit characterized by the self-inductance L generates a magnetic field in the space near the circuit. The energy of this field is

$$U_m = \frac{LI^2}{2}. \qquad (7)$$

Our experiment illustrates the case where the energy of the magnetic field induced by the current was created by an external source of electrical energy (a battery). Clearly the energy of the magnetic field must not depend on how the field was generated, much as the kinetic energy of a moving object doesn't depend on how the speed was imparted to the object.

To make this point clear, let's consider another example that illustrates how it's possible to directly transform the kinetic energy of a moving body into the energy of a magnetic field. This example will help us understand how a magnetic field is generated by current at the "microscopic level," so to speak. By "microscopic" we mean that the appearance of the current will be explained by analyzing equations that describe the motion of electrons and ions.

Let's consider a metal ring of circumference l uniformly rotating about an axis passing through the center of the ring perpendicular to its plane (figure 4). We assume that the ring is sufficiently narrow so

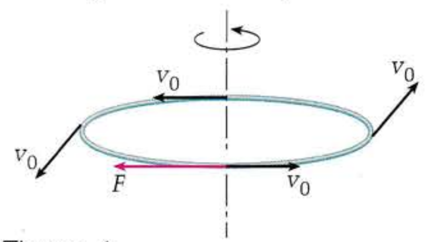


Figure 4

that the linear speed of all its segments is constant and equal to v_0. At time $t = 0$, a constant tangential braking force F is applied to the ring. The ring eventually stops rotating. However, the force F produces more than just a braking effect.

It's easy to see that the current must increase in the ring. Indeed, recall how solid metals are constructed. The atoms are ionized and located at the nodes of the crystal lattice, where they oscillate ever so slightly near the equilibrium position. It looks as if the crystal lattice forms the solid framework of the metallic body. In contrast to the ions, the conducting (free) electrons can move freely between the nodes of the lattice, forming a kind of "electron gas."

Now let's return to our metal ring. We can think of it as being composed of two rings, one inside the other (figure 5). In this figure blue corresponds to the crystal lattice, while red represents the electron gas.

At first, when the ring is rotating with constant speed, both the crystal lattice and the electron gas take part in this motion. However, the braking force F is applied only to the solid framework, not to the free elec-

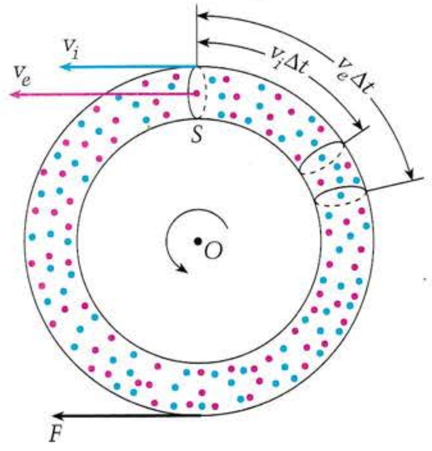


Figure 5

trons. The electron gas will be slowed only as a result of its interaction with ions in the lattice. This is the very process that underlies electrical resistance in the metal. For the sake of simplicity, we can neglect this interaction, proceeding as if, for example, the ring had been cooled to an extremely low temperature, at which metals lose their electrical resistance and become superconductors.

Therefore, the conducting electrons will pass the crystal lattice while the ring is braking. The difference in the speeds of the ions and the electrons generates a current. Let's calculate the value of this current.

Denote by v_i the speed of the ions in the crystal lattice and by v_e the speed of directed motion of the conducting electrons. The total charge flowing through a cross section of the ring during a time interval Δt consists of the charge of the crystal lattice confined in an arc of length $v_i\Delta t$ and of the electron gas located in an arc of length $v_e\Delta t$ (figure 5). If the total charge of all the ions that donated their electrons to the electron gas is Q, the total charge of the conducting electrons will be $-Q$, because the ring as a whole is electrically neutral. Therefore, the charges in the corresponding arcs are

$$Q\frac{v_i\Delta t}{l} \text{ and } -Q\frac{v_e\Delta t}{l},$$

where l is the circumference of the ring. Now we can find the total amount of electric charge flowing through a section of the metal ring per unit time—that is, we can calculate the current in the ring:

$$\begin{aligned} I &= \frac{1}{\Delta t}\left[Q\frac{v_i\Delta t}{l} - Q\frac{v_e\Delta t}{l}\right] \\ &= \frac{Q}{l}\left(v_i - v_e\right). \end{aligned} \tag{8}$$

In the analysis above we neglected the direct interaction of conducting electrons with the ions in the crystal lattice. However, in the braking process other forces come into play, which will slow the electrons even in a superconducting ring. These are forces generated by a vortical electromagnetic field.

When current arises in the ring, it generates a magnetic field. Since the strength of the current varies, this magnetic field is not constant. Thus it generates a vortical field in the ring. The forces of this vortical field affect the electrons and impede any change in the current in the ring. In other words, they act to decrease the relative speed of the electrons and ions (see equation 8). Therefore, these forces slow the electrons and accelerate the ions in the crystal lattice. The work performed by these forces in displacing a single positive charge along the ring is the self-induced emf (note that, in contrast to electrostatics, the work of a vortical field along a closed trajectory is not zero).

To picture the dynamics of the motion of the crystal lattice and conducting electrons (that is, the electron gas), it's convenient to consider them individually. Figure 6a shows the details related to the motion of the crystal lattice (the blue ring) and figure 6b does the same for the electron gas (the red ring). Now we can write Newton's second law for the crystal lattice and the electrons. Let the masses of the blue and red rings be M and m, respectively, and let their accelerations be a_i and a_e. Then

$$Ma_i = QE - F,\ ma_e = -QE, \tag{9}$$

where E is the strength of the vortical electric field, which we will now calculate.

Since the work performed by the vortical field to displace a unit charge along the ring is equal to the self-induced emf ($El = \mathcal{E}$), then

$$\begin{aligned} E &= \frac{\mathcal{E}}{l} = -\frac{1}{l}\frac{\Delta\Phi}{\Delta t} = -\frac{L}{l}\frac{\Delta I}{\Delta t} \\ &= -\frac{QL}{l^2}\left(\frac{\Delta v_i}{\Delta t} - \frac{\Delta v_e}{\Delta t}\right) = -\frac{QL}{l^2}\left(a_i - a_e\right). \end{aligned}$$

Thus equation (9) describing the motion of the crystal lattice and the electron gas will be transformed as follows:

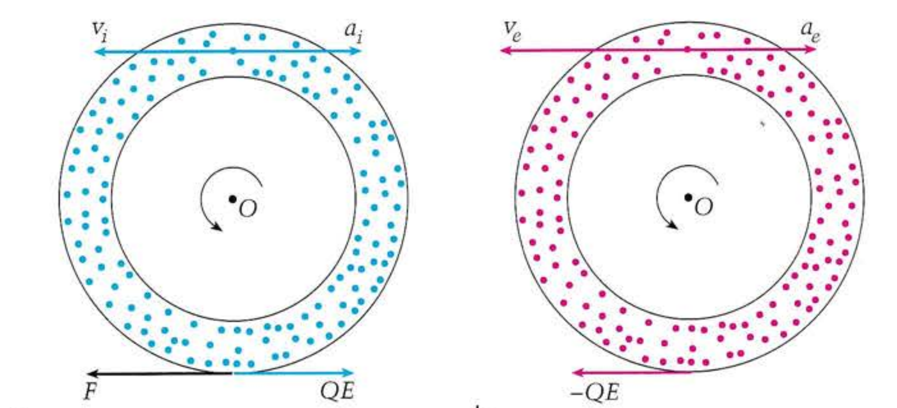
$$\begin{aligned} Ma_i &= -\mu(a_i - a_e) - F, \\ ma_e &= \mu(a_i - a_e), \end{aligned} \tag{10}$$

where for the sake of simplicity we introduced a parameter

$$\mu = \frac{Q^2L}{l^2}.$$

Solving system (10) for a_i and a_e, we get the equations for the acceleration of the rings:

$$a_i = -\frac{F}{M + \frac{m\mu}{m+\mu}};\ a_e = \frac{\mu}{m+\mu}a_i. \tag{11}$$



a

b

Figure 6

The minus sign means that both the crystal lattice and the electron gas are slowed during the braking process.

Using equations (11), we can find the time that passed from the start of braking to full stoppage of the crystal lattice:

$$t = -\frac{v_0}{a_i} = \frac{v_0}{F}\left(M + \frac{m\mu}{m+\mu}\right). \qquad (12)$$

Now it's easy to find the speed $v_e = v$ the electrons will have when braking of the lattice is complete:

$$v = v_0 + a_e t = \frac{m}{m+\mu} v_0. \qquad (13)$$

Thus the final value of the current in the ring (at $v_e = v$, $v_i = 0$) is

$$I = -\frac{Q}{l} v = -\frac{Q}{l}\frac{m}{m+\mu} v_0. \qquad (14)$$

Now we can return to the problem equationted above and find the energy of the magnetic field generated by the current induced in the ring during braking. Here we'll find it convenient to use the energy conservation.

Initially ($t = 0$), all the energy of the ring consisted of the kinetic energy of the electrons and crystal lattice—that is,

$$U_0 = \frac{mv_0^2}{2} + \frac{Mv_0^2}{2}.$$

During braking, a portion of the energy was used for work against the force F. Clearly this "loss" of energy is $\Delta U = W = Fs$, where s is the path traveled by any point of the ring (crystal lattice) during braking:

$$s = v_0 t + \frac{a_i t^2}{2} = \frac{v_0^2}{2F}\left(M + \frac{m\mu}{m+\mu}\right).$$

Therefore,

$$W = \frac{v_0^2}{2}\left(M + \frac{m\mu}{m+\mu}\right).$$

The current arising in the ring during braking generates a magnetic field. Consequently, at the end of the braking period, when the speed of the crystal lattice becomes zero, the energy of the ring consists of the kinetic energy of the electrons

$$K_e = \frac{mv^2}{2} = \frac{mv_0^2}{2}\left(\frac{m}{m+\mu}\right)^2$$

and the energy of the magnetic field U_m. According to the conservation of energy,

$$U_0 = K_e + U_m + W,$$

we obtain

$$U_m = U_0 - W - K_e = \frac{Mv_0^2}{2} + \frac{mv_0^2}{2} - \frac{v_0^2}{2}\left(M + \frac{m\mu}{m+\mu}\right) - \frac{mv_0^2}{2}\left(\frac{m}{m+\mu}\right)^2 = \frac{mv_0^2}{2}\frac{m\mu}{(m+\mu)^2}. \qquad (15)$$

Plugging the parameter μ into this equation and taking into consideration that

$$\frac{mv_0}{m+\mu} = -I\frac{l}{Q}$$

(see equation 14), we get

$$U_m = I^2 \frac{l^2}{Q^2}\frac{Q^2 L}{2l^2} = \frac{LI^2}{2}.$$

So we've arrived again at the equation we obtained earlier for the magnetic field energy of a circuit carrying current I.

In conclusion, we should make note of an interesting feature of the phenomenon we've been investigating. We know that the arithmetic average of two positive numbers cannot be smaller than their geometric mean—that is,

$$m + \mu \geq 2\sqrt{m\mu};$$

consequently,

$$\frac{m\mu}{(m+\mu)^2} \leq \frac{1}{4}.$$

Therefore, it can be concluded from equation (15) that

$$U_m \leq \frac{1}{4}\frac{mv_0^2}{2}.$$

Thus, no more than 25% of the initial kinetic energy of the directed motion of the conducting electrons can be transformed into magnetic energy by braking the revolving circuit. This restriction doesn't depend on the self-inductance L of the circuit.

Many's the time *Quantum* has demonstrated the inner relations of seemingly different phenomena. Now we see how the mechanical properties of a metal are linked to the magnetic field. ◘

Quantum on magnetic fields:

Kaleidoscope, "Electromagnetic Induction," March/April 1991, pp. 32–33.

A. Chernoutsan, "Michael, Meet Albert," September/October 1993, pp. 43–44.

S. Eatman, F. Muir, and H. Hickman, "Phlogiston and the Magnetic Field," March/April 1994, pp. 35–37.

J. Wylie, "Magnetic Monopoly," May/June 1995, pp. 4–9.

A. Ruzmaykin, "On the Nature of Space Magnetism," September/October 1995, pp. 12–17.

A. Stasenko, "Magnets, Charges, and Planets," May/June 1997, pp. 42–45.

A. Mitrofanov, "Can You See the Magnetic Field?" July/August 1997, pp. 18–22.

D. Tselykh, "Magnetic Fieldwork," September/October 1998, pp. 46–47.

A. Stasenko, "A Rotating Capacitor," May/June 1999, pp. 34–36.

V. Kartsev, "Magnetic Personality," May/June 1999, pp. 42–46.

HOW DO YOU FIGURE?

Challenges

Physics

P311

Rubber meridians. A strong rubber net envelops an elastic ball. The threads of the net lie along the meridians of the ball. What shape will the ball assume if the internal pressure is increased?

P312

Chopping drops. To obtain identical water drops, a capillary tube is connected to a large vessel of water equipped with a piston (figure 1).

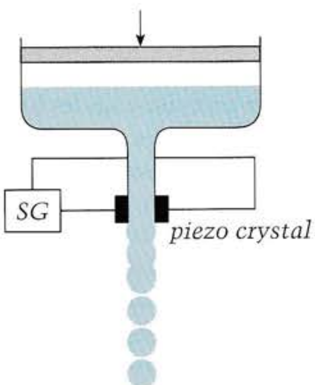


Figure 1

Water flows from the capillary when the piston moves in the vessel. A piezoelectric crystal is attached to the outer surface of the capillary's free end and connected to a low-frequency sound generator, which induces oscillations in the stream of water. At sufficiently large oscillation amplitudes, the stream is chopped into absolutely identical drops. Find the radius of these drops if the internal diameter of the capillary is 0.2 mm, the speed of the ejected water $v = 2$ m/s, and the sound frequency $f = 1000$ Hz.

(Y. Chernyshev)

P313

Moisture in a hydrogen tank. A cylinder of volume $V = 2$ liters contains hydrogen ($m = 2$ g) and a small amount of water. The pressure in the cylinder is $P_1 = 17$ atm. The cylinder is heated in such a way that the pressure rises to $P_2 = 26$ atm. How much water evaporates? What are the initial and final temperatures? *Hint*: Use a table giving data on the pressure-temperature relationships of water saturated vapor.

(A. Buzdin)

P314

Two coils and a capacitor. Two coils with self-inductance L_1 and L_2 are connected in parallel. What will be the maximum currents in the

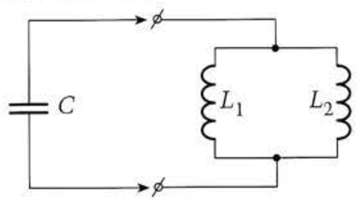


Figure 2

coils if a capacitor with capacitance C and charged to a voltage V is connected to them in parallel as shown in figure 2? (O. Savchenko)

P315

Conical reflector. A point source P is located on the axis of a hollow cone with a reflecting inner surface (figure 3). An image of P is formed on a screen S by rays making a single reflection from the cone and passing

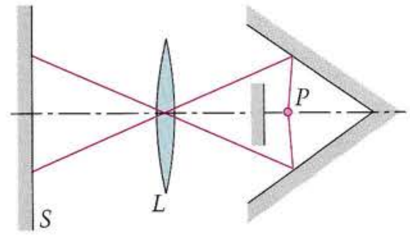


Figure 3

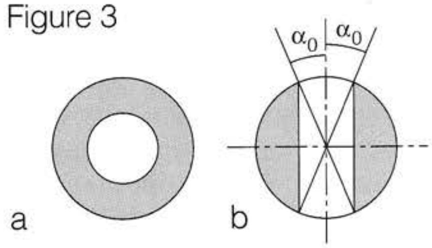


Figure 4

through a lens L. The rays cannot reach the lens directly, because they are blocked by a diaphragm. How will the image be transformed if the lens is screened with the diaphragms shown in figures 4a and 4b?

(D. Belov)

Math

M311

Internal identity. Prove that in any arithmetic progression consisting of natural numbers there exist two numbers with identical digit sums. (S. Genkin)

M312

Triangles on the march. For a given chord MN of a circle, consider all triangles ABC such that AB is a diameter of the given circle not intersecting MN and the sides AC and BC pass through the endpoints of

CONTINUED ON PAGE 33

The birth of low-temperature physics

Properties of helium near absolute zero

by A. Buzdin and V. Tugushev

THE PROPERTIES OF SUBstances at low temperatures have attracted scientific attention for a long time. In the 18th century the renowned French scientist Antoine Lavoisier (1743–1794) wrote that if the Earth could be subjected to such severe cold that all the rivers and oceans became mountains of ice and the air turned into a visible liquid, this transformation would open up new possibilities for obtaining unusual new liquids with unheard-of properties. Lavoisier's premonition of a fascinating new low-temperature world has come true in full measure. Low-temperatures research is now a major branch of modern physics.

The behavior of substances near absolute zero has nothing to do with their usual properties at ordinary temperatures. Many beautiful phenomena appear only at extremely low temperatures, while under standard conditions they are masked, as a rule, by the thermal motion of atoms. For example, one needs very low temperatures to see superconductivity—the ability of a substance to carry electric current without even the slightest resistance to its flow. This amazing phenomenon, which has no parallel in classical physics, was discovered by the outstanding Dutch scientist Heike Kamerlingh-Onnes (1853–1926). Kamerlingh-Onnes was the first to obtain liquid helium, and he undertook a comprehensive examination of its properties. It was the possibility of performing routine work at low temperatures that allowed Kamerlingh-Onnes to discover superconductivity. This article tells the story of how liquid helium was produced and how superconductivity was discovered.

Liquefaction of gases

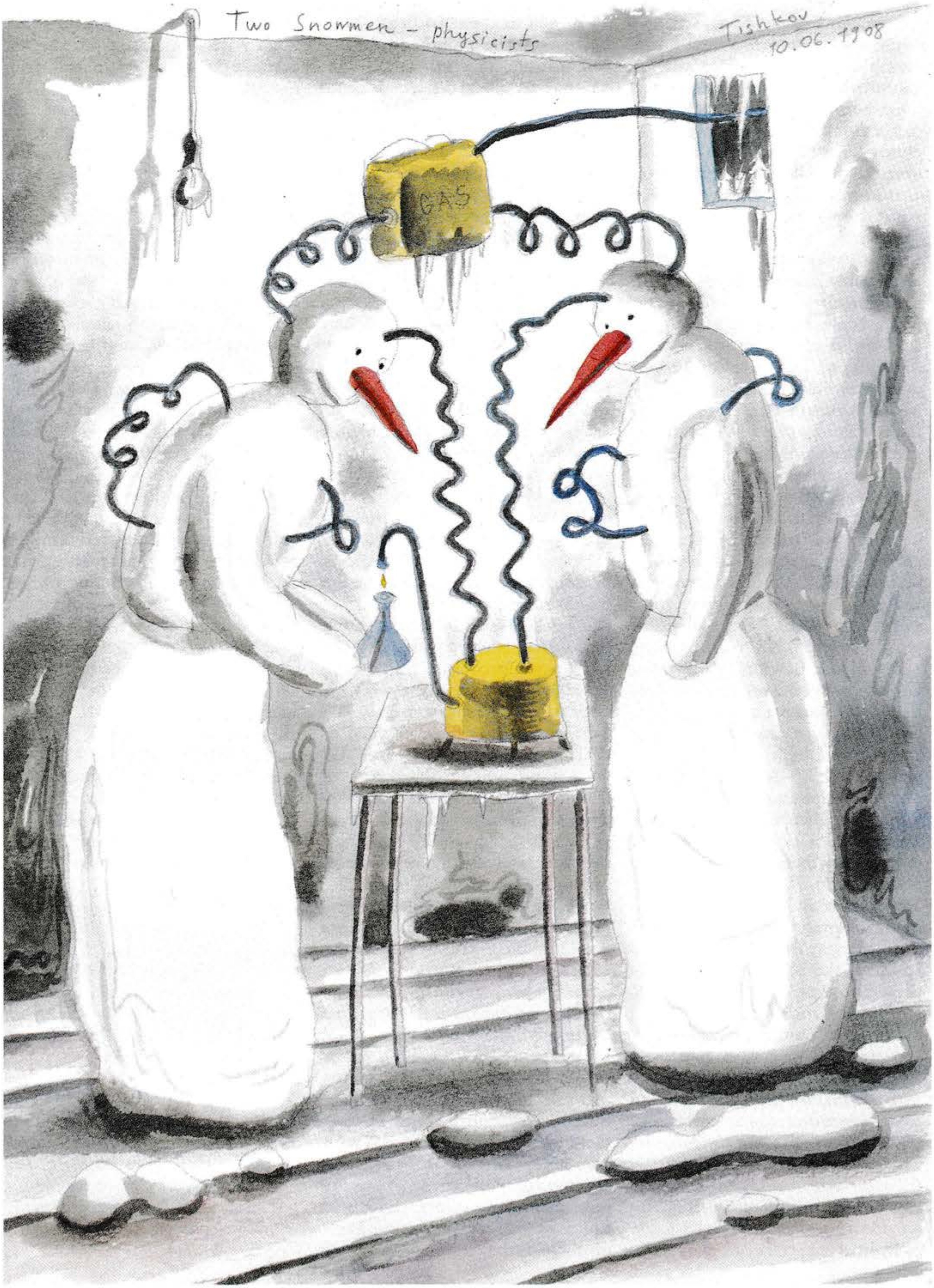
Investigations at low temperatures are closely related to the liquefaction of gases. As you know, many gases become liquids only at very low temperatures. When a liquid "gas" (that is, a substance that is gaseous under standard conditions) is poured into a vessel, it evaporates and, in doing so, absorbs heat from the surroundings. As long as there is even a drop of liquid gas in the vessel, the temperature remains constant (equal to its boiling point).

The history of the liquefaction of atmospheric gases begins in 1877 and is marked by the names of the French scientist Louis Paul Cailletet (1832–1913) and the Swiss physicist Raoul-Pierre Pictet (1842–1929). Cailletet's scientific approach allowed him to achieve liquefaction of the basic atmospheric gases a few weeks before Pictet did, while Pictet's engineering skills led to the creation of devices based on the cascade principle subsequently used in most cryogenic research.

In one experiment Cailletet tried to liquefy acetylene under high pressure. A seal in the apparatus failed and the compressed gas began to leak out. Cailletet, who was observing the experiment carefully, noticed that as soon as the leak began, a thin cloud formed in the vessel and then immediately disappeared. The researcher suggested that the pressure drop caused by the leakage had abruptly cooled the gas and produced the cloud, which was nothing but condensed acetylene. Painstaking experiments confirmed this hypothesis.

Without wasting any time, Cailletet began liquefying atmospheric gases. The first target was oxygen. This gas was compressed to 300 atm in a thick-walled glass vessel, where it was cooled with evaporating sulfur dioxide to –29°C. When the vessel was rapidly depressurized, Cailletet could see a small cloud of

Art by Leonid Tishkov


Two Snowmen - physicists
Tishkov
10.06.1908
GAS

the condensed gas. There was no doubt—this cloud was liquid oxygen. In December 1877, Cailletet submitted a report on the results of his experiments on liquefaction of oxygen to the Paris Academy of Sciences. Soon Cailletet managed to obtain liquid nitrogen.

It isn't difficult to understand why the gases were cooled in Cailletet's experiments. When a gas expands rapidly enough, there is no appreciable heat exchange between the system and its environment. So under such conditions the expansion of the gas is almost adiabatic. According to the first law of thermodynamics, in an adiabatic process $\Delta U = -W$—that is, the change in the internal energy of the gas is equal to the negative of the work performed by the gas on its environment. When the gas expands, it performs positive work on its surroundings ($W' > 0$). Correspondingly, external forces perform negative work during this process. Thus $\Delta U < 0$, which means the gas is cooled, since its internal energy is proportional to its absolute temperature.

In addition to temperature, there is another very important parameter that radically affects the liquefaction of gases. At the end of the 18th century the Dutch physicist Martin von Marum (1750–1837) performed a series of experiments with ammonium to verify the Boyle–Marriotte law. He was a careful observer and couldn't help noticing a paradoxical phenomenon. At first, the pressure increase was inversely proportional to the volume occupied by the compressed gas. When the pressure reached 7 atm, something unexpected occurred: it couldn't be increased any more by further compression. Liquid ammonium appeared in the cylinder, and additional compression only produced more liquid, not higher pressure.

Compression of gases was the main tool used to liquefy gases in experimental thermodynamics in the first half of the 19th century. However, in spite of many attempts, scientists couldn't liquefy such gases as hydrogen, nitrogen, and oxygen. These substances remained gaseous however high the pressure used to compress and liquefy them. No wonder the researchers considered these gases "permanent." In other words, they believed that hydrogen, nitrogen, and oxygen could not exist in the liquid state. However, it was baffling why some gases were permanent while others were not. This nut was cracked not by experimentalists but by the famous Dutch theoretical physicist Johannes Van der Waals (1837–1922).

In the second half of the 19th century Van der Waals deduced a very simple equation to describe the behavior of real gases. The "ideal gas" so often used in physics calculations is a rough model that approximates the properties of real gases only at high temperatures and low densities: the higher the temperature and the lower the density, the better the model functions. In contrast, at low temperatures and high densities a very important role is played by the interactions among the molecules of the gas, as well as the size of the molecules (both neglected in the ideal gas approach). The Van der Waals equation for real gases took both facts into consideration.

Figure 1 shows experimental isotherms for real gases (the solid curves) and some theoretical curves from the Van der Waals equation. Clearly the Van der Waals isotherms describe the behavior of real gases more precisely at higher temperatures. In other words, the Van der Waals equation is also an approximate model of real gases. However, this equation describes the phenomenon of two different states (phases) of matter—liquid and gaseous. The equation predicts that there is some critical temperature above which the difference between these phases disappears.

The main feature of isotherm *I* in figure 1 is the presence of the segment where pressure increases with volume. It's not hard to see that this segment of isotherm *I*, where the pressure changes in the same direction as the volume, means the gas is unstable in that region.

Indeed, let's consider gas in a cylinder with a piston: In the region between points *a* and *b* on isotherm *I*, any random decrease in the volume of the gas would decrease the pressure exerted on the piston, which means a further compression of the gas. Therefore, in this region the Van der Waals theory cannot be applied: The transition of gas to liquid (liquefaction) occurs here. The pressure in the system remains constant during condensation and is equal to the saturated vapor pressure. This constant pressure corresponds to the horizontal part of the real isotherm between valves V_a and V_b. When condensation is complete, we have no more gas in the system. Only experimentation yields the precise value of V at which condensation begins.

When the temperature rises, the maximum and minimum of the unstable segment of the isotherm approach each other and at some temperature they "fuse" (isotherm *II* in figure 1). This critical temperature T_c is a characteristic parameter of each gas. When the temperature is higher than T_c, the gas cannot be liquefied, however great the pressure applied to it. In other words, T_c is the minimum temperature at which a substance can exist in the liquid state. The corresponding value for the critical pressure P_c is the maximum pressure of the saturated gas, while the critical volume V_c determines the maximum density of the

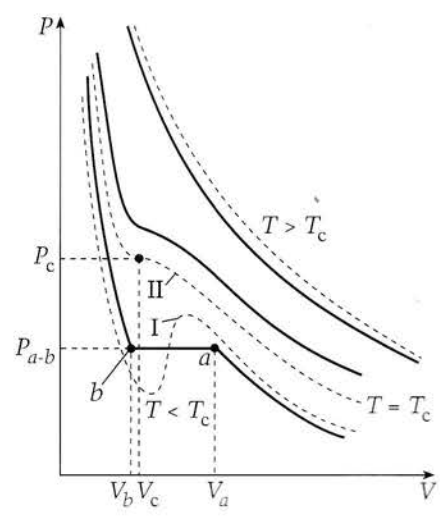


Figure 1

gas. As the gas approaches the critical point, the difference in the densities of the gas and liquid becomes small, and it disappears entirely at the critical conditions. Note that for an ideal gas $T_c = 0$, which means that condensation doesn't occur at any (finite) pressure. This is explained by the fact that the interactions between molecules aren't taken into account in the ideal gas model.

Now it's clear that researchers were trying (and failing) to liquefy some gases at temperatures higher than their critical points. Since this cannot be done at any pressure, these gases were erroneously considered "permanent."

Getting close to liquid helium

Although Cailletet managed to obtain liquid oxygen, he was still confronted with the problem of keeping an appreciable amount of it in the liquid phase, because oxygen evaporates rapidly under normal conditions. He never managed to solve this problem. The next step—from condensation to true liquefaction—was taken by the Polish physicists Karol Stanislav Olszewski (1846–1915) and Zygmunt von Florenty Wroblewski, who replaced the inefficient and low-output apparatus of Cailletet with Pictet's cascade machines and obtained liquid oxygen, "calmly boiling in the test tube." However, despite repeated attempts, Wroblewski and Olszewski could not expand the list of liquefied gases, because hydrogen strongly resisted all attempts at liquefaction. The first person to observe hydrogen in the liquid state was the British physicist Sir James Dewar in 1898.

In Holland the baton in the gas liquefaction race was passed to Heinke Kamerlingh-Onnes, a friend and disciple of Van der Waals. In 1882 Kamerlingh-Onnes headed the Cryogenic Laboratory at the University of Leyden. His first step in this field was to verify experimentally the theoretical predictions of the Van der Waals theory of the critical values of the temperature and pressure at which a gas can be converted to a liquid.

Kamerlingh-Onnes tested many different gases, cooling them to lower and lower temperatures and measuring their isotherms. The scientist was most intrigued by helium, which at that time resisted repeated attempts at liquefaction. The fact that helium remains gaseous down to extremely low temperatures makes it possible to obtain data on the interactions between gas molecules from experimental isotherms (because at low temperatures the deviation of a real gas from the ideal gas model due to molecular interaction manifests itself most clearly).

In 1907 Kamerlingh-Onnes published the results of his measurements of helium isotherms in a wide temperature range from –216° to +100°C. This was followed soon after by measurements at the temperature of liquid hydrogen (–259°C).

Research at such low temperatures required special equipment and led to the development of the technology of gas liquefaction. And this would be a good time to say a few words about the special talents of Kamerlingh-Onnes, who played such a decisive role in the progress made in low-temperature physics at the Leyden laboratory.

Kamerlingh-Onnes was one of those innovators who understood that sophisticated experimental techniques called for skilful and professionally trained specialists. He felt that the time of amateur professors performing experiments with clumsy home-made devices had passed forever. Penetration into the deeper levels of natural mechanisms could be made only by means of a special "industry" devoted to scientific instrumentation. In 1901 Kamerlingh-Onnes organized a glass-blowing workshop at his laboratory. He continually stressed that physical observations must be performed with astronomical accuracy: "Door meten tot weten" ("Through measuring to knowing"). These demands could only be met by radically improving the training and professional skill of the researchers. In essence, the Kamerlingh-Onnes's cryogenic laboratory became the prototype and model for research institutes in the 20th century.

Kamerlingh-Onnes was also well aware of the vital importance of timely information exchange among scientists. To this end he founded the journal *Communications from the Physics Laboratory at the University of Leyden*, which published the results of experiments conducted at his institution. Moreover, the doors of his lab were wide open to anyone who wished to work in cryogenic physics and technology. This was an absolutely new scientific style and a new type of relationship between scientists from different countries and diverse scientific schools.

Liquefaction of helium

Above all, Kamerlingh-Onnes was interested in the critical parameters of helium, mainly at its critical temperature T_c. By measuring helium isotherms at ever lower temperatures, he concluded that T_c was somewhere between 5 K and 6 K. Kamerlingh-Onnes felt that helium could be chilled to subcritical temperatures by rapid expansion of the gas after it had been compressed to 100 atm and cooled by liquid hydrogen. Indeed, during such an experiment Kamerlingh-Onnes observed a dense gray cloud. This cloud would seem to confirm the condensation of helium.

However, further experiments showed that the actual reason for the appearance of this cloud was the presence of a tiny amount of hydrogen, which was not removed from the helium despite careful purification. When the experiment was repeated with helium subjected to additional processing, no cloud appeared in the system. When Kamerlingh-Onnes increased the rate of expansion, he again observed a light hazy cloud. However, this cloud was not dense and disappeared within seconds. Thus the question of the value of helium's critical temperature was still open.

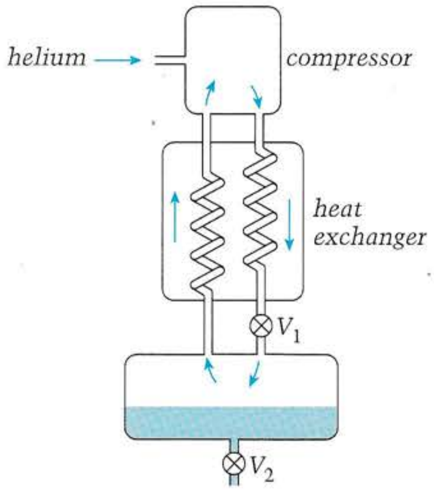


Figure 2

On July 10, 1908, the crucial experiment was carried out. The idea behind this experiment is shown in figure 2. Strongly compressed helium passes from the compressor to a heat exchanger, where it is chilled by liquid hydrogen. The compressed and cooled helium goes to a receiving tank through a special valve V_1, where it expands drastically and cools. While the temperature of the helium in the receiving tank is higher than T_c, no condensation occurs there and the chilled helium returns to the compressor through the heat exchanger. In the heat exchanger this portion of the helium provides additional cooling to the compressed helium going in the opposite direction, so the new portion of gas entering the receiving tank has a lower temperature than the previous one. This step-by-step cooling decreases the temperature of the helium expanding in the receiving tank down to the critical value, where condensation begins. The helium produced is collected into a Dewar flask through a special valve V_2.

The experiment started at 5:45 A.M. with the liquefaction of 20 liters of hydrogen needed to cool the helium. This work was finished at 1:30 P.M. Extreme caution was required during the preliminary cooling of the system with liquid hydrogen. The smallest amount of air leaking into the system would ruin the experiment. The air would condense due to contact with the liquid hydrogen and freeze onto the glass vessel containing helium, making all further observation impossible. Circulation of helium began at 4:20 P.M., and at this moment the liquifier's internal cryostat entered the still unknown region of extremely low temperatures.

For a long time the temperature indicator hardly moved. After additional adjustment of the apparatus, the researchers observed a gradual decrease in temperature, but soon it stopped. Almost the entire supply of liquid hydrogen had been used up, but there was still no sign of helium liquefaction. At 7:30 P.M. it seemed as if their attempt to liquefy helium had failed, but at this critical juncture one of Kamerlingh-Onnes's colleagues suggested that the absence of further progress of chilling helium could simply be caused by the fact that the thermometer was immersed in an invisible boiling liquid.

The researchers illuminated the collecting tank from below, and suddenly they clearly saw that the vessel was filled almost entirely with a liquid, which became visible due to the reflection of the light. They had produced almost 60 cm^3 of liquid helium. Here is how Kamerlingh-Onnes described the historic moment: "It was a wonderful sight—the first view ever of a liquid with an almost immaterial appearance. Its entry into the vessel went unnoticed. It was not noted until it had filled the vessel, and then its surface was outlined as sharply as the edge of a knife. I was happy that I could show liquefied helium to my respected friend Van der Waals, whose theory was the thread that guided our experiments to the very end."

As that famous experiment drew to a close, Kamerlingh-Onnes tried to obtain solid helium by decreasing the pressure in the vessel where the liquid was boiling. To obtain the lowest possible temperature, he let the liquid boil until only 10 cm^3 was left in the receiving tank and then connected this helium cryostat to a vacuum pump that decreased the pressure over the boiling helium to 0.01 atm. However, there were no signs of solid helium, so Kamerlingh-Onnes supposed that the freezing point of helium could not be reached with his apparatus.

Later, he made two other attempts to solidify helium, but again he failed. In 1922 Kamerlingh-Onnes made his final attempt to solidify helium. Using a dozen pumps to remove the vapor above the liquid helium, he decreased the pressure down to 0.013 mmHg and attained a temperature of 0.83 K! However, the helium remained liquid even at this record low temperature.

In his first experiment with liquid helium, Kamerlingh-Onnes was astonished by its very low density. It turned out to be less dense than water by a factor of eight. This means that the helium atoms are located at relatively large distances from one another. Therefore, it's much more difficult to solidify such a substance in comparison with ordinary liquids.

This is why helium remains a liquid down to extremely low temperatures. However, it would be liquid even near absolute zero, because in this temperature region its solidification is prohibited by the laws of quantum mechanics. Indeed, in classical physics the thermal oscillations of particles become weaker and weaker as the temperature drops. The intermolecular attractive forces must eventually lead to solidification of the cooled substance. However, according to quantum physics, the notion that atoms stop moving completely at absolute zero is erroneous. In reality, the atoms perform so-called "zero-point" vibrations even at zero temperature, and these oscillations have nothing to do with thermal motion.

Thus the unique behavior of liquid helium at low temperatures is at odds with classical physics. At normal pressure this substance doesn't freeze at temperatures near absolute zero. A huge pressure of p = 30,000 atm is needed to solidify helium at T = 1.78 K.

Kamerlingh-Onnes noticed another strange property of liquid helium in his very first experiment with helium liquefaction. When the temperature was decreased below the boiling point, the density of helium increased as expected, which meant that helium can be compressed by cooling (or expanded by heating). However, at 2.2 K there was a sharp peak in the density–temperature plot, and further cooling decreased the density of the liquid helium. Research conducted by Kamerlingh-Onnes in the last years of his life uncovered other anomalous properties of helium chilled below 2.2 K.

Later, after the death of Kamerlingh-Onnes in 1926, his students published the results of their studies of liquid helium. In 1932, Keesom and Clausius found a drastic increase in the specific heat for helium near $T = 2.2$ K. A few years later it became clear that this temperature is the landmark of a phase transition, which changes not only the specific heat but many other properties of liquid helium. In the temperature range from $T = 2.2$ K to its boiling point, helium is just a liquid with more or less ordinary properties. This liquid is referred to as He I.

In contrast, at $T < 2.2$ K helium exhibits unexpected and paradoxical properties. In 1938, the outstanding Russian physicist Pyotr Kapitsa found that at $T < 2.2$ K liquid helium (referred to as He II) has no viscosity and can flow in a narrow capillary without resistance. In other words, friction is unknown to He II. This property of He II was called " superfluidity." The theory of superfluidity was created by Kapitsa's colleague, the great Soviet theoretician Lev Landau. The creative unity of their experimental and theoretical work was highly appreciated by the scientific world: Both researchers became Nobel laureates.

In his experiments with liquid helium, Kamerlingh-Onnes actually observed the transition to superfluidity, but he had no conceptual framework to appreciate and describe the significance of this event. It should not be forgotten that quantum physics was at that time in its infancy. To explain the behavior of liquid helium at low temperatures, the old classical views had to be thrown out and a new theory created. This work could be performed only by physicists of the next generation. In Kamerlingh-Onnes's time, the very process of helium liquefaction was a great scientific victory—one that seemed to be the last stage on the road to absolute zero.

Discovery of superconductivity

Having created a special technique for attaining temperatures around 1 K, Kamerlingh-Onnes and his collaborators performed numerous experiments to test the properties of various substances at "helium temperatures." It's relatively easy to measure the electrical resistivity at any temperature. Quite understandably, Kamerlingh-Onnes began to study the behavior of electrical resistivity at low temperatures. However, it wasn't just experimental simplicity that led him to do it.

At the beginning of the 20th century there were two electrical theories that provided opposite predictions of the behavior of electrical resistivity at very low temperatures. One of them proclaimed that resistivity should disappear at absolute zero temperature. Indeed, resistivity in metals is inversly proportional to the time between two successive collisions of an electron with oscillating atoms. When the temperature drops, the amplitude of the atomic oscillations drops (it looks as if the atoms "shrink"), so the chance of encountering an atom becomes negligibly small. This means the resistivity should tend to zero as $T \to 0$.

The other theory said that low temperatures freeze anything, so electrons stop "wandering" in the metal and condense onto the host atoms. In this case, no carriers of electricity would remain in the metal, which translates to an infinitely high resistivity as $T \to 0$.

Kamerlingh-Onnes chose platinum as his first test material when he began measuring resistivities at low temperatures. Surprisingly, his results didn't fit either theory: The resistivity of extremely cold platinum did not depend on temperature at all! However, one might come to expect anything in this crazy frozen world. Kamerlingh-Onnes noticed that the resistivity of various platinum samples was lower as their purity increased. He concluded that existence of a finite resistivity as $T \to 0$ (so-called *residual resistivity*) results from impurities, while a clean metal should have a vanishingly small resistivity at zero temperature. However, Kamerlingh-Onnes erroneously supposed that the resistivity of pure metals should gradually decrease to zero even at "helium temperatures."

Therefore, the problem was to examine the resistivity of the cleanest possible specimens. For example, gold can be purified much better than other metals, and better than platinum in particular. As expected, the residual resistivity of gold was much lower than that of platinum and it decreased more as the specimens became purer.

Exploring this problem further, Kamerlingh-Onnes began to investigate the resistivity of mercury at low temperatures. This substance is a liquid at room temperature, so it's a rather simple task to obtain highly purified mercury specimens by successive distillation. The results of the seemingly routine experiments that followed were utterly unexpected. The resistivity of mercury did not gradually decrease as it was cooled—when the temperature was slightly lower than the boiling point of helium, the resistivity dropped drastically and became immeasurably small.

On April 28, 1911, Kamerlingh-Onnes reported the results of his research to the Royal Academy of Sciences in Amsterdam. He called the newl phenomenon "superconductivity."

CONTINUED ON PAGE 27

Where do problems come from?

The art of problem composition

by I. Sharygin

A WIFE PUTS DOWN A BOOK indignantly and turns to her husband: "Do you remember the poem you wrote for me on our wedding day? They're saying in this book that it was written by some guy named Petrarch!"

I'd like to begin this article with two contradictory statements. On the one hand, I don't recommend that every student who reads this article immediately start composing problems. On the other hand, I invite you to take a crack at composing problems (starting with geometry problems) and send them to me. They might appear in our magazine someday.

In this article I'd like to give you the benefit of my long experience in composing geometry problems, to tell you some of my secrets, and to formulate some aesthetic and even ethical principles.

For convenience, we can divide all problems into three categories: textbook exercises, competition problems, and Olympiad-style problems. We might also add another category: "creative" problems, but this is more of a "subtext" than a formal feature, since the "creativity" has more to do with how the problem is solved than with the problem itself.

Now let's peek into the problem maker's "bag of tricks."

Reformulation

Here's an example of a problem that makes use of the technique of reformulation.

Problem 1. A circle is circumscribed about a triangle, and a diameter is drawn that is perpendicular to one side. This diameter is then projected onto a second side. Prove that the length of the projection is equal to that of the third side of the triangle.

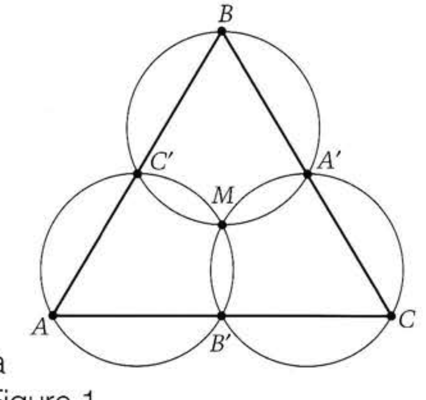
The solution is left to the reader (who may want to recall that a segment of length s projected onto a line with which it forms an angle θ, is equal to $s \cos \theta$).

The following theorem is well known.

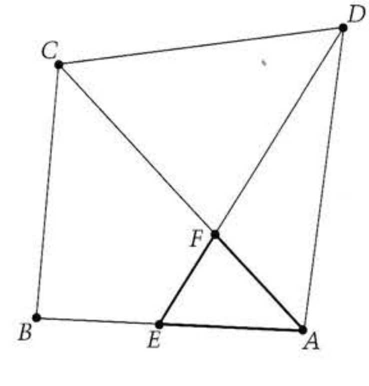
Theorem. *If arbitrary points A', B', and C' are taken on lines AB, BC, and CA, respectively (distinct from the vertices of triangle ABC), then the circles passing through points A, B', and C', A', B, and C', and A', B', and C have a common point.*

This theorem is sometimes called Miquel's theorem, and the common point of the circles is called the Miquel point and is denoted by M.

The proof of this theorem is not very complicated. The only difficulty, when no oriented angles are used, is in examining all the different cases of various mutual positions of the points C', A', and B'. In the situation shown in figure 1a, calling M the point of intersection of the circles passing through points A, B', C' and A', B, C', it is not difficult to prove that the points A', B', C, and



a



b

Figure 1

Art by Vadim Ivanyuk

M lie on a circle. (For example, we can show that $\angle C$ and $\angle B'MA'$ are supplementary.)

Here is a problem suggested at an All-Union Olympiad.

Problem 2. A point E (different from A and B) is given on side AB of a convex quadrilateral $ABCD$. Segments AC and DE intersect at point F. Prove that the circles circumscribed about triangles ABC, CDF, and BDE have a common point.

Solution. A close examination of figure 1b and the statement of the problem shows that problem 2 coincides with the theorem above. Indeed, it's sufficient to reformulate the condition of the problem in terms of triangle AEF, renaming the points as follows: $E \to B$, $B \to C'$, $F \to C$, $C \to B'$, and $D \to A'$. Certainly the statement of the problem is less natural and thus is less attractive than that of the theorem. Maybe I'm wrong about the origin of problem 2—so much the worse for the organizers of the Olympiad!

Elegant and effective problems can arise when a geometrical problem is translated into the language of algebra. For example, consider the well-known problem of constructing a triangle given its three altitudes (are you able to solve it?). The idea is that the triangle with sides a, b, and c is similar to the triangle with sides $1/h_a$, $1/h_b$, and $1/h_c$.

Let the lengths of the triangle's altitudes be a, b, and c, and let its sides be x, y, and z. If this triangle is acute, we immediately obtain a system of equations for x, y, and z. Thus we have the following problem (figure 2).

Problem 3. Solve the system of equations

$$\begin{cases} \sqrt{x^2-c^2}+\sqrt{y^2-c^2}=z, \\ \sqrt{y^2-a^2}+\sqrt{z^2-a^2}=x, \\ \sqrt{z^2-b^2}+\sqrt{x^2-b^2}=y. \end{cases}$$

Knowing the origin of this system, we easily find the condition for its consistency (this is the condition that the triangle with sides $1/a$, $1/b$, and $1/c$ be acute) and then solve the system. Prove that both

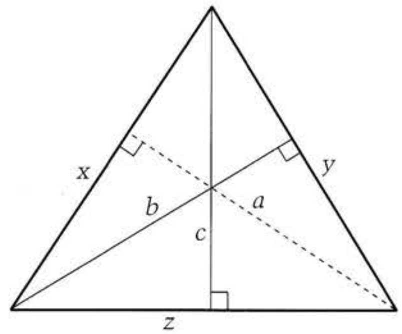


Figure 2

the system and the construction problem cannot have more than one solution.

The change in the problem statement obtained by passing from the direct proposition to the inverse one can also be classified with this kind of problem, although it's a bit of a stretch. (We should stress that the boundary between different kinds of problems is rather fuzzy. One and the same problem may illustrate different methods, especially since the final problem is often obtained by a combination of several methods.) This method has many variations, but here I'll give only one example that demonstrates how a trivial direct proposition can give rise to a problem with rich geometric content. By looking at various right triangles, we can see that the orthocenter H of acute triangle ABC (the point of intersection of the altitudes) possesses the property

$$\angle HAB = \angle HCB,$$
$$\angle HBA = \angle HCA,$$
$$\angle HAC = \angle HBC.$$

So the following problem arises quite naturally.

Problem 4. Find the locus of points M for which $\angle MAB = \angle MCB$ and $\angle MBA = \angle MCA$, where ABC is the given acute triangle.

Solution. It's clear that the point of intersection of the triangle's altitudes belongs to the locus in question. The nontrivial problem is to find if there exists only one such point inside the triangle. We prove that this is indeed the case. Let's extend AM, BM, and CM until they intersect the triangle's sides at points A_1, B_1, and C_1 (figure 3). Points A, C,

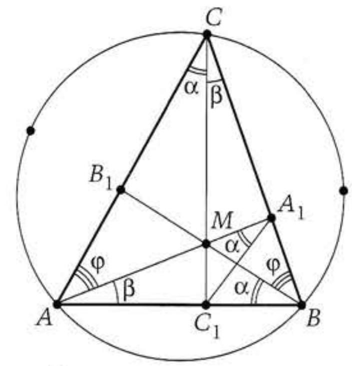


Figure 3

A_1, and C_1 lie on a circle (since $\angle A_1AC_1 = \angle A_1CC_1$). Therefore, $\angle MA_1C_1 = \angle MCA = \angle MBC_1$ and $\angle MAC = \angle MC_1A_1$. Thus points M, B, A_1, and C_1 also lie on a circle and $\angle MBA_1 = \angle MC_1A_1 = \angle MAC$. Denoting the angles by α, β, and ϕ, as in figure 3, we find that $\alpha + \beta + \phi = \pi/2$, which implies that AA_1, BB_1, and CC_1 are the altitudes of the given triangle. However, the locus of points we seek is not exhausted by the single point of the altitudes' intersection. The arc AB of the circle circumscribed about triangle ABC and the midpoints of arcs BC and CA also belong to this locus (prove this fact).

Problems built on other problems

Often in geometry, the structure of a problem involves working other geometric figures or results into a given diagram. Solid geometry problems of this kind are often used in Russian university entrance exams. When the structure of such problems is complicated, the solution usually comes step by step, stacked like books on a bookshelf or a set of nested wooden dolls.

However, such problems don't necessarily have a complex underlying structure. Here is a simple example made up of two (or three) subproblems.

Problem 5. The diagonals of a convex quadrilateral divide it into four triangles. Prove that the product of the areas of two opposite triangles is equal to the product of the areas of the two other triangles.

Problem 6. Prove that the square has the maximum area among all quadrilaterals inscribed in a given circle.

We invite the reader to solve these two problems, while we pose a new one based on them.

Problem 7. Quadrilateral $ABCD$ is inscribed in a unit circle. The diagonals of this quadrilateral intersect at point M. Find its area if it is given that the product of the areas of triangles ABM and CDM is 1/4.

Solution. To solve this problem, it's sufficient to note (figure 4) that $S_1S_3 = S_2S_4 = 1/4$ (problem 5) and that

$$\begin{aligned} S &= S_1 + S_2 + S_3 + S_4 \\ &\geq 2\sqrt{S_1S_3} + 2\sqrt{S_2S_4} = 2 \end{aligned}$$

(the arithmetic mean is greater than or equal to the geometric mean).

In addition, $S_{ABCD} \leq 2$ (problem 6), since the area of the square inscribed in a unit circle is 2. All this implies that $ABCD$ is a square and its area is 2.

Many problems are constructed so that they can be solved using a particular idea. And wouldn't you know, it often turns out that these problems have a different (and sometimes simpler) solution.

Here's one of my problems of this sort. I wanted to compose a problem in which the reasoning that three bisectors of a triangle meet at a point (more exactly, that the third bisector passes through the point of intersection of two other bisectors of angles of a triangle, not necessarily internal ones) was repeated twice. In addition, the second stage had to be essentially dependent on the first stage of the solution. I can't say the resulting problem was particularly successful, since it's based on a well-known construction; nevertheless, here it is.

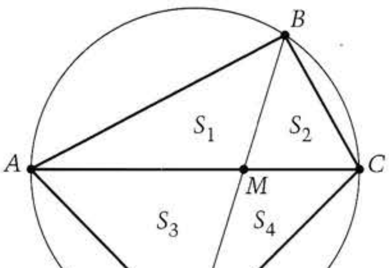


Figure 4

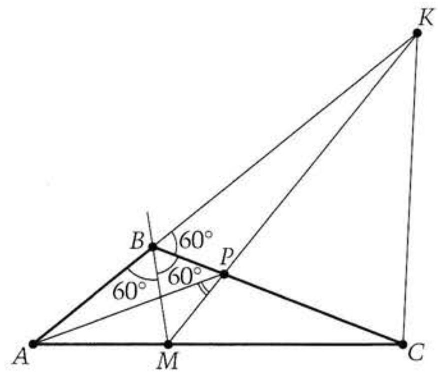


Figure 5

Problem 8. The measure of angle B of triangle ABC is 120°. A point M is taken on side AC and point K is on line AB such that BM is the bisector of angle ABC and CK is the bisector of the angle adjacent to angle ACB. Segment MK intersects side BC at point P. Prove that $\angle APM = 30°$.

Solution. Our reasoning will involve two steps. For triangle BMC, lines BK and CK are the bisectors of the external angles B and C, respectively (figure 5). Therefore, MP is the bisector of angle BMC, and P is the point of intersection of the bisectors of external angles that are adjacent to angles B and M of triangle ABM. Thus AP is the bisector of angle BAC. Finally, we obtain

$$\begin{aligned} \angle APM &= \angle PMC - \angle PAM \\ &= {}^1/_2(\angle BMC - \angle BAM) \\ &= {}^1/_2(\angle ABM) = 30°. \end{aligned}$$

We used the fact that the measure of an exterior angle of a triangle is equal to the sum of the two remote interior angles.

Finally, a problem may be constructed so as to obtain a required result (this trick is often used in exercises). By way of example, here is a "trap" (such problems are encountered very rarely). The numerical data are selected so as to create an unusual geometrical structure. In spite of some drawbacks (I had to resort to a trick in the statement), I like this problem. However, it can be classified neither as an Olympiad-style problem (it's unusual to suggest problems with numerical data at Olympiads) nor as a competitive-type problem (even the best students can get caught in a trap, which contradicts the purpose of a competition). Rather, it's an instructional problem.

Problem 9. A convex quadrilateral with two sides 6 units long and two other sides 10 units long lies at the base of a pyramid. The height of the pyramid is 7 units, and the lateral faces are inclined at the angle of 60° to the basal plane. Find the volume of the pyramid.

Solution. From the statement of the problem, the measure of the dihedral angles at the base is 60° or 120° (but not necessarily 60°—this is the invisible "fine print" in the problem statement!). The projection of pyramid's vertex onto the plane of the base is equidistant from the sides of the quadrilateral (more exactly, it is equidistant from the lines to which the sides of the quadrilateral belong). Therefore, this quadrilateral cannot be a parallelogram. Its two adjacent sides are 6 units long, and the two other sides, also adjacent, are 10 units long. Now, if $AB = BC = 10$ and $AD = DC = 6$ (see figure 6), there exist two points (O_1 and O_2) that are equidistant from its sides (figure 6). The conditions of the problem imply that the distance

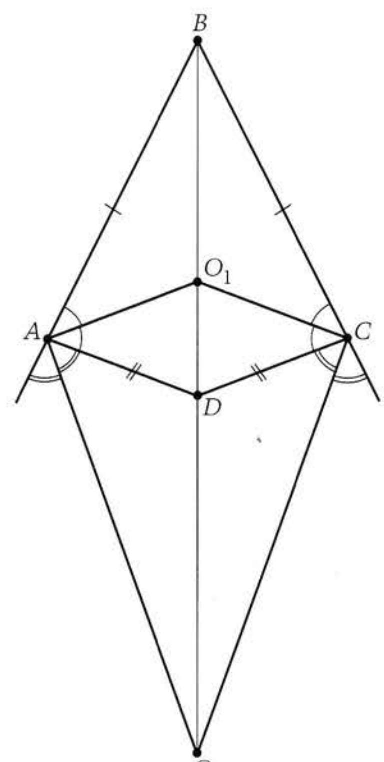


Figure 6

from the projection of the pyramid's vertex to its sides is $7/\sqrt{3}$. If the vertex is projected into point O_1, which is the center of the circle inscribed in $ABCD$, the area of $ABCD$ must be

$$16\cdot\frac{7}{\sqrt{3}}.$$

However, this area cannot exceed 60 (it equals 60 when angles A and C are 90°), and

$$16\cdot\frac{7}{\sqrt{3}}>60.$$

Thus the pyramid's vertex is projected into point O_2, and the distance from this point to the sides of quadrilateral $ABCD$ is $7/\sqrt{3}$.
Now we easily calculate the area of $ABCD$, which is

$$(10-6)\frac{7}{\sqrt{3}}=\frac{28}{\sqrt{3}},$$

and the volume of the pyramid, which is

$$\frac{64}{\sqrt{3}}.$$

A specific case

Many general theorems that give us a powerful instrument for solving problems, such as Ceva's theorem in geometry or inequalities about the various means in algebra, may be also used to compose problems. Consider, for example, Pascal's theorem: *If six points A, B, C, D, E, and F lie on a circle, then three points at which the pairs of lines AB and DE, BC and EF, and FA and CD intersect lie on a line.* Here is a problem linked to Pascal's theorem.

Problem 10. Let sides AB and DE of an inscribed quadrilateral $ABDE$ intersect at point M and let sides BD and AE intersect at point K. Prove that the tangent to the circle drawn through points B and E meet on line KM.

It's easy to see that this theorem is a particular case (more exactly, a limiting case) of Pascal's theorem, where points B and C coincide, as do points E and F. Two sides of the hexagon referred to in Pascal's theorem, in the limiting case, have become tangents.

Professional mathematicians who help organize mathematical olympiads often derive elegant and interesting problems from their scientific work. Particular cases of fundamental theorems and numerous lemmas that arise in proving almost any theorem can often be reformulated as problems for high school students. Such examples rarely occur in geometry, since the modern view of this subject differs considerably from the version taught in school (in saying this, I mean no slight to geometry). In this article I'll give only one—perhaps not the most characteristic or shining—example.

There exists a "zigzag" theorem. Two circles (possibly in space) are given. It's known that there exists a set of $2n$ points $A_1A_2\,..A_{2n}$ such that the points with odd numbers lie on one circle, the points with even numbers lie on the other, and $A_1A_2 = A_2A_3 = \ldots = A_{2n}A_1$. Given all this, there exist infinitely many such sets of $2n$ points, any point of the first circle may be taken for point A_1, and the distance between the consecutive points in these sets is the same as for the basic set.

I don't know of any elementary proof of this theorem. However, its particular cases can be used as elementary problems. Here is one of them.

Problem 11. Two circles with radii R and r are given in the plane; the distance between their centers is a. Find the side of a rhombus such that its two opposite vertices lie on one circle and two other vertices lie on the other.

I leave this problem as a challenge to the reader, giving only the answer here:

$$\sqrt{R^2+r^2-a^2}.$$

Varying the problem statement

The following set of problems illustrates what can happen when you vary the statement of a problem just a bit: Construct a triangle (a) given its three sides; (b) given its three medians; (c) given its three altitudes; (d) given its three angle bisectors. This examples show how a small change in the statement of a problem can result in an enormous change in its level of difficulty. Problem (a) is a standard textbook exercise. Problem (b) is only a little more difficult, although quite interesting. Problem (c) is much more difficult, and problem (d) cannot be solved with straightedge and compass.

An interesting method for writing series of problems was suggested by the high school teacher V. Kutsenok. Consider a geometric relationship—for example, the equality

$$ah_a = bh_b \qquad (1)$$

—and ask the question: What are the properties of a triangle for which the relation obtained from (1) by replacing the altitudes by medians or bisectors holds? As a result, we obtain the following problem.

Problem 12. Two points, A and B, are given in the plane. Find the locus of points C in the plane such that the following equality holds for triangle ABC:

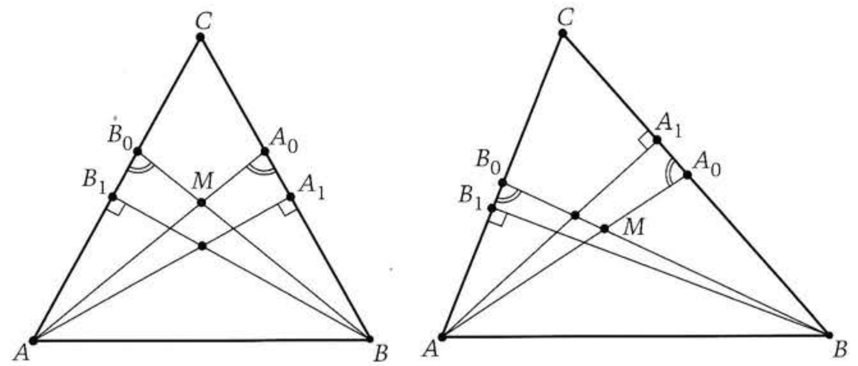
(a) $am_a = bm_b$ (m_a and m_b are medians of triangle ABC);

(b) $a\beta_a = b\beta_b$ (β_a and β_b are bisectors of triangle ABC).

The solutions for both cases are similar. Consider case (a).

Let AA_1 and BB_1 be altitudes of the triangle, and let AA_0 and BB_0 be its medians. It follows from the statement of the problem that the right triangles AA_0A_1 and BB_0B_1 are similar. The two possible arrangements of points A_1, A_0, B_1, and B_0 on the sides of triangle ABC are shown in figures 7a and 7b. In the first case, points A, B, A_0, and B_0 lie on a circle. Since A_0B_0 is parallel to AB, this fact implies that the trapezoid AB_0A_0B is isosceles, so $AC = BC$. In this case C is on the perpendicular bisector of AB.

In the second case, points C, M, A_0, and B_0 lie on a circle. In the inscribed quadrilateral CA_0MB_0, diagonal A_0B_0 is half of AB and is bisected by diagonal CM. Diagonal CM equals $^2/_3 m_c$, so CM divided by diagonal A_0B_0 in the ratio 3:1. Now we use the fact that if a quadrilateral is inscribed in a circle, then the products of the segments of its two di-



a b

Figure 7

agonals are equal. We obtain $m_c^2 = 3AB^2$, which is a constant. Thus the desired locus consists of the perpendicular bisector of segment AB and the circle centered at the midpoint of AB with the radius $AB\sqrt{3}$.

In case (b), point C must be located on the perpendicular bisector of AB or on an arc of the circle from which AB subtends an angle of 60°.

Sometimes problems can be obtained by varying the statement of a well-known result. Here I'll give just two examples. The well-known Steiner–Lehmus theorem states that a triangle is isosceles if its two bisectors are congruent. This theorem is quite natural, though the proof is rather complicated (in contrast to similar theorems about congruent medians and bisectors in a triangle, which are easily proved). Here we see the "bad behavior" of bisectors, which is displayed in full measure in the following problem.

Problem 13. In triangle ABC,the angle bisectors are AA_1, BB_1 and CC_1.Is the triangle isosceles if it has

(a) congruent bisectors of the external angles A and B?

(b) congruent segments KA_1 and KB_1, where AA_1 and BB_1 are the bisectors of the interior angles of the triangle and K is their intersection point?

(c) equal distances from point C_1 to the midpoints of sides CA and CB?

(d) $C_1A_1 = C_1B_1$?

(e) the circle passing through points A_1, B_1, and C_1 is tangent to side AB of the given triangle?

The answer is no in all five cases. Cases (d) and (e) are the most difficult. In this article I'll consider only case (e), which I got to know through some gossip in mathematical circles. I'll demonstrate how a nonisosceles triangle possessing the desired property can be constructed, but I won't explain how this solution was cooked up.

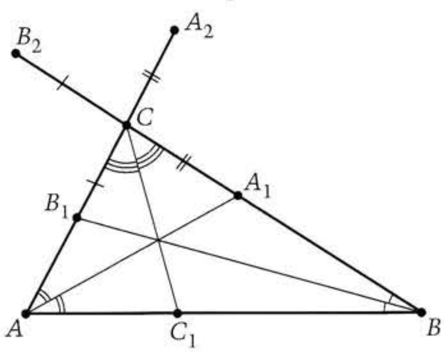


Figure 8

Let AA_1, BB_1, and CC_1 be the bisectors of triangle ABC (figure 8). Consider points A_2 and B_2 on the extensions of sides AB such that $CA_2 = CA_1$ and $CB_2 = CB_1$. It's clear that points A_1, A_2, B_1, and B_2 lie on a circle (check that the angles $\angle B_2A_1A_2$ and $\angle A_2B_1B_2$ are equal). If it turns out that AC_1 and BC_1 are equal to the tangents drawn from points A and B to this circle, then this circle touches AB at point C_1. If the length of a segment equals the sum of the tangents drawn from its endpoints to a circle, then this segment is tangent to the circle. The proof of this fact is left to the reader. Thus the following equalities must hold:

$$AC_1^2 = AB_1 \cdot AA_2,$$
$$BC_1^2 = BA_1 \cdot BB_2.$$

It is well known that an angle bisector of a triangle divides the side to which it is drawn into two segments which are proportional to the other two sides. This fact allows us to express each of the segments in terms of the lengths of the three sides. We get

$$AB_1 = \frac{bc}{c+a}$$

and so on. Then we find that each of these two equations is equivalent to

$$(a + b + c)(a + b)^2 = c(c + a)(c + b).$$

It remains to prove that a nonisosceles triangle exists such that its sides satisfy this relation. A numerical example is easy to find. Let $c = 1$, $a + b = 1 + \lambda$. Then, rewriting the previous equation in terms of λ, we have $ab = \lambda(2 + \lambda)^2$. If λ is sufficiently small, we can find numbers a and b sutisfying these equations, with $a \neq b$.

Another interesting series of problems is related to the situation in which a tetrahedron has congruent faces. Many necessary and sufficient conditions are known for the tetrahedron to have congruent faces. Not trying to be complete (or even to set a record), I'll formulate the following problem.

Problem 14. Which of the following conditions are necessary and sufficient for a tetrahedron $ABCD$ to have congruent faces?

(a) For each pair of opposite edges, the edges are equal.

(b) The perimeters of all the faces are equal to each other.

(c) The sums of the plane angles adjacent to all the vertices are equal to 180°.

(d) The following equality holds: $\angle BAD = \angle BCD = \angle ABC = \angle ADC$.

(e) The following equalities hold: $\angle BAC = \angle BDC$, $\angle ABD = \angle ACD$, $\angle BAD = \angle BCD$.

(f) The radii of the circles circumscribed about all the faces are equal.

(g) The radii of the circles inscribed in all the faces are equal.

(h) The areas of all the faces are equal to each other.

(i) The segments connecting the midpoints of the opposite edges are perpendicular pair by pair.

(j) The center of the circumscribed sphere coincides with the center of mass of the tetrahedron.

(k) The center of the inscribed sphere coincides with the center of mass of the tetrahedron.

(l) The centers of the circumscribed and inscribed spheres coincide.

(m) The sum of the cosines of the dihedral angles is –2.

(n) There exist four spheres with centers located at the sphere circumscribed about the tetrahedron such that each of these four spheres touches a face in an interior point and the planes of three other faces.

That's enough. Generally speaking, it's possible to formulate several dozen such conditions, especially if we take into account that some of the conditions may be mixed. For example, the first eight conditions are written as a combination of three equalities. We can combine them in a different way, taking one equality from, say, item (a) (two opposite edges are equal) and two equalities from item (c) (the sums of the plane angles adjacent to two vertices are equal to 180°).

In this problem, almost all the conditions are necessary and sufficient for the tetrahedron to have congruent faces. You may have guessed that item (g) is an exception (again, bisectors mess things up). And this is indeed the case. Try to construct a counterexample. (You might want to start by taking two pairs of noncongruent isosceles triangles which possess the desired property.)

Item (n) is very difficult to prove—at least, I don't know of any elementary proof.

By the way, item (e) occurred to me as I was writing this article. It's worth noting that here the fact that the tetrahedron is a spatial figure is essential: if A, B, C, and D lie in a plane, the property indicated in item (e) is not sufficient for the triangles to be congruent.

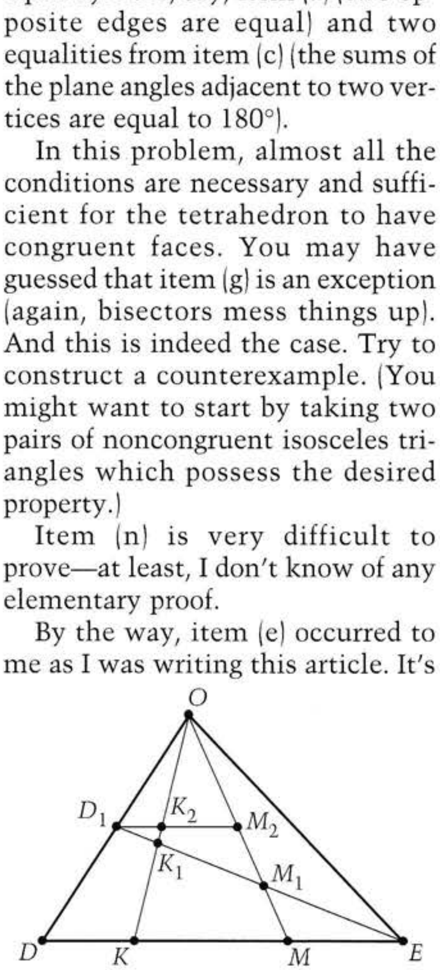


Figure 9

Generalizations

The development of mathematics is associated with a continuous chain of generalizations. When using generalization in composing elementary problems, we don't expect to obtain fundamental results. Nevertheless, it's important to understand the significance of this method.

Generalizations can follow different directions. Sometimes it's possible to remove some constraints in a problem and extend the proposition to a wider set of objects. For example, I once encountered the following problem in an old mathematics journal.

Problem 15. Side AD of an inscribed quadrilateral $ABCD$ is a diameter of the circle, and the bisectors of angles B and C meet on side AD. Prove that $AB + CD = AD$.

I didn't like the solution to this problem presented in the journal. After thinking about it, I found another solution that didn't use the condition that AD was a diameter of the circle; it turned out to be redundant. As a result, a new problem was born.

Problem 15′. The bisectors of angles B and C of an inscribed quadrilateral $ABCD$ meet on side AD. Prove that $AB + CD = AD$.

My solution (there are many others) used the following reasoning. Consider point P on AD at which the bisectors of angles B and C meet. Draw the circle circumscribed about triangle BCP and denote by M the second point of intersection of this circle with AD. Taking into account the congruence of the various angles in inscribed quadrilaterals $ABCD$ and $BCPM$, we can prove that triangles ABM and CDM are isosceles ($AB = AM$ and $CD = CM$, respectively). I invite the reader to finish the proof.

This example also shows that it's very useful to try different methods when solving a problem, paying special attention to more geometrical methods, since they allow for deeper understanding of the intrinsic properties of figures and make it easier to distinguish between the essential and the secondary. Let's look at another example. In 1990, at the All-Union Mathematical Olympiad, the following problem was proposed.

Problem 16. Points D and E lie on sides AB and BC, respectively, of triangle ABC. Points K and M divide segment DE into three equal parts. Lines BK and BM intersect side AC at points T and P, respectively. Prove that

$$TP \le \frac{1}{3} AC.$$

The very statement of the problem left me cold—it was too tedious and contained too many different letters. I didn't like the solution either—it didn't agree with the principles I had already worked out. Having found a better solution (from my point of view), I was able to reformulate the problem and make it more general.

Problem 16′. Two rays emanate from the vertex of an angle and lie inside this angle. A line intersects the sides of the angle at points D and E and it intersects the rays at points K and M. Prove that the ratio KM/DE attains its maximum if $DK = ME$.

To prove this fact, we consider another line that intersects the sides of the angle at points E and D_1 and intersects the given rays at points K_1 and M_1 (figure 9). Set $DK = ME = a$, $KM = b$, and $OD_1 = \lambda OD$. Draw line DM_2 parallel to DE through D_1. We have

$$D_1K_2 = \lambda a,\ K_2M_2 = \lambda b,$$

$$\frac{D_1M_1}{M_1E} = \frac{\lambda(a+b)}{a},$$

(from similar triangles $D_1M_2M_1$, EMM_1). This is algebraically equivalent to

$$D_1M_1 = \frac{\lambda(a+b)}{\lambda(a+b)+a} \cdot D_1E,$$

$$D_1K_1 = \frac{\lambda a}{\lambda a + b + a} \cdot D_1E,$$

$$K_1M_1 = \left(\frac{\lambda(a+b)}{\lambda(a+b)+a} - \frac{\lambda a}{\lambda a + b + a} \right) D_1E.$$

Here we used several pairs of similar triangles, and well-known properties of proportions.

Ultimately, the problem is reduced to proving a simple algebraic inequality:

$$\frac{\lambda(b^2 + 2ab)}{(\lambda(a+b)+a)(\lambda a + a + b)} \le \frac{b}{2a+b},$$

which can be written as

$$(\lambda - 1)^2 a(a+b) \ge 0.$$

Another possible direction of generalization is carrying over a geometric fact from one object to another—in particular, extending geometric properties from plane to three-dimensional figures. That's what gave rise to the following problem.

Problem 17. Two tangents AB and CD are drawn to two spheres such that points A and C lie on one sphere and B and D on the other. Prove that the projections of AC and BD onto the line connecting the centers of the spheres are equal.

For the case of the plane, this problem is simple (here AB and CD are the common external and internal tangents, respectively, to two circles). For the spatial case, it isn't very difficult either. The solution relies on the fact that the midpoints of the tangents common to two spheres lie in a plane that is perpendicular to the line connecting the sphere's centers (the proof of this fact is left to the reader). In my view, this problem is remarkable, because here the three-dimensional analogue of a proposition formulated for plane figures remains valid, which is a rare occurrence. Often it's necessary to construct a counterexample that disproves such a generalization. For example, the simple proposition of plane geometry stating that the foot of at least one altitude in any triangle lies on the corresponding side (rather than on its extension) gives the following problem for the three-dimensional case.

Problem 18. Is it true that the foot of at least one altitude of any tetrahedron belongs to the corresponding face?

The answer is no. A counter-example is provided by a tetrahedron in which two dihedral angles corresponding to two skew edges are obtuse.

Rather often a problem can be generalized in different directions, generating a series of problems. Consider this well-known theorem: *The sum of distances from an arbitrary point inside an equilateral triangle to its sides is constant.* (For those of you who are unfamiliar with this statement, I'll sketch out the proof here. The area of the equilateral triangle under consideration equals the sum of the areas of the three triangles that have the sides of the given triangle as their bases, and whose common vertex is an arbitrary point inside the equilateral triangle.)

The assertion of this theorem can easily be extended to the case of any convex equilateral polygon. It's less obvious that it also can be extended to the case of any equiangular polygon. Indeed, let $A_1A_2 \ldots A_n$ be an equiangular polygon (see figure 10, where n is taken to be 5). Consider the regular n-gon $A_1'A_2' \ldots A_n'$ whose sides are parallel to the sides of the initial n-gon containing it. For any point M inside $A_1A_2 \ldots A_n$, the sum of distances to the sides of $A_1'A' \ldots A_n'$ is constant. The distance from M to any side of the initial n-gon is less than the distance from M to the parallel side of the enveloping regular n-gon by a constant amount. Thus the sum of the distances from M to the sides of $A_1A_2 \ldots A_n$ differs from the sum of distances from M to the sides of $A_1'A_2' \ldots A_n'$ by a constant. Therefore, it is itself a constant.

We can make a further generalization that combines the preceding ones.

Problem 19. Given n different unit vectors in the plane that add up to zero, consider the convex n-gon with sides perpendicular to those vectors. For any point inside this n-gon, the sum of distances to its sides is the same.

Other ways of generalizing the initial theorem about the equilateral triangle are possible—for example, we may consider the three-dimensional case. It would be interesting to learn if the three-dimensional analogue of problem 19 is valid.

Discoveries and problems

The examples in the preceding sections illustrated certain techniques. However, the main source of new problems is inquisitiveness, the desire to reveal the essence of a problem, the ability to look at a well-known fact from an unusual point of view. This is when the most interesting geometric problems appear, ones that can be called discoveries. Here is one of the most elegant Olympiad-style problems that has appeared in recent years.

Problem 20. Is it possible to saw three regular tetrahedrons with a unit edge from a wooden unit cube?

This problem was suggested at the All-Russia Olympiad in 1989. It's interesting that the problem of cutting two unit tetrahedrons from a unit cube had been discussed in many collections of olympiad problems, and it turned out that as many as three tetrahedrons can be cut! Indeed, consider three edges of the cube that cross in pairs. Each of them will be the edge of a tetrahedron. The midpoints of the opposite edges of every tetrahedron coincide with the center of the cube. Now it remains to prove that these tetrahedrons have no other common points.

Whether or not an elegant fact that you've discovered was known earlier is not really that important. It sometimes happens that an old geometric theorem comes as a surprise to experts in geometry. Unfortunately, much has been lost in geometry over the past thousands of years.

I consider the following problem one of my best geometric discoveries.

Problem 21. What is the maximum number of lines that can be drawn through a point in three-di-

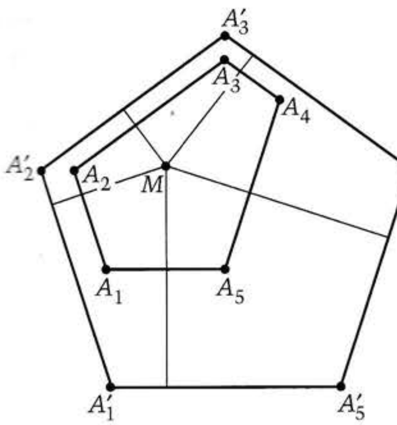


Figure 10

mensional space such that all the angles between them, taken in pairs, are equal?

The answer is six. I'm certainly aware that this fact had probably been discovered in ancient times—for instance, Archimedes may have known it.

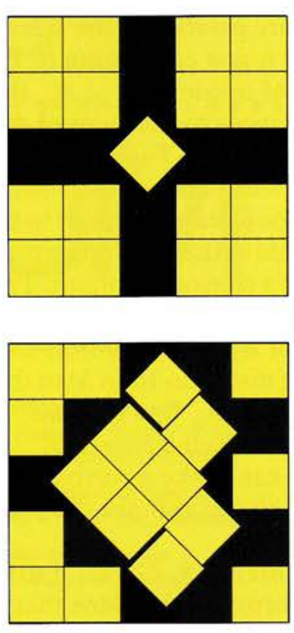
It's not difficult to prove that the number of such lines cannot exceed six. Indeed, let l_1 and l_2 be two lines from the set of lines under consideration passing through point O. Then all other lines of this set must belong to the intersection of two conical surfaces: for the first surface, l_1 is the axis and l_2 is its generating line; for the other surface, l_2 is the axis and l_1 is the generating line. Such conical surfaces cannot intersect on more than four lines.

An example of the six lines in question is provided by the diagonals of an icosahedron (the regular 20-hedron that has 12 vertices). If it's difficult for you to imagine an icosahedron, you can construct the example as follows. Take six vectors $(a, \pm b, 0)$, $(\pm b, 0, a)$, and $(0, a, \pm b)$ and use them as the directing vectors of the lines to be constructed. Since all these vectors have the same length, the absolute values of their paired scalar products must be equal as well. Assuming that $a \geq b > 0$, we obtain the equation $a^2 - ab - b^2 = 0$. For example, we can set $b = 1$ and

$$a = \frac{1+\sqrt{5}}{2}.$$

In geometry, as in no other field of mathematics, the distance between a textbook problem and an open problem can be short. For example, the statement of the following problems doesn't seem to differ much: *Find the minimum value of the area of a triangle containing a unit circle* and *find the figure of minimum area that can be used to cover any plane figure of unit diameter.* Nevertheless, the first problem is a simple exercise, while the second is the open Lebesgue problem on minimum covering. To state such an interesting and meaningful problem, one must possess a deep understanding of mathematics. You should keep in mind the proverb that one fool can ask a question that a hundred clever men can't answer.

Also in contrast to many other fields of mathematics, geometry allows experimentation in the direct, physical sense of this word. Many geometric discoveries in ancient times were the result of observations and experiments. It's possible that the remarkable contemporary geometer Connelli, who constructed a deformable n-hedron (an n-hedron that can change its shape so that every face remains unchanged), experimented in the course of his work—that is, created physical models. The Connelli n-hedron solved one of the oldest mathematical problems. The fact that this solution turned out to be quite elementary seems fantastic, taking into account the level of development of modern mathematics. This is possible very rarely, and perhaps only in geometry.



a

b

Figure 11

I'd like to give an example of a small discovery made experimentally. Here is a problem in elementary geometry: *What is the maximum number of unit squares that can be cut from a square with side length* $4 + \alpha$, *where* $0 < \alpha < 1$? This problem attracted the attention of many remarkable mathematicians—for example, the Hungarian Paul Erdös. It was the source of another problem that was suggested in the journal *Mathematics in School.*

It's known that 17 unit squares can be cut from a square with side length $4 + \alpha$. Find the minimum possible α for which this can be done.

Maybe this problem isn't very elegant, but nevertheless it makes sense. We certainly didn't expect our readers to give an exact answer. The majority of them gave the answer illustrated in figure 11a, for which

$$\alpha = \frac{\sqrt{2}}{2}.$$

An unexpected solution was sent in by members of the mathematics club of school 51 in Kiev (headed by V. N. Shkolnik). The accompanying letter indicated that the members experimentally found that the minimum α is obtained from the arrangement shown in figure 11b. You can verify that, for this arrangement, α is several hundredths smaller than that shown in figure 11a. Since I don't have a detailed description of the experiment, it's difficult to judge the validity of this statement. Besides, my mathematical training prevents me from accepting this "proof." ("And why is that?" you may ask. It's worth discussing.) It's quite possible that the arrangement shown is indeed the optimal one for the case under consideration. The point is, you don't have to be a budding mathematical genius to make geometric discoveries—this problem shows that any student can do it. And that includes you!

CONTINUED FROM PAGE 17

In 1913 Kamerlingh-Onnes was awarded the Nobel Prize. In his ceremonial lecture he suggested that the phenomenon of superconductivity might be related to the quantization of energy discovered by Max Planck (1858–1947) at the end of the 20th century. But nearly a half-century would pass before the complete theory of superconductivity was created on the basis of quantum mechanics, as the great scientist predicted.

Having detected the transition of matter to a previously unknown state, Kamerlingh-Onnes tried to thoroughly investigate its properties. The first question was: How small does the resistivity of a superconducting material become? To answer it, he needed to invent a method of measuring infinitely small resistivity. As usual, he succeeded brilliantly, producing a clever new method of measurement.

The apparatus consisted of a coil made of lead wire (figure 3) that could be connected to a battery via a switch (1), while another switch (2) could short-circuit the coil. At the start of the experiment switch 1 was closed and switch 2 was open. The coil was cooled by liquid helium in a Dewar flask and maintained in the superconducting state. The current from the battery passing through the coil created a magnetic field around it, which was easily detected by a compass needle placed outside the Dewar flask. Switch 2 was then closed (guess how the electric current is distributed between the branches of the circuit) and after a while switch 1 was opened. Now the superconducting coil was short-circuited by switch 2. In an ordinary coil this would lead to damping of the current in the coil and fading of the generated magnetic field around it. However, the compass needle stayed deflected, which confirmed the existence of an electric current in the coil, even though it was disconnected from the battery.

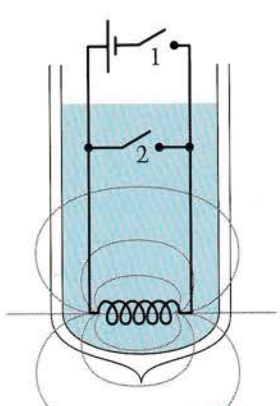


Figure 3

By observing the needle over the course of many hours (until all the helium evaporated from the bath), Kamerlingh-Onnes could not detect any change in the position of the needle. This meant there was no electrical resistance in the coil—otherwise the energy would be dissipated, current would decrease, and the magnetic field would disappear, causing the needle to return to its initial position.

Kamerlingh-Onnes was able to estimate the upper limit of the resistivity of the superconducting lead coil, which turned out to be lower than its resistivity in the normal (nonsuperconducting) state by at least a factor of 10.

To date, the longest-lived undamped current is about two years. Perhaps this current would have flowed into the 21st century if transportation workers had not gone on strike and interrupted the supply of liquid helium. Even after two years of current circulation, there were no indications that it was being damped, from which we might infer that the resistivity of a superconductor is practically zero.

The practical application of superconductivity shows great promise. A superconducting electromagnet consumes no energy, so it could provide an easy way of producing extremely strong magnetic fields. Such fields are usually generated by huge currents flowing in electromagnets, which dissipate a tremendous amount of heat in the coils. This heat impedes any further increase of current necessary for stronger fields. The use of superconductors in transformers, electric motors, and generators also promised great advantages that would more than compensate for all the expenses involved in working at helium temperatures.

Kamerlingh-Onnes was the first to construct a superconducting electromagnet. However, he was disappointed in his endeavor. The experiments carried out at the Leyden Cryogenic Laboratory definitely showed that the phenomenon of superconductivity disappears for magnetic fields greater than some critical value. These critical (threshold) values were rather small—a few hundred gauss, which is significantly lower than the intensity of the fields generated even in small electrical machines. Large currents also destroyed superconductivity, because the related magnetic field surpassed the critical value even at rather moderate values of electric current.

Many tanks of liquid helium later, superconducting materials were invented, which were capable of resisting strong magnetic fields and carrying huge currents without destroying the superconductivity. More than forty years of tenacious work was necessary to produce the first superconducting magnets that were of practical importance. At present, the mainstream of superconductivity studies is aimed at creating materials that become superconducting at high temperatures—perhaps, even at room temperatures.

Research into the physical properties of substances at low temperatures plays a prominent role in modern science. The prediction of Kamerlingh-Onnes has indeed come true: "From every field of physics come problems whose solution is to be found at helium temperatures. Looking into the future, I see measurements being made everywhere in cryostats filled with liquid helium, which will flow as freely as water. This work will certainly pull off the veil used by thermal motion to hide the internal world of atoms and electrons at ordinary temperatures." ◘

Dielectrical m

AS OUR EPIGRAPH HINTS, we'll be looking at an interaction between two types of matter that involves electricity. From this broad area of physics we'll start with questions of the behavior of nonconducting objects—insulators (or dielectrics)—in electric fields. In fact, the first electrical phenomenon that people noticed was the movement of such objects in an external, non-uniform electric field.

Thousands of years of observation and experimentation have given us a good grasp of the properties of many dielectrics. Some unusual phenomena were also discovered along the way—for instance, pyroelectricity (electrification of crystals during heating) and segnetoelectricity (spontaneous electric polarization due to mechanical stress). Eventually each of these phenomena led to applications in science, technology, and everyday life—in extremely sensitive devices for finding defects in machine parts or for listening to the heart. They are exploited in the manufacture of miniature capacitors and sensing transducers, microphones, and telephones. They help us visualize heat radiation and obtain "thermal photographs." Piezoelectrics are used in cigarette lighters, and dielectrics are the foundation of acoustical and optical electronics, which may provide future computers with new components to replace outdated electronic chips. A new technical term has even been coined for them: "smart dielectrics."

However, it should be noted that successful research into the physical properties of various dielectrics would be impossible without a deep understanding of their molecular structure. In turn, data on the specific features of electrical processes occurring at the microscopic level provide the key to obtaining new information about the structure of atoms and molecules.

Let these questions and problems serve as your admission ticket to this fascinating field.

"...ordinary matter is a kind of sponge with respect to the electrical fluid ..."

—Benjamin Franklin

Questions and problems

1. Why are small pieces of paper attracted to a charged plastic comb, but not to either plate of a parallel-plate capacitor?

2. Why do all attempts to remove the charge of a dielectric to ground end unsuccessfully?

3. The plates of a vertical parallel-plate capacitor, charged and disconnected from a battery, are partially immersed in a liquid dielectric. Where is the electric field more intense—in the air or in the dielectric?

4. Why do electrolytic capacitors have such a large capacity?

5. A metal plate is inserted in the space between the plates of a charged capacitor. Would the charges induced in the plate change if the space between the capacitor plates is filled with kerosene?

6. A dielectric bar is situated between the plates of a charged parallel-plate capacitor (figure 1). Draw the electric field lines, neglecting the distortion at the edges of the plates.

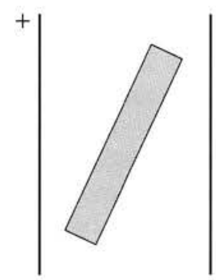


Figure 1

7. A charged metal ball is surrounded by a thick spherical dielectric layer. Plot the pattern of the lines of force inside and outside the dielectric. Why does the electric field change at the metal-dielectric boundary?

8. Two small metal balls are connected to a distant voltage source. How will the attractive force between the balls change if you im-

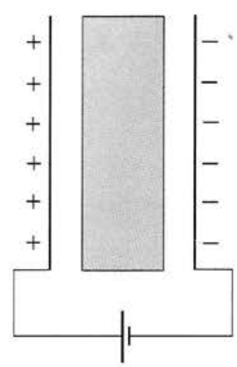


Figure 2

OSCOPE

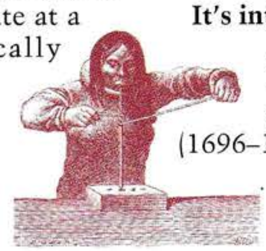
materialism

merse them in a liquid dielectric without changing the distance between them?

9. When the plates of a parallel-plate capacitor are connected to a battery, they start to attract each other. How will the attractive force change if a dielectric plate is inserted into the capacitor (figure 2)?

10. Given two capacitors of the same capacity, in which the same dielectric material is used, why is the one designed to operate at a higher voltage physically larger?

11. The voltage drop across a charged capacitor disconnected from a battery doubled when its dielectric leaked out. What is the relative dielectric constant of this liquid dielectric?



12. A parallel-plate capacitor is charged, disconnected from the voltage source, and lowered into kerosene. How does the energy stored in the capacitor change?

13. The plates of a dielectric-filled and charged capacitor are short-circuited for the briefest moment. When the potential difference across the capacitor decreases by a factor of three, the plates are disconnected. Then the potential difference slowly increases to 2/3 of its initial value. Why?

14. Which substances reflect electromagnetic waves better—metals or dielectrics?

15. Does an electron interact with a neutral atom?

16. What can you say about the structures of the following molecules: (a) carbon dioxide, which has no dipole moment (that is, it's a nonpolar molecule); (b) water, which is a strongly polarized molecule with a pronounced dipole moment?

Microexperiment

Electrify a plastic comb by friction and watch how it attracts small pieces of paper. But when you place the same pieces near the terminal of a charged battery, nothing happens. Why?

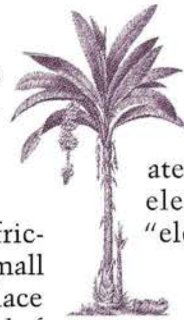
It's interesting that ...

...at the beginning of the 18th century the English physicist Stephan Grey (1696–1736) found that friction-induced electrification of an object proceeds more efficiently if it is first heated. This is because the heat causes moisture to evaporate, which reduces conductivity and improves the object's dielectric properties.

...the first description of pyroelectricity appeared as early as the fourth century B.C. The details of this phenomenon were investigated in the middle of the 18th century by the German physicist Franz Maria Aepinus (1724–1802), who worked in St. Petersburg. He showed that electrification of tourmaline crystals induced by heating is fundamentally different from electrification by friction, which was well known at that time.

...the concept of dielectric permeability, which characterizes the attenuation of an electric field in dielectric bodies, was introduced by Michael Faraday (1791–1867) in 1837 under the name of "specific inductive capacity."

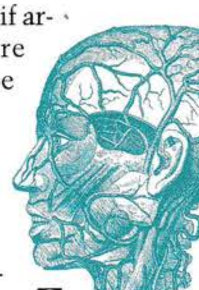
...it may have been Faraday who predicted the existence of electrostatic analogues of permanent magnets. These dielectrics, which generate a permanent and constant electric field, were termed "electrets" at the end of the 19th century by the renowned English physicist Oliver Heaviside (1850–1925). The first artificial electret was obtained about 80 years ago from a mixture of palm resin and rosin.



...at the beginning of the 19th century, the French mineralogist René Juste Hauey (1743–1822) found that pyroelectric crystals can be electrified by pressure. He used this property to construct a sensitive electroscope. Later this effect, which is an intrinsic property of many crystals, was called "piezoelectric."

...the maximum possible polarization of dielectrics was achieved in 1918 in crystals of Seignette (Rochelle) salt. Thus a new term, *seignette-electric*, was coined. Nowadays the dielectric permeability of certain ceramic materials can attain huge values—as high as 20,000.

...many tissues in living organisms (for example, blood vessels) are electrets. This fact must be taken into consideration if artificial vessels are used—they must be pretreated in an electric field to prevent an increase in coagulation, which could lead to the formation of a blood clot.



ANSWERS, HINTS & SOLUTIONS ON PAGE 53

PHYSICS CONTEST

A good theory

by Arthur Eisenkraft and Larry D. Kirkpatrick

> *In my work, I have always tried to unite the true with the beautiful; but when I had to choose one or the other, I usually chose the beautiful.*
> *—Hermann Weyl*

CREATION CAPTIVATES THE mind. We view creation as a series of miracles. The birth of a human being is an act of creation. The first simple thought of a young child is another act of creation. We create and we admire those individuals whose creations extend our own perceptions. As a culture, we struggle with the creation of the world by inventing stories and theories as to how everything we know came into being. "In the beginning, God created the Heavens and the Earth" is one such story. The Big Bang is another such story. "Non-being existed not, nor being" from the Upanishads is still another story.

The painter, the writer, and the composer share their personal conceptions of the world and allow us to peer into their minds. The scientist must also create a personal conception. The scientist, however, bears the pressure of a large constraint. As fertile as the imagination may be, the scientist's creation must be consistent with measurements of the physical world.

The greatest scientists create their world views and help us to see our world through their lenses. As noted in the quote above, Weyl describes how the inner vision of the scientists guides their work. The great scientists have such a superb intuition about the world that when they conceive of a world different from the one we know, we often discover that the "real" world embodies manifestations of their vision and that society's earlier view was myopic. Of course, that vision is further corrected centuries later by a new set of corrective lenses, which further clarifies the blur.

How does a good theory get judged? It must first be able to explain what the prevailing theory has successfully explained. It must also be able to explain some known phenomenon that the prevailing theory is unable to explain. When that theory is able to predict something that nobody has perceived, and that something is then discovered, we realize that we have a very good theory.

One such good theory is Newton's theory of gravity. When Newton proclaimed that every mass attracts every other mass with a force that is inversely proportional to the distance between them, he was able to explain the motion of the planets about the Sun. Kepler had described this quite well, but Newton's synthesis did more. It was able to explain why the ratio of the square of the period to the cube of the distance was a constant for all planets orbiting the Sun. Newton provided a means to weigh the Sun. He was also able to describe the motion of the tides. What was Newton's surprise prediction? Long after Newton's death, slight perturbations were measured in Neptune's orbit. If Newton was correct, these perturbations were signaling the existence of a planet that nobody knew about. Uranus' discovery was an enormous support for Newton's theory of gravitational attraction.

Einstein's theory of gravity supplanted Newton's 200-year success story. Einstein's warping of a space-time continuum was able to explain the motions of the planets and was as successful as Newton's explanation. Einstein's creation was also able to explain the precession of Mercury's orbit about the Sun. The precession was well known but it could not be explained using Newton's theory. Einstein succeeded in shedding light on this puzzle. What did Einstein predict that nobody knew about? He predicted the bending of light as the light passed by a large mass—a bending that was larger than one might expect from Newton's gravity and Einstein's earlier $E = mc^2$. When Arthur Eddington viewed the eclipse of 1919, his experimental team

Art by Tomas Bunk


BANG!
EMERGENCY EXIT
NON BEING
OUT TO LUNCH
GENESIS
TAO
CREATE YOUR OWN WORLD IN 7 DAYS!
BATTERIES NOT INCLUDED
INTUITION
N. BOHR
ACTIVE PHYSICS
ARTHUR
TOMAS BUNK

found that Einstein had indeed predicted a phenomenon that nobody had envisioned.

Niels Bohr and all of his colleagues were aware that hydrogen had a distinct spectrum. Rydberg found a mathematical means of calculating the wavelengths of the emitted light. Bohr's theory of electron orbits about the nucleus was able to do a better job than Rydberg. Bohr postulated that the electrons were restricted to orbits whose angular momentums were whole number multiples of Planck's constant divided by 2π:

$$mvr = \frac{nh}{2\pi}.$$

Combining this with the notion that the electrons move in circular orbits due to a Coulomb force

$$\frac{mv^2}{r} = \frac{kq_1q_2}{r^2},$$

Bohr was able to find both the radii and the energy levels of the different orbits. The energy levels could be succinctly described in terms of the lowest energy assigned to the tightest orbit:

$$E_n = \frac{E_1}{n^2},$$

where E_1 is the ground state energy of -13.6 eV and n is the number of the orbit.

Bohr was able to explain the spectrum of hydrogen by noting that electrons jumping from one energy level to a lower energy level would emit photons with an energy corresponding to the difference in energy of the two levels. If the electron jumps from the $n = 3$ orbit to the ground state $n = 2$,

$$E_3 - E_2 = hf,$$

where $h = 6.63 \times 10^{-34}$ J·s is Planck's constant and f is the frequency of the light.

The visible spectrum of hydrogen emerged from jumps of the electron from the $n = 3$ level to the $n = 2$ level, the $n = 4$ level to the $n = 2$ level, and the $n = 5$ level to the $n = 2$ level. What if the electrons could also jump from higher levels to the $n = 1$ level? Bohr could calculate those frequencies of light—frequencies in the ultraviolet part of the spectrum. When explorations of the ultraviolet took place, Bohr's predictions of the frequencies were right on.

Bohr also predicted frequencies of light for the hydrogen spectrum, which would be emitted if an electron jumped to the $n = 3$ level. These infrared frequencies were also right on. Some of these frequencies had been known. Some had not. The frequencies corresponding to jumps to the $n = 4$ and $n = 5$ levels had never been observed. As humans, our eyes limit our world view to the visible spectra. Bohr, with the assistance of his very good theory, helped us peer into the invisible world.

What else did Bohr's theory tell us that we never expected? Bohr was able to create an understanding of the periodic table and predict the chemical properties of element 72, which had previously been a blank in the table. The discovery of element 72 (named hafnium after Bohr's home town) came just hours before Bohr was to accept his Nobel Prize.

Our first contest problem for this month draws from Newton's and Bohr's theories.

1. This problem comes from a wonderful, slim book *Thinking Like a Physicist*, edited by N. Thompson. A small moon of mass m and radius a orbits a planet of mass M while keeping the same face towards the planet. Show that if the moon approaches the planet closer than

$$r_c = a\sqrt[3]{\frac{3M}{m}},$$

loose rocks lying on the surface of the moon will be lifted off.

2. This problem was first given at the 7th International Physics Olympiad in Warsaw, Poland, in 1974. A hydrogen atom in its ground state collides with another hydrogen atom in its ground state at rest. What is the least possible speed for which the collision is inelastic? If the speed is greater, a photon is emitted that can be observed in the direction of the initial velocity or in the opposite direction. How much do the frequencies of these photons differ from the frequency they would have if emitted from an atom at rest? The mass of the hydrogen atom is 1.67×10^{-27} kg and its ionization energy $E = 13.6$ eV $= 2.18 \times 10^{-18}$ J.

Batteries and bulbs

In the July/August 2000 issue of *Quantum* we asked a series of questions about identical light bulbs wired to an ideal battery. Our faithful contributor Art Hovey from Amity Regional High School in Connecticut, and Güney Gönenç from Ankara, Turkey, submitted correct solutions for these circuits.

In figure 1, bulb A is the brightest; it has the standard brightness as it has its own path to the battery. Bulbs B and C are equally bright because they are wired in parallel and have the same current. They are dimmer than bulb A because their path has more resistance. For real bulbs that do not obey Ohm's law, the current through these bulbs will not be one-half of that through bulb A.

We now look at the questions about removing or short-circuiting various bulbs. (1) If bulb A is removed from its socket, the other two bulbs do not change brightness as they have an independent path to the battery. Note that we are assuming that the battery is ideal—that is, it can provide any amount of current demanded by the circuit. (2) If bulb C is removed from its socket, bulb B will go out as its path to the battery is broken. Bulb A is not affected. (3) If a wire is connected across the terminals of bulb A to short-circuit it, all of the bulbs will go out as the short-circuit across bulb A also shorts out the path through bulbs B and C. (4) If bulb C is short-circuited, it goes out. Bulb A is not affected. Bulb B gets brighter because there is less resistance on its path. In fact, it becomes as bright as bulb A as the two are now wired in parallel.

In figure 2, bulb A is the brightest as all of the current passes through it. Bulbs C and D are the dimmest and bulb B has an intermediate brightness because its path has less

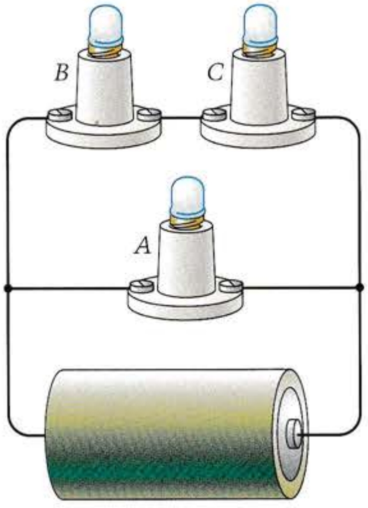


Figure 1

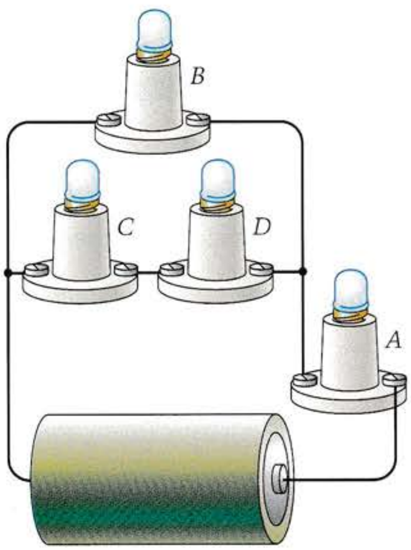


Figure 2

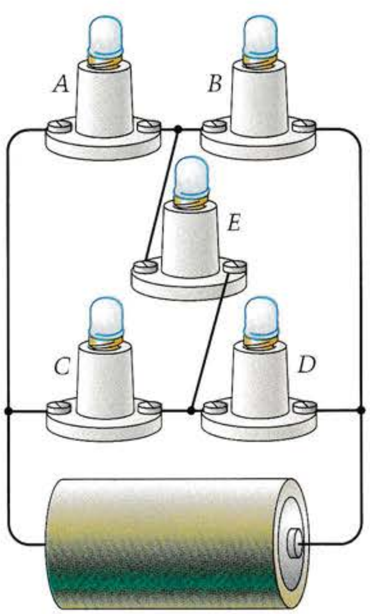


Figure 3

resistance. (1) Removing bulb A breaks the circuit to the battery and all of the bulbs go out. (2) Removing bulb C causes bulb D to go out as its path is broken. Bulb A will get dimmer because removing one of the parallel paths increases the resistance in A's path. Because bulb A and bulb B are now in parallel, they will be equally bright. (3) Shorting bulb A causes bulb A to go out and converts this circuit into the circuit of figure 1. Removing the resistance of bulb A causes all of the other bulbs to increase in brightness. (4) Shorting bulb C removes some of the resistance in the path of bulb A, so it will brighten. Bulbs B and D are now in parallel and have the same brightness. Two competing effects determine the change in brightness of bulb B. There is more current from the battery but only half of it flows through bulb B. If we introduce the concept of voltage, we can argue that bulb A getting brighter requires bulb B to get a bit dimmer.

If we wire up the bulbs as shown in figure 3, we immediately notice that all of the bulbs have the same brightness except for bulb E, which does not glow. We can use symmetry to argue that there should be no current through bulb E. (1) If bulb A is removed, the circuit becomes the same as that in figure 2. The symmetry for bulb E is removed and it will now glow. The resistance of the circuit increases so there is less current from the battery. But all of the current now must pass through bulb C. The latter effect wins and bulb C brightens. Bulbs B and D dim. (2) Removing bulb E from its socket has no affect, as there was no current through bulb E in the original circuit. (3) Removing bulbs A and E causes bulb B to go out. Bulbs C and D will not be affected. (4) Removing bulbs A and D leaves us with three bulbs wired in series. Bulb E will brighten and bulbs B and C will dim. (5) Shorting-circuiting bulb A provides a direct path for bulb B to the battery. Therefore, it will brighten to the standard brightness. Bulbs C and E are now in parallel so bulb E will brighten. Bulb C was in series with one other bulb, but is now in parallel with bulb E and the combination is in series with bulb D. This means that bulb C will dim. The resistance in the path of bulb D decreases—the resistance of one bulb to that of two bulbs in parallel—so bulb D will brighten. (6) Short-circuiting bulb E has no affect on the brightness of the other three bulbs, as there was no current through this branch. (7) Short-circuiting bulbs C and E at the same time also short-circuits bulb A. Therefore, these three bulbs will go out. The remaining bulbs each have a direct path to the battery and will brighten to the standard brightness. (8) Shorting bulbs A and D leaves us with three bulbs in parallel, so each will brighten to the standard brightness. ◘

CONTINUED FROM PAGE 11

MN. Prove that the altitudes of all such triangles drawn from vertex C to side AB meet in a point.

(E. Kulanin)

M313

Underlying n. Given n real numbers $x_1, x_2, \ldots, x_n$ satisfying the conditions $x_1 + x_2 + \ldots + x_n = 0$ and $x_1^2 + x_2^2 + \ldots + x_n^2 = 1$, prove that there exist i and j such that $x_i x_j \le -1/n$.

(N. Vasilyev and E. Stolov)

M314

More than four. Prove that if the product of two positive numbers is greater than their sum, then the sum is greater than 4. (N. Vasilyev)

M315

Exactly one-fourth. A point D is taken on the base AC of isosceles triangle ABC such that the circle inscribed in triangle ABD has the same radius as the circle that touches the extensions of segments BC and BD and segment CD (the escribed circle of triangle BCD). Prove that this radius equals 1/4 of the triangle's altitude drawn to a leg.

(I. Sharygin and N. Vasilyev)

ANSWERS, HINTS & SOLUTIONS ON PAGE 49

AT THE
BLACKBOARD I

The sum of minima and the minima of sums

by R. Alekseyev and L. Kurlyandchik

IN THIS ARTICLE WE PRESENT a general method that makes it possible to prove many well-known inequalities. The underlying idea is quite simple. Consider two functions, $f(x)$ and $g(x)$, defined on an interval $[a, b]$ that attain their respective minima at two (in the general case, different) points (figure 1).

The following inequality holds:

$$\min f(x) + \min g(x) \le \min (f(x) + g(x)).$$

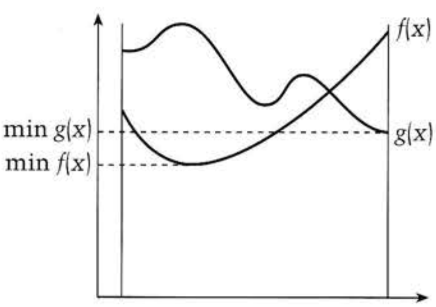


Figure 1

Indeed, the minimum of the sum $f(x) + g(x)$ is attained at a certain point of the interval $[a, b]$. The values of $f(x)$ and $g(x)$ at this point are not less than the respective minima of these functions. It's clear that the inequality becomes an equality if the minima of $f(x)$ and $g(x)$ are attained at the same point.

In the case of n functions, we have the same situation:

$$\min f_1(x) + \ldots + \min f_n(x) \le \min (f_1(x) + \ldots + f_n(x)).$$

We can prove this easily by induction. We single out one function after another on the right-hand side to obtain

$$\min (f_1(x) + \ldots + f_n(x)) \ge \min (f_1(x) + \ldots + f_{n-1}(x)) + \min f_n(x),$$

and so on.

Note that a similar principle holds for maxima:

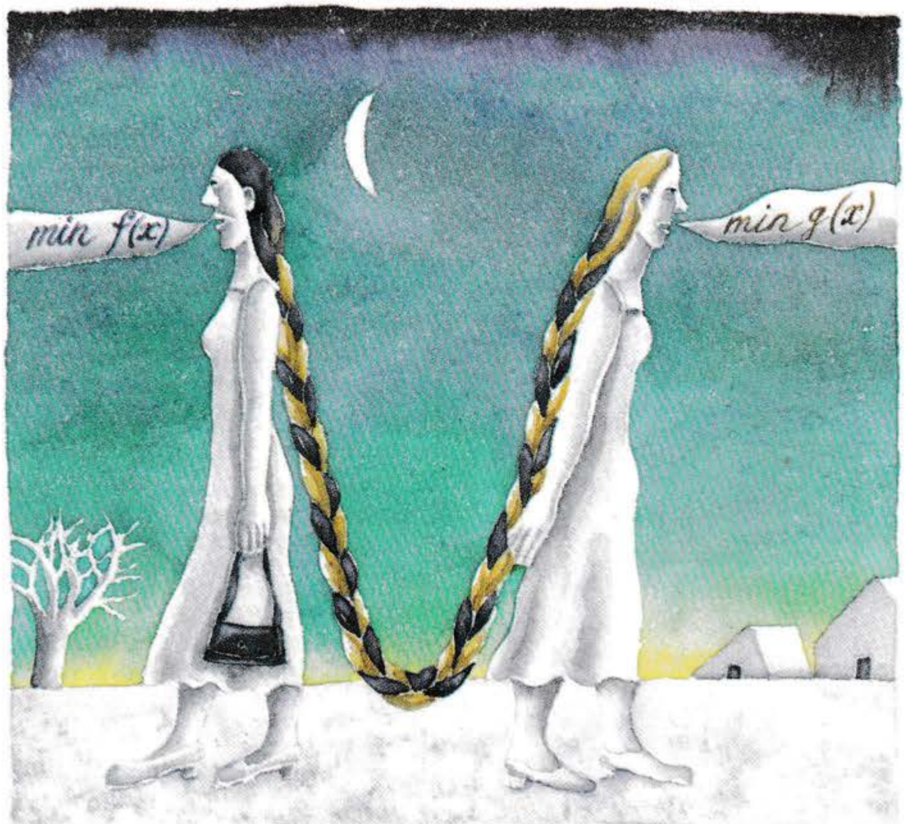
$$\max (f_1(x) + \ldots + f_n(x)) \le \max f_1(x) + \ldots + \max f_n(x).$$

Minima of quadratic functions and quadratic inequalities

We begin with the functions $f(x) = ax^2 + 2bx$, where $a > 0$. To find the minimum, let's complete the square:

$$ax^2 + 2bx = a\left(x^2 + 2\frac{b}{a}x + \frac{b^2}{a^2} - \frac{b^2}{a^2}\right)$$

$$= a\left(\left(x + \frac{b}{a}\right)^2 - \frac{b^2}{a^2}\right) = a\left(x + \frac{b}{a}\right)^2 - \frac{b^2}{a}.$$



Art by Sergey Ivanov

Thus the minimum equals

$$f\left(-\frac{b}{a}\right)=-\frac{b^2}{a}$$

and is attained at the point

$$x=-\frac{b}{a}.$$

Now consider n functions

$$f_1(x) = a_1x^2 + 2b_1x,$$
$$f_2(x) = a_2x^2 + 2b_2x,$$
$$\dots,$$
$$f_n(x) = a_nx^2 + 2b_nx$$

and apply to them the basic inequality. The function $f_i(x)$ attains its minimum,

$$-\frac{b_i^2}{a_i},$$

at the point

$$x=-\frac{b_i}{a_i}.$$

The minimum of the sum is attained at the point

$$x=-\frac{b_1+\ldots+b_n}{a_1+\ldots+a_n}$$

and equals

$$-\frac{(b_1+\ldots+b_n)^2}{a_1+\ldots+a_n}.$$

As a result, we obtain the inequality

$$\frac{b_1^2}{a_1}+\ldots+\frac{b_n^2}{a_n}\ge\frac{(b_1+\ldots+b_n)^2}{a_1+\ldots+a_n}$$

(for the sake of convenience, we changed the signs on the left- and right-hand sides).

Now we can easily prove several well-known inequalities.

Exercise 1. From the above formula derive

(a) the Cauchy–Bunyakovsky–Schwarz inequality

$$\sqrt{\left(c_1^2+\ldots+c_n^2\right)\left(d_1^2+\ldots+d_n^2\right)}\ge c_1d_1+\ldots+c_nd_n;$$

(b) the inequality between the quadratic and arithmetic means

$$\sqrt{\frac{b_1^2+\ldots+b_n^2}{n}}\ge\frac{b_1+\ldots+b_n}{n};$$

(c) the inequality between the arithmetic and harmonic means

$$\frac{a_1+\ldots+a_n}{n}\ge\frac{n}{\frac{1}{a_1}+\ldots+\frac{1}{a_n}}.$$

Hints: (a) Make the change of variables $a_1 = c_1^2, \dots, a_n = c_n^2$, $b_1 = c_1d_1, \dots, b_n = c_nd_n$; (b) set $a_1 = a_2 = \dots = a_n = n$; (c) set $b_1 = b_2 = \dots = b_n = 1$.

Exponential functions

In this section, we consider a more complicated function $f(x) = ae^x - bx - b$, where $a > 0$ and $b > 0$.

Exercise 2. Find the minimum value of this function.

Answer. The minimum value is

$$-b\ln\frac{b}{a},$$

and it is attained at

$$x=\ln\frac{b}{a}.$$

Thus, for n functions

$$f_1 = a_1e^x - b_1x - b_1,$$
$$\dots,$$
$$f_n = a_ne^x - b_nx - b_n,$$

(where all coefficients a_i and b_i are positive), we obtain from the basic inequality

$$b_1\ln\frac{b_1}{a_1}+\ldots+b_n\ln\frac{b_n}{a_n}$$
$$\ge(b_1+\ldots+b_n)\ln\frac{b_1+\ldots+b_n}{a_1+\ldots+a_n}.$$

Now a bit of algebraic transformation yields

$$\left(\frac{b_1}{a_1}\right)^{b_1}\cdot\ldots\cdot\left(\frac{b_n}{a_n}\right)^{b_n}\ge\left(\frac{b_1+\ldots+b_n}{a_1+\ldots+a_n}\right)^{b_1+\ldots+b_n}.$$

Exercise 3. From the above formula derive

(a) Cauchy's inequality between the arithmetic and geometric means

$$\frac{a_1+\ldots+a_n}{n}\ge\sqrt[n]{a_1\cdot\ldots\cdot a_n};$$

(b) the elegant inequality

$$b_1^{b_1}\cdot\ldots\cdot b_n^{b_n}\ge\left(\frac{b_1+\ldots+b_n}{n}\right)^{b_1+\ldots+b_n};$$

(c) the very useful inequality

$$c_1b_1+\ldots+c_nb_n\ge c_1^{b_1}\cdot\ldots\cdot c_n^{b_n}$$

if $b_1 + \dots + b_n = 1$.

Hints: (a) set $b_1 = b_2 = \dots = b_n = 1/n$; (b) set $a_1 = a_2 = \dots = a_n = 1$; (c) set

$$c_1=\frac{a_1}{b_1},\ \ldots,\ c_n=\frac{a_n}{b_n}.$$

There are certainly many other methods of proving these inequalities, but every proof is based on a different idea. Here all the inequalities were obtained by a unified method. (Another rather general method of deriving inequalities was described in the November/December 1999 issue of *Quantum* in the article "Obtaining Symmetric Inequalities.")

Functions of two variables and Hölder's inequality

The basic principle—the sum of the minima of several functions does not exceed the minimum of their sum (and the sum of their maxima is not less than the maximum of their sum) remains valid for functions of several variables. The proof of the principles is also the same.

Consider the simplest linear function of two variables:

$$f(x, y) = ax + by.$$

Surprisingly, this simple function makes it possible to derive nontrivial inequalities.

Problem. Prove the inequalities:

(a)

$$\sqrt{a_1^2+b_1^2}+\ldots+\sqrt{a_n^2+b_n^2}$$
$$\ge\sqrt{(a_1+\ldots+a_n)^2+(b_1+\ldots+b_n)^2};$$

(b)

$$\left(a_1^p+\ldots+a_n^p\right)^{\frac{1}{p}}\cdot\left(b_1^q+\ldots+b_n^q\right)^{\frac{1}{q}}$$
$$\ge a_1b_1+\ldots+a_nb_n$$

if $a_i \ge b_i$, $b_i \ge 0$, for all i, and if $p > 0$ and $q > 0$ are such that $1/p + 1/q = 1$ (Hölder's inequality).

Solution. Consider n functions

$$f_1(x, y) = a_1x + b_1y,$$
$$\dots,$$
$$f_n(x, y) = a_nx + b_ny.$$

The main thing is to make a proper choice of the set on which the functions will be considered. Since we deal with functions of two variables defined on a plane, we can consider their minimum on any geometric figure.

(a) Consider a circle of unit radius centered at the origin of the coordinate system (figure 2). What is the

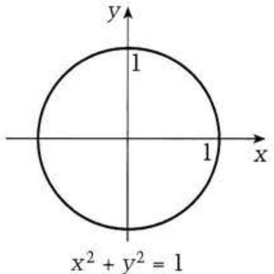


Figure 2

geometric meaning of the function $f(x, y) = ax + by$? It's clear that it is the scalar product of the vectors (a, b) and (x, y). But the scalar product equals the product of the vector's lengths and the cosine of the angle between them! The cosine is maximal for the zero angle, and the length of the variable vector (x, y) does not vary when its endpoint lies on a circle. Thus the function $f(x, y)$

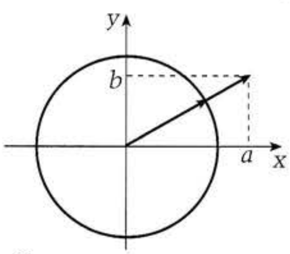


Figure 3

attains its maximum on the vector (x_0, y_0) that is collinear to (a, b) (figure 3). This vector is equal to

$$(x_0, y_0) = \left(\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}}\right)$$

and the scalar product is

$$(x_0, y_0)\cdot(a, b) = \sqrt{a^2+b^2}.$$

Thus, in the case of the unit circle, we have

$$\max(ax+by) = \sqrt{a^2+b^2}.$$

Now we can easily prove inequality (a). On the left-hand side of this inequality, we have the sum of the maximum values of the functions $f_1, f_2, \ldots, f_n$ on the unit circle; on the right-hand side, we have the maximum of the sum of these functions.

(b) Consider the set consisting of points (x, y), $x, y > 0$, for which

$$x^{\frac{1}{p}} \cdot y^{\frac{1}{q}} = 1.$$

Let's find the minimum of the function $f(x, y) = ax + by$ on this set. This can be done analytically (we invite the reader to do it), but a result obtained earlier in this article allows us to find a simpler solution. Let's apply inequality (c) in exercise 3:

$$ax + by = \frac{1}{p}\cdot pax + \frac{1}{q}\cdot qby$$
$$\geq (pax)^{\frac{1}{p}}\cdot(qby)^{\frac{1}{q}} = (pa)^{\frac{1}{p}}\cdot(qb)^{\frac{1}{q}}.$$

Verify that this inequality becomes an equality when $x/y = b/a$. Thus the minimum of $f(x, y)$ on the set under consideration is

$$(pa)^{\frac{1}{p}}\cdot(qb)^{\frac{1}{q}}.$$

If we now set

$$a = \frac{c^p}{p}$$

and

$$b = \frac{d^q}{q};$$

that is, if we consider the linear function

$$\frac{c^p}{p}x + \frac{d^q}{q}y,$$

this minimum will be cd. Setting

$$a_1 = \frac{c_1^p}{p},\ b_1 = \frac{d_1^q}{q},\ \ldots,\ a_n = \frac{c_n^p}{p},\ b_n = \frac{d_n^q}{q}$$

for each of the functions $f_1, f_2, \ldots, f_n$, we obtain Hölder's inequality!

We now want to compare the inequalities of Cauchy-Bunyakovsky and of Hölder. It's clear that the first one is a particular case of the second for $p = q = 2$. Thus if we define a "length" for the vector (a, b) as

$$|(a, b)|_p = \left(a^p + b^p\right)^{\frac{1}{p}},$$

then we find that the scalar product of two vectors does not exceed the product of their "lengths."

This is the geometric interpretation and the meaning of Hölder's inequality.

On the basis of the method described, we have proved a number of inequalities (including several classical ones) using only a modest set of functions. We invite the reader to try proving other inequalities using this method.

Exercise 4. Prove the following inequalities for positive numbers $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$:

(a)

$$\sqrt[n]{a_1 \ldots a_n} + \sqrt[n]{b_1 \ldots b_n}$$
$$\leq \sqrt[n]{(a_1+b_1)\ldots(a_n+b_n)};$$

(b)

$$(1+a_1)\ldots(1+a_n) \geq \left(1+\sqrt[n]{a_1\ldots a_n}\right)^n$$

(Huygens' inequality);

(c)

$$\left(a_1^p + \ldots + a_n^p\right)^{\frac{1}{p}}\cdot\left(b_1^q + \ldots + b_n^q\right)^{\frac{1}{q}}$$
$$\leq a_1b_1 + \ldots + a_nb_n$$

if $q < 0$ and $1/p + 1/q = 1$ (the inverse Hölder inequality);

(d)

$$\left(a_1^p + \ldots + a_n^p\right)^{\frac{1}{p}} + \left(b_1^p + \ldots + b_n^p\right)^{\frac{1}{p}}$$
$$\geq \left((a_1+b_1)^p + \ldots + (a_n+b_n)^p\right)^{\frac{1}{p}},$$

where $p > 1$ (Minkowski's inequality). Prove that for $p < 1$ the inequality is in the opposite sense.

Hints: (a) Consider the linear function $y = a_1x_1 + \ldots + a_nx_n$ of n variables on the set consisting of points $(x_1, \ldots, x_n)$ such that $x_1 \cdot \ldots \cdot x_n = n$ and use the inequality between the arithmetic and geometric means to find the minimum. (b) Set $b_1 = b_2 = \ldots = b_n = 1$ in the preceding inequality. (d) Again, consider the linear function $y = a_1x_1 + \ldots + a_nx_n$ on the set defined by the condition

$$x_1^q + \ldots + x_n^q = 1,$$

where the number q is determined by the condition

$$\frac{1}{p} + \frac{1}{q} = 1$$

and use Hölder's inequality to find the minimum. ◘

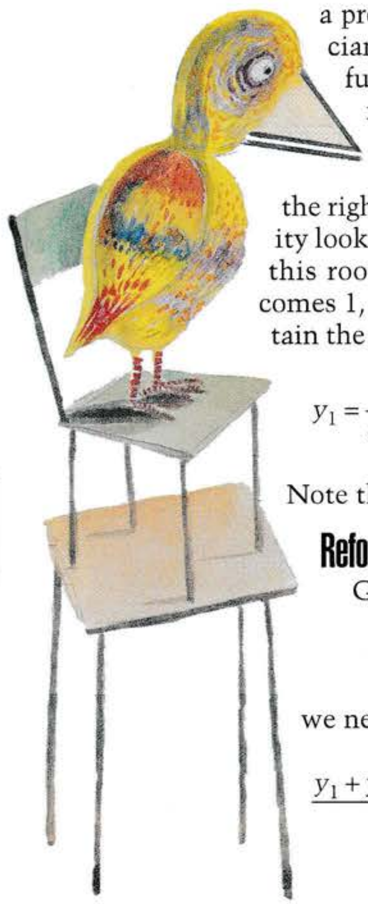
Cauchy and induction

by Y. Solovyov

THE ARITHMETIC MEAN OF A SET OF POSItive numbers is not less than their geometric mean:

$$\frac{x_1 + x_2 + \ldots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \ldots x_n}.$$

This famous inequality, first proved by the French mathematician Cauchy, was published in 1821. Since then it has considered one of the most difficult numerical inequalities. Over 180 years a few dozen different proofs of this inequality were found. The proof given by Cauchy occupied several pages of complicated manipulations. I propose that you learn the simplest proof I know. It's both elegant and instructive, I think. But first we need to reformulate the Cauchy inequality.



What is the purpose of reformulating a problem? Experienced mathematicians know that it can be very useful. Simple proofs of complex theorems often consist of a chain of reformulations.

At the outset, the nth root on the right-hand side of the given inequality looks scary. Let's divide both sides by this root. Then the right-hand side becomes 1, and on the left-hand side we obtain the arithmetic mean of the numbers

$$y_1 = \frac{x_1}{\sqrt[n]{x_1 x_2 \ldots x_n}}, \ \ldots, \ y_n = \frac{x_n}{\sqrt[n]{x_1 x_2 \ldots x_n}}.$$

Note that $y_1, y_2, y_3, \ldots, y_n = 1$.

Reformulation

Given the conditions

$$y_1 > 0, \ldots, y_n > 0,$$
$$y_1 \ldots y_n = 1,$$

we need to prove that

$$\frac{y_1 + y_2 + \ldots + y_n}{n} \geq 1 \qquad (*)$$

Now recall the method of mathematical induction. We will prove the inequality step by step, increasing n by one every time. For $n = 1$, the inequality is evident (it becomes a strict equality). Assume that we have managed to prove it for a certain n and try derive it for $n + 1$.

The inductive step

We need to prove that if

$$z_1 > 0, \ldots, z_n > 0, z_{n+1} > 0,$$
$$z_1 \ldots z_n z_{n+1} = 1,$$

then $z_1 + \ldots + z_n + z_{n+1} \geq n + 1$. We operate under the assumption that inequality $(*)$ holds for any suitable set of n numbers. Let $y_1 = z_1, \ldots, y_{n-1} = z_{n-1}, y_n = z_n z_{n+1}$. Then both conditions

$$y_1 > 0, \ldots, y_n > 0,$$
$$y_1 \ldots y_n = 1.$$

hold, and we assume that the inequality $y_1 + \ldots + y_n \geq n$ is proved; that is,

$$z_1 + \ldots + z_n \cdot z_{n+1} \geq n. \qquad (**)$$

Note that up to this point we only made a series of reformulations of the problem. But where is the proof? Here it is.

If necessary, reorder the numbers $z_1, \ldots, z_{n+1}$ so that $z_n > 1$ and $z_{n+1} < 1$. This is clearly possible if not all z_i are equal to one. Now add 1 to both sides of inequality $(**)$ and replace $z_n \cdot z_{n+1} + 1$ by the sum $z_n + z_{n+1}$ on the left-hand side. To justify this replacement, we must prove that $z_n + z_{n+1} \geq z_n \cdot z_{n+1} + 1$, or

$$z_n + z_{n+1} - z_n \cdot z_{n+1} - 1 \geq 0.$$

The left-hand side factores as

$$z_n(1 - z_{n+1}) - (1 - z_{n+1}) = (z_n - 1)(1 - z_{n+1}) \geq 0.$$

This inequality is certainly true, since $z_n > 1$ and $z_{n+1} < 1$. This concludes the proof.

Now take another look at the proof. The only nontrivial step in it is the choice of the numbers z_n and z_{n+1}. Everything else is the use of the mathematical induction and a chain of reformulations. ◘

Art by Leonid Tishkov

Calculus and inequalities

by V. Ovsienko

HENRI POINCARÉ ONCE SAID THAT MATHematics is a way of making the same statement about completely different things. This is especially true of calculus, which is a sort of lock-picking tool that will open many doors to the astute problem-solver.

We'll consider three different problems and solve them by the same method. The method is based on a well-known fact: If the derivative of a function is positive, then it is an increasing function (and vice versa—if its derivative is negative, it is a decreasing function).

Inequalities involving sines or cosines

The inequalities

$$\sin x < x,$$

$$\cos x > 1 - \frac{x^2}{2},$$

which are valid for $x > 0$ (figure 1), are traditionally proved in school by geometric means. However, the analytic proof is simpler and doesn't require much inventiveness.

Consider the first inequality. The functions $y_1(x) = \sin x$ and $y_2(x) = x$ coincide at zero. Therefore, it's sufficient to prove that the second function grows more rapidly than the first one for $x > 0$ (in other words, we need to prove that $y(x) = x - \sin x$ is an increasing function). But the derivative $y'(x) = 1 - \cos x$ is greater than zero. Thus the inequality is proved.

To prove the second inequality, we proceed in exactly the same way. The function

$$y(x) = \cos x - 1 + \frac{x^2}{2}$$

is zero for $x = 0$. Its derivative is $y'(x) = -\sin x + x$. Thus, $y'(x) > 0$ for $x > 0$ (by virtue of the first inequality); that is, $y(x)$ is an increasing function and, therefore, is positive.

Art by Leonid Tishkov

Exercises

1. Prove that for $x > 0$,

(a) $\sin x > x - \frac{x^3}{6}$;

(b) $\cos x < 1 - \frac{x^2}{2} + \frac{x^4}{24}$.

Hint: follow the above proofs. After differentiating, use the preceding inequalities.

2. Continue the chain of inequalities for $\sin x$ and $\cos x$.

Inequalities in a triangle

For a triangle with angles α, β, and γ (figure 2), the following inequality holds:

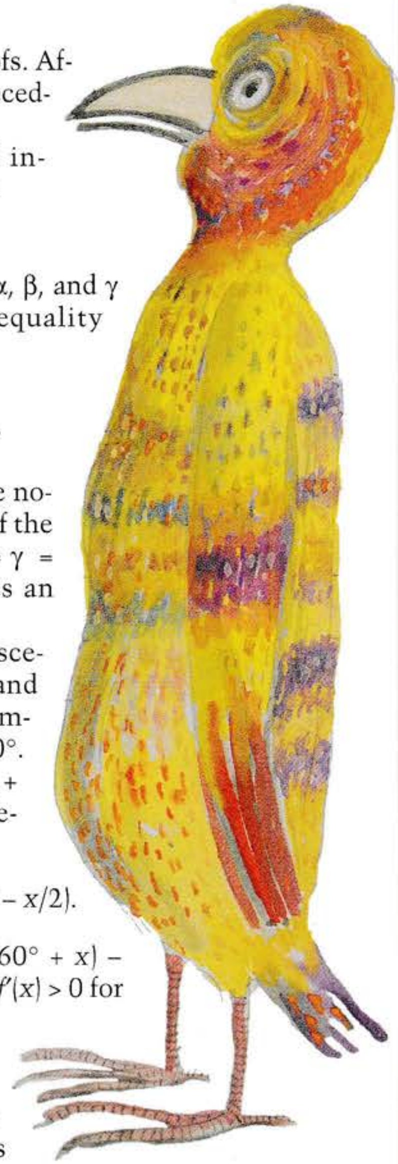
$$\sin\alpha + \sin\beta + \sin\gamma \le \frac{3\sqrt{3}}{2}.$$

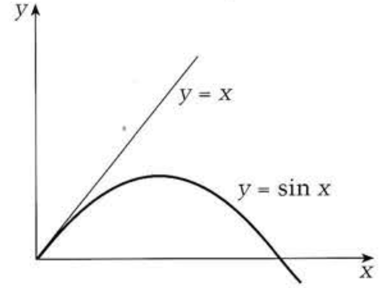
Here is a proof. First of all, we notice that the inequality holds if the triangle is equilateral ($\alpha = \beta = \gamma = 60°$)—in this case it becomes an equality.

Now let the triangle be isosceles with angles $\alpha = 60° + x$ and $\beta = \gamma = 60° - x/2$, where x is a number such that $-60° < x < 120°$. Then, the expression $\sin\alpha + \sin\beta + \sin\gamma$ is a function depending on x:

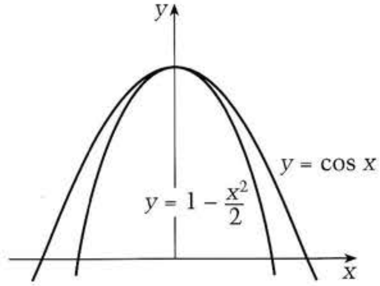
$$f(x) = \sin(60° + x) + 2\sin(60° - x/2).$$

Its derivative is $f'(x) = \cos(60° + x) - \cos(60° - x/2)$. It's clear that $f'(x) > 0$ for $-60° < x < 0$ and $f'(x) < 0$ for $0 < x < 120°$. Thus the function $f(x)$ attains its maximum at the point $x = 0$ and is smaller the more x deviates from zero.





a



b

Figure 1

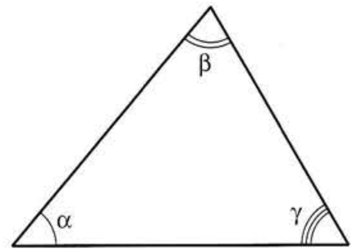


Figure 2

Now let α be fixed and the angles β and γ vary: $\beta = \delta - x$ and $\gamma = \delta + x$. Again, the given expression is a function of x:

$$f(x) = \sin(\delta - x) + \sin(\delta + x) + \sin\alpha.$$

Its derivative is $f'(x) = \cos(\delta + x) - \cos(\delta - x)$. Again, $f'(x)$ is greater than zero for negative x and is less than zero for positive x. In other words, the expression $\sin\alpha + \sin\beta + \sin\gamma$ is smaller the more the triangle under consideration differs from an isosceles triangle.

The inequality is thus proved, since any triangle can be obtained from an equilateral triangle by varying the angle α first and then the angles β and γ. We proved that under such a variation, the expression $\sin\alpha + \sin\beta + \sin\gamma$ always decreases.

Exercise 3. Prove the inequalities

(a) $\cos\alpha + \cos\beta + \cos\gamma \le \dfrac{3}{2}$;

(b) $\sin\alpha \cdot \sin\beta \cdot \sin\gamma \le \dfrac{3\sqrt{3}}{8}$;

(c) $\cos\alpha \cdot \cos\beta \cdot \cos\gamma \le \dfrac{1}{8}$

if α, β, and γ are the angles of a triangle.

Cauchy's inequality

For any nonnegative numbers $a_1, \ldots, a_n$, the inequality between the arithmetic and geometric means

$$\frac{a_1 + \ldots + a_n}{n} \ge \sqrt[n]{a_1 \cdot \ldots \cdot a_n}$$

holds. This is called *Cauchy's inequality*. We prove it by the same method as above—that is, by calculating the derivative of a function.

To begin with, consider the simple case of $n = 2$. Certainly this inequality can be proved by various geometric means (some of them are very elegant), but we'll use calculus to prove it. For definiteness, let $a \le b$. Then $b = a + x$, where x is a nonnegative number. Now the expression

$$\frac{a+b}{2} - \sqrt{ab}$$

is a function of x:

$$f(x) = \frac{2a + x}{2} - \sqrt{a(a+x)}.$$

For $x = 0$, $f(x) = 0$ and the derivative

$$f'(x) = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{a}{a+x}} > 0.$$

Thus $f(x)$ is an increasing function, the arithmetic mean is greater than the geometric mean, and the difference between them is greater the larger the difference $b - a = x$.

Now we prove Cauchy's inequality for the general case. In fact, this is a repeat of the proof for $n = 2$. We only need to know the derivative of the function

$$\sqrt[n]{x^k} = x^{\frac{k}{n}}.$$

This derivative can be found in any textbook on calculus and is

$$\left(x^{\frac{k}{n}}\right)' = \frac{k}{n} x^{\frac{k}{n} - 1}.$$

For definiteness, let $a_1 \le a_2 \le \ldots \le a_n$. It's clear that for $a_1 = \ldots = a_n$, the inequality becomes an equality. Let's take n equal numbers $a_1 = a_1 = \ldots = a_1$ and increase them step by step beginning from the second number. The process takes $(n - 1)$ steps. At the first step, we have

$$a_1,\ \underbrace{a_1 + z,\ \ldots,\ a_1 + z}_{n-1},$$

where $a_1 + z = a_2$.

At the second step, we have

$$a_1,\ a_2,\ \underbrace{a_2 + y,\ \ldots,\ a_2 + y}_{n-2},$$

where $a_2 + y = a_3$.

Finally, we obtain our given set of numbers $a_1, a_2, \ldots, a_n$.

We'll prove that at every step the difference between the arithmetic and geometric means increases. Thus we'll not only prove the Cauchy inequality but also see its "dynamics." At the mth step, we have the following set of numbers:

$$a_1,\ \ldots,\ a_m,\ \underbrace{a_m + x,\ \ldots,\ a_m + x}_{n-m}.$$

CONTINUED ON PAGE 48

Ptolemy's trigonometry

by V. Zatakavai

ABOUT 2,000 YEARS AGO, while performing astronomical observations, the ancient Greeks noticed that the length of a chord of a circle depended on the degree-measure of the subtended arc. The Greek astronomer Ptolemy measured the length of a chord with a given central angle, then took the ratio of this length to the radius of the circle, and drew up a table. If the central angle is θ, and the length of the chord is c, then $c = 2R \sin \theta$ (see figure 1). So Ptolemy was in fact making a table of $\sin 2\theta$ for $0° \le \theta \le 90°$.

Ptolemy also extended his table to chords of obtuse angles. For example, if he wanted the length of AC (figure 2), he first computed the length of BC, then, noting that AB is a diameter, Ptolemy used the Pythagorean theorem to get AC. Since $AC = 2R \sin \alpha$, and $BC = 2R \sin \beta$, this is equivalent to using the relationship $\sin^2 \alpha + \cos^2 \alpha = 1$.

Another of Ptolemy's achievements was the theorem that now bears his name—Ptolemy's theorem.

Theorem. *The product of the diagonals of a quadrilateral inscribed in a circle equals the sum of the products of its opposite sides* (figure 3).

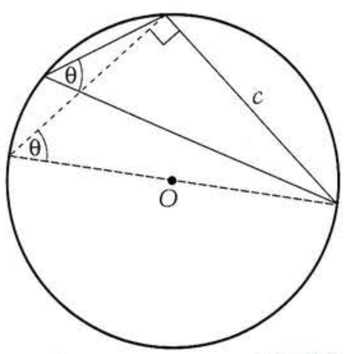


Figure 1

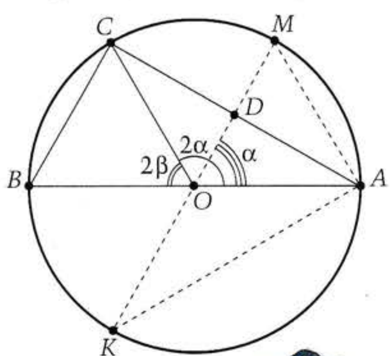


Figure 2

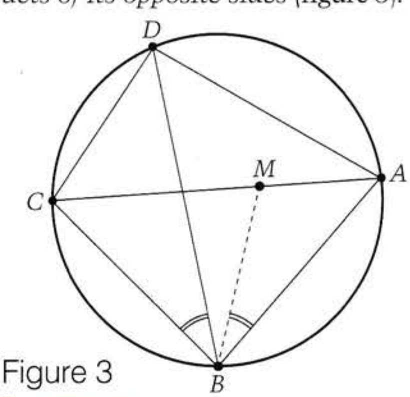
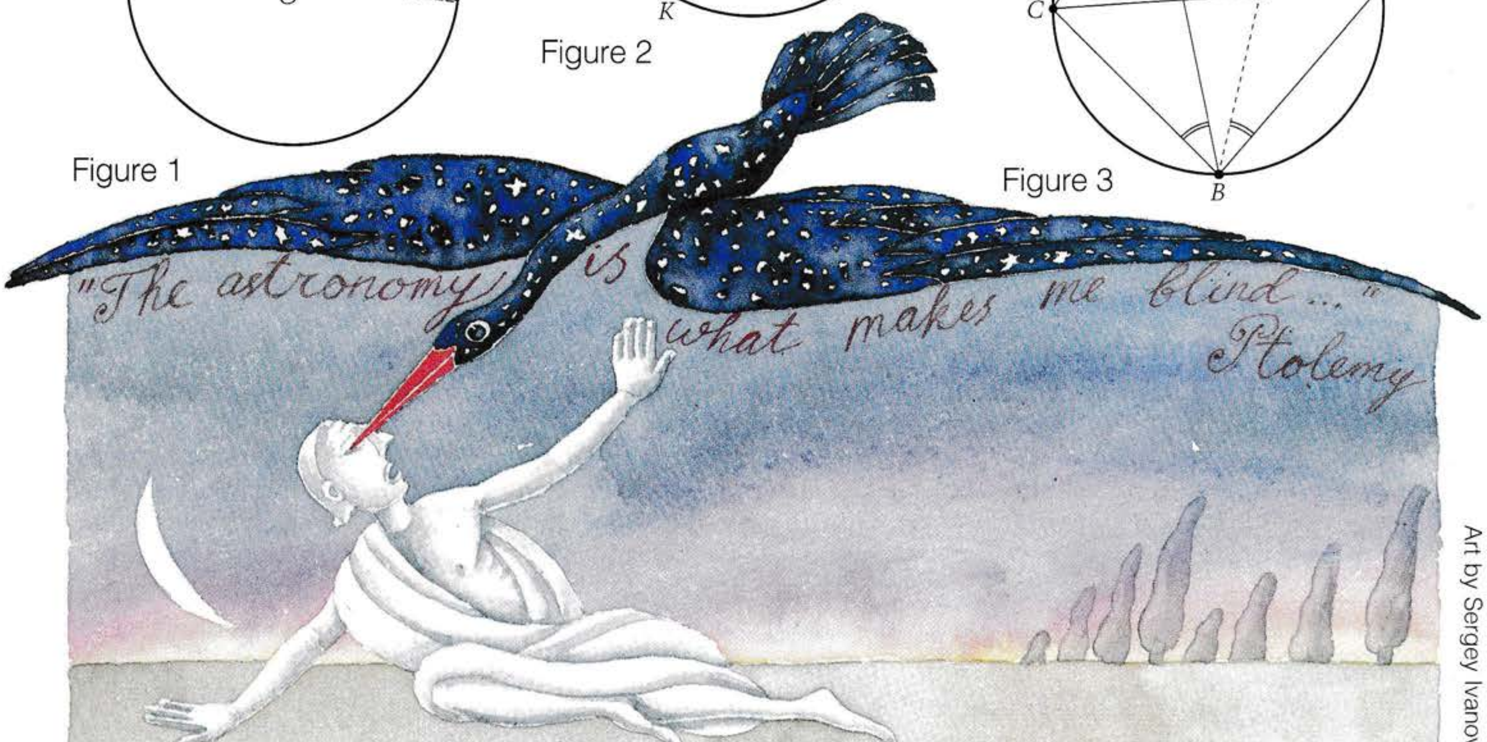


Figure 3



Art by Sergey Ivanov

Proof. Consider a point M on diagonal AC such that $\angle ABM = \angle CBD$. Since $\angle CDB$ and $\angle MAB$ are both inscribed in the circle, and both intercept arc BC, $\angle CDB = \angle MAB$. Hence triangles BCD and ABM are similar. Thus,

$$\frac{BD}{AB} = \frac{CD}{AM},$$

so

$$AB \cdot CD = AM \cdot BD. \qquad (1)$$

Now $\angle ABD = \angle MBC$ by construction, and $\angle BCM = \angle ADB$ (both are inscribed angles intercepting arc AB). Thus triangles ABD and MBC are similar.
Hence

$$\frac{AD}{CM} = \frac{BD}{BC},$$

and

$$AD \cdot BC = BD \cdot CM. \qquad (2)$$

Adding equations (1) and (2) term by term, we obtain

$$AB \cdot CD + AD \cdot BC = BD(AM + CM) = BD \cdot AC,$$

which is what we wanted to prove.

We can use Ptolemy's theorem to derive a number of trigonometric formulas.

Let us take diameter AB of a given circle, and choose points C and D on one of the semicircles it determines

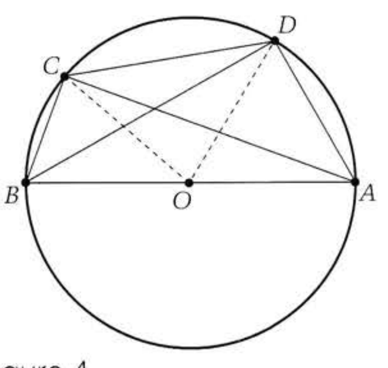


Figure 4

(figure 4). Let $\angle AOC = 2\alpha$ and $\angle AOD = 2\beta$; then $AC = 2R \sin \alpha$, $AD = 2R \sin \beta$, $BC = 2R \cos \alpha$, $BD = 2R \cos \beta$, and $CD = 2R \sin(\alpha - \beta)$.

Using Ptolemy's theorem and taking into account that $AB = 2R$, we obtain

$$2R \cdot 2R \sin(\alpha - \beta) + 2R \cos \alpha \cdot 2R \sin \beta = 2R \sin \alpha \cdot 2R \cos \beta,$$

or

$$\sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta.$$

We invite the reader to derive a similar formula for $\sin(\alpha + \beta)$. (Hint: inscribe the quadrilateral in a circle such that its diagonal is a diameter.)

Thus we have derived two important trigonometric formulas:

$$\sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta,$$
$$\sin(\alpha + \beta) = \sin \alpha \cos \beta + \cos \alpha \sin \beta.$$

Combine them term by term to obtain

$$\sin(\alpha - \beta) + \sin(\alpha + \beta) = 2 \sin \alpha \cos \beta,$$

or

$$\sin \alpha \cos \beta = {}^1/_2(\sin(\alpha - \beta) + \sin(\alpha + \beta)). \qquad (3)$$

Combining term by term the formulas for the cosine of the sum (or difference) of two angles (try to derive them yourself), we obtain two more formulas:

$$\cos \alpha \cos \beta = {}^1/_2(\cos(\alpha - \beta) + \cos(\alpha + \beta)), \qquad (4)$$

$$\sin \alpha \sin \beta = {}^1/_2(\cos(\alpha - \beta) - \cos(\alpha + \beta)). \qquad (5)$$

Formulas (3), (4), and (5) are called the *products-to-sums* formulas. We can derive from them some corresponding *sums-to-products* formulas, by a change of variables.

Let $\alpha = (x + y)/2$, $\beta = (x - y)/2$, and substitute into formula 3. We obtain

$$2\sin\frac{x+y}{2}\cos\frac{x-y}{2} = \sin x + \sin y.$$

Similarly, we can derive formulas for $\sin x - \sin y$, $\cos x + \cos y$, and $\cos x - \cos y$ as products. These formulas are very useful for proving trigonometric identities.

Example 1. Calculate, without using tables or a calculator, the value of the expression

$$\frac{1}{2\sin 10^\circ} - 2\sin 70^\circ.$$

Solution. The given expression is equal to

$$\frac{1 - 4\sin 70^\circ \sin 10^\circ}{2\sin 10^\circ}$$
$$= \frac{1 - 2\left(\cos 60 - \cos 80^\circ\right)}{2\sin 10^\circ}$$
$$= \frac{2\cos 80^\circ}{2\sin 10^\circ} = 1.$$

Example 2. Calculate, without using tables or a calculator, the value of the expression $16 \sin 20^\circ \sin 40^\circ \sin 60^\circ \sin 80^\circ$.

Solution. Using the well-known values of $\sin 60^\circ$ and $\cos 60^\circ$, we find that the given expression is equal to

$$16\sin 20^\circ \sin 40^\circ \frac{\sqrt{3}}{2}\sin 80^\circ$$
$$= 4\sqrt{3}\left(\cos 20^\circ - \cos 60^\circ\right)\sin 80^\circ$$
$$= 4\sqrt{3}\left(\cos 20^\circ \sin 80^\circ - \frac{1}{2}\sin 80^\circ\right)$$
$$= 4\sqrt{3}\left(\frac{1}{2}\left(\sin 100^\circ + \sin 60^\circ\right) - \frac{1}{2}\sin 80^\circ\right)$$
$$= 2\sqrt{3}\left(\sin 80^\circ + \frac{\sqrt{3}}{2} - \sin 80^\circ\right) = 3.$$

Example 3. Solve the equation

$$4 \sin 2x \sin 5x \sin 7x = \sin 4x.$$

Solution. The given equation is equal to

$$4 \sin 2x \sin 5x \sin 7x - 2 \sin 2x \cos 2x = 0,$$

or

$$2 \sin 2x (2 \sin 5x \sin 7x - \cos 2x) = 0,$$

or

$$\sin 2x = 0 \to x = \frac{\pi}{2}n,\ n \in Z;$$

or

$$\cos 12x = 0 \to x = \frac{\pi}{24}(2n+1),\ n \in Z.$$

Here are some more exercises.

1. Calculate, without using tables or a calculator, the value of the expressions

(a) $\dfrac{1}{2\cos 20^\circ} - 2\sin 50^\circ$;

(b) $\cos 10^\circ \sin 20^\circ \cos 50^\circ$;

(c) $\tan 20^\circ \tan 40^\circ \tan 60^\circ \tan 80^\circ$.

2. Solve the equations

(a) $2 \cos 2x \sin x + \sin 2x \cos x = \sin 4x \cos x$;

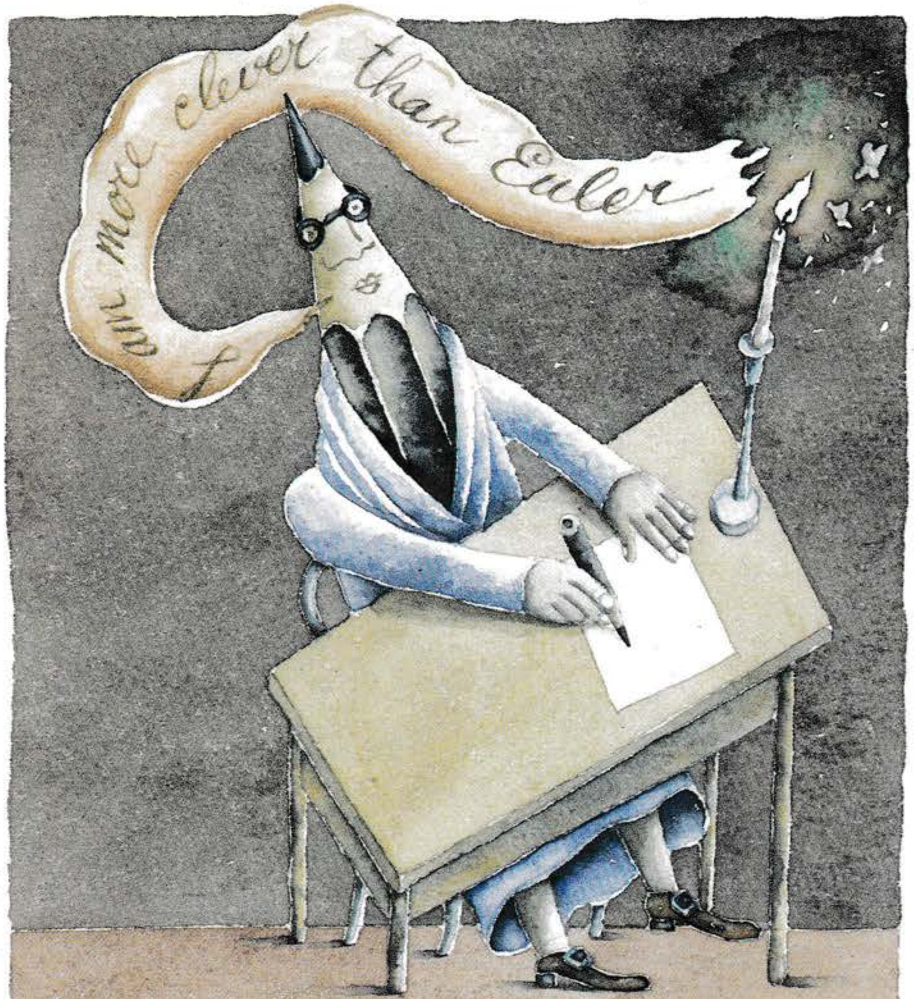
(b) $\sin 2x \sin x + \cos^2 x = \sin 5x \sin 4x + \cos^2 4x$;

(c) $4 \sin x \sin 2x \sin 3x = \sin 4x$.

Problems teach us how to think

by V. Proizvolov

YOU CAN'T LEARN MATHematics by watching someone else do it. Active learning involves solving problems of increasing difficulty—if you keep solving problems of the same difficulty, they simply become routine exercises. If a particular problem puts up a lot of resistance you can stare at the ceiling or knit your brow (there's no law against it), but the best thing to do is to take a sheet of paper and a pencil and start experimenting: make some estimates, consider particular cases, sketch out your ideas, and so on. Leonard Euler once said, "My pencil is sometimes more clever than my head."



Art by Sergey Ivanov

"Freeze, die, come back to life"—this catch-phrase from a Russian children's game would be good advice for anyone who wants to solve a math problem. To overcome the problem you need to focus your attention deep into its conditions and its setting, until you reach the first glimmer of an idea and the hope of success. Solving a problem involves not only an intellectual challenge but a test of will—you need a "fighting spirit."

It's not necessary, (and not even possible) to solve all the problems known to mathematics. So you need to select whatever you find appealing, instructive, interesting, and within your powers. In the process, you'll develop your taste and become more mathematically "cultured."

Among other things, mathematics teaches you to be honest—to yourself and to others. You can't beat around the bush while answering a mathematical question. And honesty is a necessary condition for rigorous thinking. But besides that, in solving problems we not only learn how to prove what's true, we also learn how to guess at the truth. And the ability to guess is a necessary part of productive thinking.

The beautiful world of mathematical problems is continually being replenished, which shows that mathematics is indeed a living science.

Try to solve the following problems without looking at the solutions.

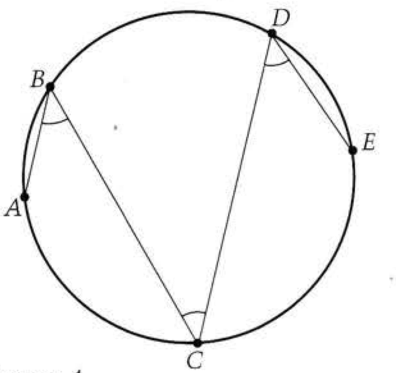


Figure 1

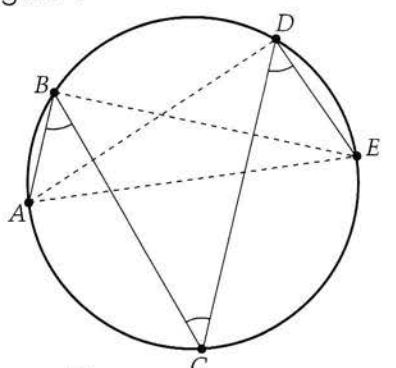


Figure 2

Problem 1. *Is it possible?* All the positive integers from 1 through 100 are written in ascending order. Twenty-five of them have been crossed out. Is it possible to cross out 25 more numbers such that the sum of all the remaining numbers equals the sum of all the numbers that were crossed out?

Solution. Yes, it's possible. The first hundred positive integers can be broken down into 50 pairs such that the sum of every pair is 101. Assume that there are k complete pairs in the 25 numbers that were crossed out; the remaining $25 - 2k$ numbers in this set do not create such a pair. To these $25 - 2k$ unpaired numbers we add $25 - 2k$ numbers that make pairs with them and cross them out. Then we cross out k pairs of numbers from those still not crossed out. We end up with a total of 25 complete pairs of crossed-out numbers, which amounts to half the sum of all positive integers from 1 to 100.

Problem 2. *Zigzag in a circle.* The vertices of the polygonal line $ABCD$ lie on a circle (figure 1). The measures of angles B, C, and D are each 45°. Prove that $AB^2 + CD^2 = BC^2 + DE^2$.

Solution. Since $\angle ABC = \angle EDC = 45°$, arcs AC and CE are each 90°, so AE is a diameter of the circle. Therefore, $AB^2 + BE^2 = AE^2$ and $AD^2 + DE^2 = AE^2$ (figure 2). Now, taking into account that $BE = CD$ and $AD = BC$ (both pairs of chords intercept equal arcs), we obtain the desired equality: $AB^2 + CD^2 = BC^2 + DE^2$.

Problem 3. *Three-digit numbers.* Three three-digit numbers are given. Their decimal notation includes all the digits except zero, and their sum is 1,665. The first digit of each number was interchanged with the last one to obtain three new three-digit numbers. What is the sum of the new numbers?

Solution. The answer is 1,665. The sum of the last digits of the three given numbers must be 5, 15, or 25. But it cannot be 5 or 25 because these numbers cannot be represented as the sum of three different one-digit numbers (from 1 to 9). Therefore, the sum of the three last digits is 15. Thus the sum of the middle digits is also 15, as is the sum of the first digits. Now it's clear that the sum of the numbers with the first and last digits interchanged remains the same—that is, 1,665.

Here's one possible triad of numbers that satisfies the condition of the problem: 159, 672, 834.

Problem 4. *Equal areas.* Three segments C_1A_2, C_2B_1, and A_1B_2 with endpoints lying on the sides of triangle ABC are parallel to the sides and pass through a point P (figure 3). Prove that the areas of triangles $A_1B_1C_1$ and $A_2B_2C_2$ are equal.

Solution. The areas of triangles A_1B_1P and A_2B_1P are equal since they have the same base and equal altitudes. The areas of triangles A_2B_2P and A_2B_1P are equal for the same reason. Thus the areas of triangles A_1B_1P and A_2B_2P are also equal. Similarly, we can prove that the areas of triangles C_1A_1P and C_2A_2P are equal, as are the areas of triangles C_1B_1P and C_2B_2P. Thus the areas of triangles $A_1B_1C_1$ and $A_2B_2C_2$ are equal.

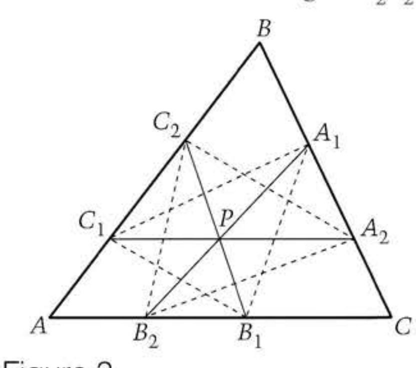


Figure 3

Problem 5. *A chess position.* A chess position possesses the following property: On every vertical and on every horizontal row, there is an odd number of pieces. Prove that there is an even number of pieces on black squares.

Solution. Let's renumber the vertical and horizontal rows of the chess board and put the letter A in the black squares of the vertical rows that received an odd "vertical" number. Put the letter B in all the other black squares (figure 4). Similarly, put the letter C in the white squares of the horizontal rows that received an odd "horizontal" number. Let the number of pieces on A squares be a, on B squares be b, and on C squares be c. The statement of the problem implies that both $a + c$ and $b + c$ are even. Therefore, $a + b$ is also even—that is, there is an even number of pieces on the black squares.

	B		B		B		B
A	C	A	C	A	C	A	C
	B		B		B		B
A	C	A	C	A	C	A	C
	B		B		B		B
A	C	A	C	A	C	A	C
	B		B		B		B
A	C	A	C	A	C	A	C

Figure 4

Problem 6. *Six points.* Find six points in the plane such that any five of them can be covered by two squares having a unit diagonal, but all six points cannot be covered by two circles having a unit diameter.

Solution. Place a unit square in the plane such that its sides are parallel to the coordinate axes. Mark its four vertices and two points inside

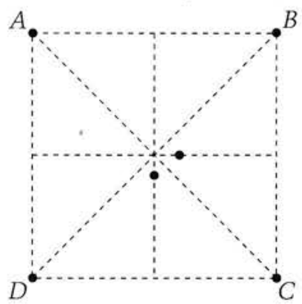


Figure 5

it: one 0.1 to the right of the center and the other 0.1 below the center (figure 5). It's easy to see that any five points of these six can be covered by two squares having a unit diagonal. Now notice that no three vertices of the square can be covered by a unit circle (indeed, if it were possible, the circle would have covered the square's diagonal). Thus each circle must cover two vertices of the square apiece. For example, let one of them cover the two left vertices and the other cover the two right vertices. Then the left and right sides of the square coincide with the diameters of these circles, and the point that is below the center would remain uncovered. Similarly, if one circle covers the two upper vertices and the other covers the two lower vertices, the point to the right of the center would remain uncovered.

Problem 7. *Chords in a circle.* Several chords are drawn in a circle such that each one of them passes through the midpoint of another. Prove that all the chords are diameters of the circle.

Solution. Consider the chord a that is farthest from the center of the circle O and call its midpoint A. By the statement of the problem, this chord contains the midpoint B of another chord b. Therefore, OB is not perpendicular to the original chord, so $OB \geq OA$. On the other hand, $OB \leq OA$ by the definition of chord a. Thus $OB = OA$, and the points A and B coincide. Since the midpoints of the different chords a and b coincide, they are diameters. Thus we see that the chord farthest from the center is a diameter. Therefore, all the given chords are also diameters.

Problem 8. *A set of weights.* A set of 100 weights is given, and it is known that they can be put in pairs such that the difference in the weights in each pair is always the same. Show that the weights can be distributed between the two pans of a scale, 50 weights per pan, so that the scale is in balance.

Solution. Divide the weights into the pairs described. Take any 25 pairs and place the lesser weight of each pair on the left pan of the scale and the other weight of the pair on the right pan. For the other 25 pairs of weights, we do it the other way around: Place the heavier weight on the left pan and the lighter weight on the right pan. The scale will be in balance.

Problem 9. *Calendar equation.* (a) Do natural numbers x, y, and z exist that satisfy the equation $28x + 30y + 31z = 365$? (b) Nonnegative integer numbers x, y, and z satisfy the equation $28x + 30y + 31z = 365$. Prove that $x + y + z = 12$.

Solution. (a) To give an affirmative answer to this question, it's sufficient to recall the calendar. An ordinary (not leap) year has 365 days: one month has 28 days ($x = 1$), four months have 30 days ($y = 4$), and seven months have 31 days ($z = 7$). This equation also has other solutions—for example, $x = 2$, $y = 1$, $z = 9$. (b) Assume that $x + y + z \leq 11$. Then $28x + 30y + 31z \leq 11 \cdot 31 = 341$. Thus $x + y + z > 11$. Assume that $x + y + z \geq 13$. Then $28x + 30y + 31z \geq 13 \cdot 28 = 364$, and the equality is possible only when $x = 13$ and $y = z = 0$; in all other cases, $28x + 30y + 31z \geq 366$. Therefore, $x + y + z = 12$.

Problem 10. *Constructing a rectangle.* A 2×2 square was cut into rectangles by lines parallel to its sides. The rectangles are painted black and white in a checkerboard pattern. It turned out that the total area of the black rectangles was equal to the total area of the white ones. Prove that it is possible to put the black rectangles together to form a 1×2 rectangle.

Solution. The lines cut the given rectangle into even and odd vertical and horizontal strips. Let's rearrange the vertical strips so that all the odd strips are on the left and all the even strips are on the right. Then we rearrange the horizontal strips so that all the odd strips are on the top and all the even strips are on the bottom (figure 6). Now the given rectangle has been broken down into four rectangles, and the sum of the areas of the black rectangles equals that of the white ones. It's clear that it is possible to produce a 1×2 rectangle from the black rectangles.

Problem 11. *Around the table.* Twelve people had a conference at a round table. After the break they sat at the table again, but in a different order. Prove that there are two persons who have same number of people between them as before the break (counting clockwise from the first person to the second).

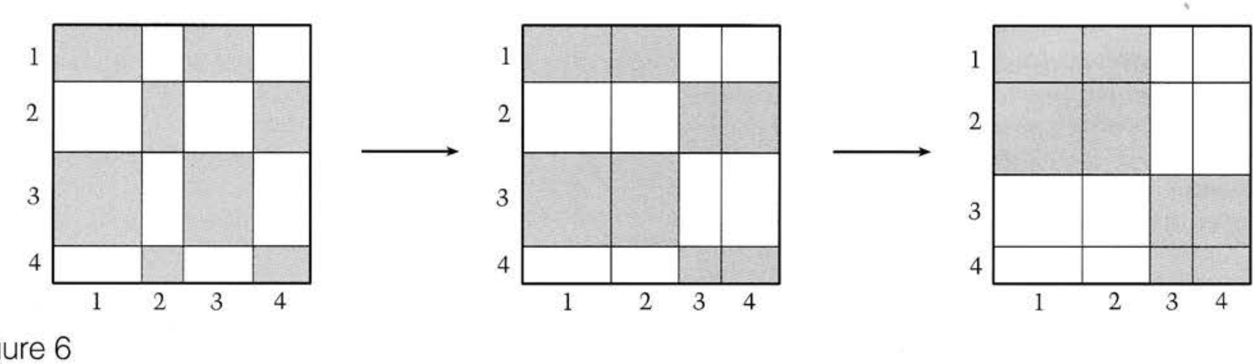


Figure 6

Solution. Assume that all pairwise distances between the participants changed after the break. Suppose they walked clockwise from their places to the new ones. Then the total distance covered would equal an integer number of complete circuits (the proof of this fact is left to the reader). On the other hand, by our assumption, all the participants covered different distances, which is possible only if one of them remained in place, the second covered 1/12 of the circuit, the third covered 2/12, and so on. The last person must have covered 11/12 of the circuit. The total distance covered is 66/12, or 5.5 full circuits. But this contradicts the fact that the total distance covered equals an integer number of complete circuits.

Problem 12. *Alternate triangles.* Two congruent equilateral triangles are arranged so that their intersection forms a hexagon. Prove that the sum of the lengths of three pairwise nonadjacent sides of the hexagon equals the sum of the lengths of its three other sides.

Solution. All the triangles Δ_1, Δ_2, ..., Δ_6 are similar (figure 7). So to prove that the sum of three pairwise nonadjacent sides of the hexagon equals the sum of its three other sides, it's sufficient to prove that the sum of the altitudes in triangles Δ_1, Δ_3, and Δ_5 drawn to the sides of the hexagon equals the sum of the corresponding altitudes in triangles Δ_2, Δ_4, and Δ_6. Notice that the sum of the areas of triangles $A_1A_2A_3$, $A_3A_4A_5$, and $A_5A_6A_1$ equals the sum of the areas of triangles $A_2A_3A_4$, $A_4A_5A_6$, and $A_6A_1A_2$ (figure 8), since each of these sums plus the area of the equilateral triangle yields the area of the hexagon $A_1A_2A_3A_4A_5A_6$. Since the bases of all six triangles are equal, we find that the sums of their altitudes are equal as well.

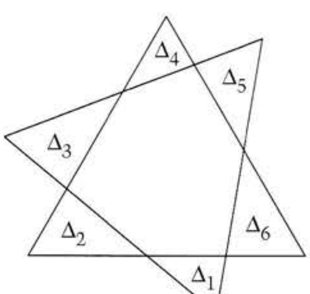


Figure 7

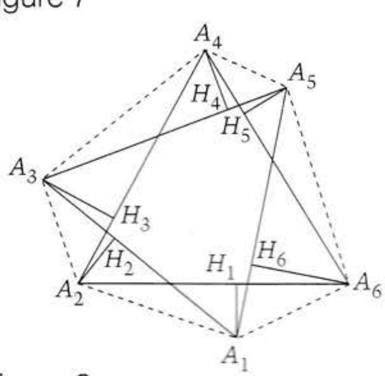


Figure 8

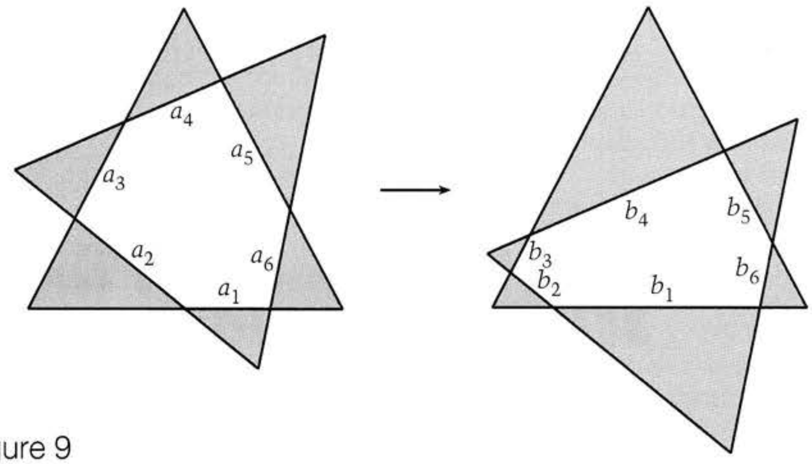


Figure 9

Problem 13. *Constant perimeter.* Two equilateral triangles (not necessarily congruent!) were arranged so that their intersection formed a hexagon. Then the triangles underwent a parallel translation, and their intersection formed another hexagon (figure 9). Prove that both hexagons have equal perimeters.

Solution. The sums of the perimeters of the six triangles situated around both hexagons are equal, since every one of them equals the sum of the perimeters of the given equilateral triangles. Thus we have $p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = q_1 + q_2 + q_3 + q_4 + q_5 + q_6$, where the numbers p_i are the perimeters of the 'old' shaded triangles, and the numbers q_i are the perimeters of the 'new' shaded triangles. All twelve bordering triangles are similar; therefore, there exists a coefficient k such that $p_i = ka_i$ and $q_i = kb_i$. We substitute these equalities in the above equation for the sums of the perimeters to obtain $a_1 + a_2 + a_3 + a_4 + a_5 + a_6 = b_1 + b_2 + b_3 + b_4 + b_5 + b_6$.

Problem 14. *Two chords.* A circle is divided into n^2 equal arcs by n^2 points. Among these points, n are painted red and the other n blue. Prove that there exist two equal chords, one of which has red endpoints and the other blue.

Solution. Let the length of the circle be n^2 units. Then the length of any arc whose ends are at the given points can take only one of the $n^2 - 1$ possible values—namely, 1, 2, ..., $n^2 - 1$. Consider all possible arcs that have differently colored endpoints and such that if we travel from the red endpoint to the blue endpoint, we are going clockwise. There are exactly n^2 such arcs. So, by the pigeon-hole principle, there are two arcs, AB and CD, of the same length. Then chord AC with red endpoints equals the chord BD with blue endpoints.

Problem 15. *Cutting a square into squares.* A square is cut into 36 smaller squares. The area of one of these small squares is distinct from 1, while the areas of all other squares equal 1. Find the area of the original square.

Solution. There must be a side of the original square that is adjacent only to squares of area 1. Therefore, the side of the original square is an integer n. So the side of the square whose area differs from 1 is an integer k. Now we have $n^2 - k^2 = 35$, or $(n + k)(n - k) = 35$. The possible values of $n + k$ and $n - k$ are the pairs of factors of 35. Since $k \neq 1$, we find that we must have $n + k = 35$, $n - k = 1$. This yields $n = 18$, and the area of the original square is $18^2 = 324$.

What happens at the boundary

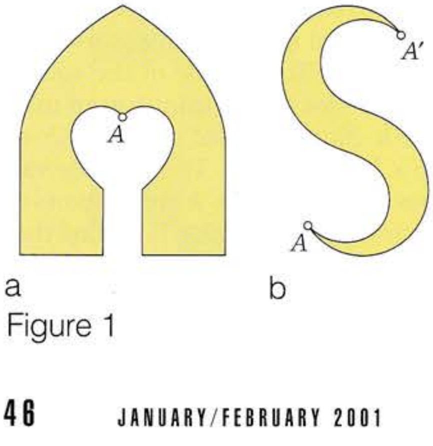
by A. Borovoy and Y. Klimov

EVENTS OCCURRING AT THE boundary of a liquid and another medium don't get a lot of attention in high school physics course. But surface tension, wetting, and capillary action underlie many interesting observations and experiments that have been known since way back when or have been discovered quite recently—some very complicated, some very simple. That's what this article is about. Let's begin with a famous demonstration that is easy to perform.

Homemade "speedboat" and diverging rings

Take a piece of paper and cut out a "boat" like the one shown in figure 1a. Place a drop of concentrated soap solution or a small piece of soap onto point A. Put the boat in a bowl of water. It should start moving forward.

Instead of the boat, you can make a spinner like the S-shaped object in figure 1b. Place soap at points A and A'. In which direction will the spinner rotate?



a b
Figure 1

The next experiment requires some practice, but your efforts will be rewarded with a beautiful demonstration. Take a container of clean, pure water and touch the surface with a fountain pen filled with black ink. It will produce a colored spot that spreads on the water's surface. Now take a rod that has been rubbed with soap, or simply a piece of soap, and touch the center of this spot—it will be disrupted and turn into a thin ring. Touch its center again with the pen (this time you might use ink of a different color) and then use the soaped rod. This cycle can be repeated many times to produce a pattern of concentric rings of different colors (figure 2). If you carefully lower a piece of blotting paper onto the water, you can transfer the image to it.


Figure 2

A practical application was found for this experiment—it is used in Japan to produce paper with an astonishing variety of beautiful patterns.

We can easily explain what's going on in these experiments if we picture the water as a stretched film. The famous English physicist Thomas Young was the first to advance this model. In December 1804, he published a paper explaining the capillary action using a stretched film to model the surface of a liquid. This model was soon confirmed by other experiments and was generally accepted.

Thus, the surface of uncontaminated, clean water is a stretched film. Water with soap or ink in it is also a film, but it is less elastic. Clean water has a stronger "desire" to shrink than soapy or tinted film, so it drags the edge of the other films, pulling the boat or inky ring toward it.

There are many ways to determine experimentally the degree of "elasticity" of surface films. This value is known as *surface tension* and is represented by the Greek letter σ. One of simplest methods is based on the process of drops forming at the end of a tube. Take a pipette with a narrow tip and fill it with water. Carefully begin to squeeze out a drop and see how it gradually grows and changes shape. Eventually the drop separates from the pipette and falls (figure 3).

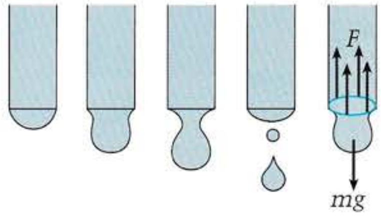


Figure 3

Note that the drop separates from the pipette only when it becomes large enough. While the drop is small, surface tension holds it at the tip of the pipette. At the moment of separation, the weight of the drop mg equals the force of the surface tension F. Or, as your physics textbook might put it,

$$mg = 2\pi r\sigma,$$

where r is the radius of the "neck" of the drop just before it separates from the pipette.

By counting, say, 100 drops and determining their total mass $M = 100m$ using a scale, we can determine the coefficient of surface tension of the test fluid:

$$\sigma = \frac{mg}{2\pi r} = \frac{Mg}{200\pi r}.$$

Unfortunately, this method is very inexact and inaccurate. First, it's not easy to measure the radius of the droplet's neck (in laboratories, the neck is optically magnified and projected onto a screen where the measurements are performed). Second, the process of separation of a liquid drop is more complicated than just the rupture of the neck. For example, a small droplet is formed simultaneously with the original large drop. Also, in a real experiment there are other sources of error.

However, if we want to obtain merely an order of magnitude for σ or to qualitatively compare the surface tension of, say, clean water and a soap solution, it's sufficient to use a pipette and assume that the radius of the separating drop's neck is equal to the inner radius of the tip of the pipette. The latter can be measured, for example, with a calibrated wire that fits snugly in the opening of the pipette.

For the sake of comparison, let's write down the values of σ for some liquids (in SI units, N/m): water at 20°C (0.073), soap and water solution (0.04), olive oil (0.033), kerosene (0.026), and mercury (it gradually decreases in air from 0.5 to 0.4).

A ripple and how to damp it

It's interesting that the forces of surface tension are responsible for so-called capillary waves (that is, ripples on water). They can be observed by using the same pipette. Fill a bowl with water, place the pipette near the surface, and release the drops into the water. The rapidly running concentric waves are actually waves of surface tension. (Generally speaking, the mechanism of wave formation and propagation on the surface of water is not that simple. The leading actors in this show are the force of gravity and surface tension. However, the model of an elastic film is adequate for explaining ripples on water.)

There's an old story that in ancient times pearl divers, who fished for pearls at the bottom of the Aegean Sea, took some olive oil in their mouths before diving. When they reached the bottom, they released the oil, which rose to the surface and eliminated the ripple, thereby producing a clear "window" that let sunlight into the depths. This made it easier for the divers to work on the floor of the sea.

Here's another interesting observation. *"In calm weather certain patterns, which are similar to moire patterns, can be clearly seen on the smooth sea surface,"* wrote the renowned Russian geophysicist V. V. Shuleikin. *"If one looks closely, it's easy to see that the ripples appear within the boundary of a brighter background, and that the bright spots and the bands observed against this background correspond to places on the sea's surface where the ripples are damped."* He and his colleagues carried out experiments in which the coefficient of surface tension was measured by the cone detachment method. These experiments showed that the coefficient of surface tension of rippling water equals that of pure water. In contrast, it was significantly smaller in the area of the bright spots. They concluded that ripples on the sea surface are damped by a film of oil or some other substance. It was no simple matter to explain this peculiar phenomenon.

Studies subsequently showed that surface films only damp small waves effectively, but they can also prevent the appearance of foam on the crests of large waves during stormy weather. This rather modest help can be of vital importance in a shipwreck, as Shuleikin pointed out: *"It is important in practice to damp only the foamy crests, which pose a great threat to the ship and even more to a lifeboat being lowered."*

What is the shape of a drop?

The answer to this seemingly simple question depends on a number of factors. If the shape were determined only by forces of surface tension, the drops would assume a shape whose surface area is minimal. Every student knows this shape—it's a sphere. However, the Earth's gravity "distorts" this ideal form, so in many cases the drops are flattened, and only very small droplets keep their spherical shape. What if we try to exclude gravity—or rather, counterbalance the force of gravity by some other force?

In 1843, the Belgian physicist Joseph Plateau conducted an experiment that was soon named after him and introduced into every school textbook. Plateau found the concentration of alcohol in a water solution that makes the density of the mixture equal to that of olive oil. He then introduced some olive oil into the solution and obtained a spherical drop. The explanation is obvious: the buoyancy counterbalances the force of gravity, while surface tension is responsible for the spherical shape.

It's interesting that in the state of weightlessness one can obtain large balls of various liquids. Such drops were observed in space in the course of experiments with electrical welding. These drops consisted of molten metal.

You can reproduce Plateau's experiment at home using castor oil and water. The density of this oil is just slightly less than that of water, so if you pour pure water in it, the slowly sinking water drops are spherical. You can stain the drops with watercolors to produce a splendid display of multicolored balls in the glass of oil.

A more difficult but even more spectacular experiment can be carried out with liquid epoxy resin and saline solution. Pour table salt into a jar filled with water. After a while the salt will diffuse upward and produce a solution in which the concentration decreases with height. Correspondingly, the upper layers of the solution will have a lower density than the lower ones. In such a solution of variable density, a drop of epoxy resin will hover at a certain height. You can use whatever method you like to color the drops.

Trace of a dried drop

A puzzling phenomenon (at least at first glance) occurs when a drop of some solution dries. This is a very simple experiment. Place several drops of not very concentrated saline solution on a glass plate and let them dry. Naturally white spots of salt will eventually appear in place of the drops. The strange thing is that the salt precipitates not uniformly, but as a set of alternating rings. The reason for this strange behavior is the intermittent, "jumpy" nature of the drying process.

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

CONTINUED FROM PAGE 39

Consider the expression

$$\frac{a_1+\ldots+a_n}{n}-\sqrt[n]{a_1\cdot\ldots\cdot a_n}$$

as a function of x:

$$f(x)=\frac{a_1+\ldots+a_m(n-m+1)+(n-m)x}{n}$$
$$-\sqrt[n]{a_1\cdot\ldots\cdot a_m\left(a_m+x\right)^{n-m}}.$$

Its derivative is easy to calculate:

$$f'(x)=\frac{n-m}{n}-\frac{n-m}{n}\sqrt[n]{\frac{a_n\cdot\ldots\cdot a_m}{\left(a_m+x\right)^m}}.$$

Clearly $f'(x) > 0$ for $x > 0$. Indeed, the radicand is less than 1:

$$\frac{a_1a_1\ldots a_m}{\left(a_m+x\right)^m}$$
$$=\frac{a_1}{a_m+x}\cdot\frac{a_2}{a_m+x}\cdot\ldots\cdot\frac{a_m}{a_m+x}<1.$$

Thus the Cauchy inequality is proved.

Two other proofs of this inequality can be found in the articles "Cauchy and Induction" and "The Sum of Minima and the Minima of Sums" in this issue of *Quantum*. Compare the advantages and disadvantages of these proofs.

Index of Advertisers

PocketScope	48

ANSWERS, HINTS & SOLUTIONS

Physics

P311

Let's consider one of the threads and choose a small segment of length l (figure 1). Denote the tension of the elastic thread by $\mathbf{T}$. The resultant force $\mathbf{Q}$ acting on the segment has a magnitude $2T \sin(\alpha/2)$. If $l \ll R$, where R is the radius of the thread's curvature along the chosen segment, then $\alpha = l/R \ll 1$, so the above expressions can be simplified:

$$\sin(\alpha/2) = \alpha/2, \text{ and } Q = (l/R)T.$$

Since our segment is in a state of equilibrium, the sum of all the forces acting on the segment is zero. This means that the ball acts on the thread with a force $\mathbf{F} = -\mathbf{Q}$. According to Newton's third law, the thread acts on the ball with an equal force directed to the center of the ball.

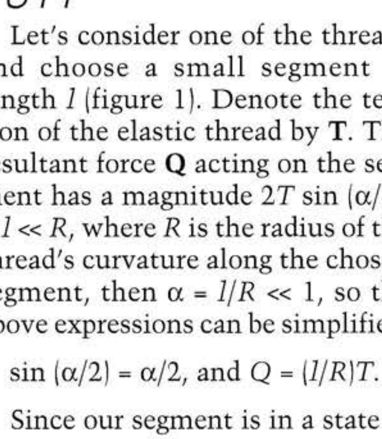


Figure 1

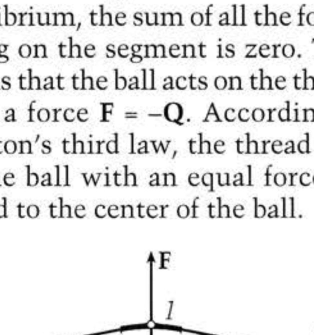


Figure 2

Any circle on the ball parallel to the "equator" is crossed by the same number of threads ("meridians") N. The circumference of such a circle with radius $r = R\cos\phi$ (where ϕ is the angle formed by the ball's radius drawn to the circle and the equatorial plane) is $2\pi r = 2\pi R\cos\phi$. Therefore, a segment of the spherical envelope with area $l \times l$ is covered by the following number of the threads:

$$n = \frac{N}{2\pi R\cos\phi} l.$$

Thus the pressure developed by the rubber net and applied to the envelope is

$$P = n\frac{Q}{l^2} = \frac{NT}{2\pi R^2 \cos\phi}.$$

If the ball's envelope retained its spherical shape, the pressure would increase with the angle ϕ. However, according to Pascal's law, the pressure is the same at any point within the ball. Thus the envelope cannot be spherical: Its equatorial diameter is greater than its vertical diameter. The radius of curvature of the meridians decreases toward the equator and increases toward the poles in such a way that $R^2\cos\phi = \text{const}$. As a result, the ball looks like a pumpkin (figure 2). Perhaps the shape of a real pumpkin is determined by a similar phenomenon—structural "support cords" aligned with the meridians?

P312

Due to the effect of the mechanical (sonic) oscillations, the stream emerging from the capillary tube will be slightly crimped. The stream would also be crimped even without the piezoelectric crystal driver, because its surface will be disturbed by the edge of the capillary tube. Of course, the capillary edge produces a stronger disturbance when it oscillates due to the vibrating piezoelectric crystal. Therefore, the stream is disturbed earlier than it would have been in the absence of the piezoelectric element. In any case, the amplitude of the crimping is small compared to the radius of the stream, so we will assume that the stream is approximately cylindrical and that its radius is equal to the internal radius of the capillary.

It's natural to assume that the volume of a drop is determined by the amount of water flowing from the capillary during one period of the sound oscillation:

$$\frac{\pi d^2}{4} vT = \frac{4}{3}\pi R^3.$$

Taking into consideration that $T = 1/f$, we get

$$R^3 = \frac{3}{16}\frac{vd^2}{f},$$

$$R = \sqrt[3]{\frac{3}{16}\frac{vd^2}{f}} = 0.25 \text{ mm}.$$

P313

A graphical solution to this problem is the simplest route. The total pressure $P(T)$ in the cylinder is the sum of the saturated vapor pressure $P_s(T)$ and the hydrogen pressure $P_H(T)$. According to the ideal gas equation,

$$\begin{aligned} P_H(T) &= \frac{m_H}{\mu_H V} RT \\ &= \frac{(2\cdot 10^{-3}\ \text{kg})(8.31\ \text{J/K}\cdot\text{mol})}{(2\cdot 10^{-3}\ \text{kg/mol})(2\cdot 10^{-3}\ \text{m}^3)} T \\ &= (4.15\cdot 10^3\ \text{Pa/K})T. \end{aligned}$$

Let's calculate $P_H(T)$ at two temperatures. For example, $P_H \cong 15.5 \cdot 10^5$ Pa at $T = 373$ K and $P_H \cong 18.8 \cdot 10^5$ Pa at $T = 453$ K. Now we plot the

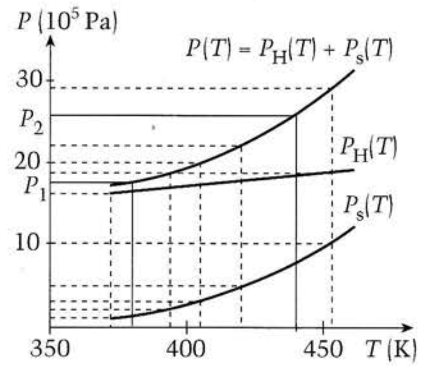


Figure 3

graph $P_H(T)$ versus T as shown in figure 3. Using the table for saturated vapor pressure, we plot the function $P_s(T)$. Adding the functions $P_H(T)$ and $P_s(T)$ gives us the dependence of pressure in the cylinder on temperature $P(T)$. From this plot we obtain the initial and final temperatures in the cylinder: $P_1 = 17 \cdot 10^5$ Pa yields $T_1 \cong 380$ K, while $P_2 = 26 \cdot 10^5$ corresponds to $T_2 \cong 440$ K.

Let's find the mass of the evaporated water. Assuming the vapor to be an ideal gas, we obtain the values for the initial vapor pressure P_{s1} and final vapor pressure P_{s2} from the curves in figure 3. At $T_1 = 380$ K, the hydrogen pressure is $P_{H1} \cong 15.5 \cdot 10^5$ Pa, so

$$P_{s1} = P_1 - P_{H1} \cong 1.5 \cdot 10^5 \text{ Pa}.$$

Similarly, at $T_2 = 440$ K, $P_{H2} \cong 18 \cdot 10^5$ Pa, so

$$P_{s2} = P_2 - P_{H2} \cong 8 \cdot 10^5 \text{ Pa}.$$

Using the ideal gas law for the water vapor at (P_{s1}, T_1) and (P_{s2}, T_2), we have

$$P_{s1}V = \frac{m_{v1}}{\mu_v} RT_1 \text{ and } P_{s2}V = \frac{m_{v2}}{\mu_v} RT_2,$$

where m_{v1} and m_{v2} are the initial and final masses of vapor in the cylinder. From these equations we get the mass of the evaporated water:

$$\Delta m_v = m_{v2} - m_{v1} = \frac{\mu_v V}{R}\left(\frac{P_{s2}}{T_2} - \frac{P_{s1}}{T_1}\right)$$

$$= \frac{\left(18 \cdot 10^{-3} \text{ kg/mol}\right)\left(2 \cdot 10^{-3} \text{ m}^3\right)}{8.31 \text{ J/K} \cdot \text{mol}}$$

$$\cdot\left(\frac{8}{440} - \frac{1.5}{380}\right) \cdot 10^5 \ \frac{\text{Pa}}{\text{K}} \approx 6 \cdot 10^{-3} \text{kg}.$$

P314

Since the coils are connected in parallel, the voltages across them are identical: $V_{L_1} = V_{L_2}$. However,

$$V_{L_1} = L_1 \frac{\Delta I_1}{\Delta t} \text{ and } V_{L_2} = L_2 \frac{\Delta I_2}{\Delta t}.$$

Thus

$$L_1 \frac{\Delta I_1}{\Delta t} = L_2 \frac{\Delta I_2}{\Delta t}. \qquad (1)$$

Initially, there were no currents in both coils, so it follows from equation (1) that at any instant in time

$$L_1 I_1 = L_2 I_2. \qquad (2)$$

This means that both currents reach maximum values simultaneously.

Evidently, I_1 and I_2 are greatest when the capacitor is discharged. At this time energy conservation requires that

$$\frac{CV^2}{2} = \frac{L_1 I_1^2}{2} + \frac{L_2 I_2^2}{2}. \qquad (3)$$

The system of equations (2) and (3) yields

$$I_1 = V\sqrt{\frac{L_2 C}{L_1(L_1 + L_2)}},$$

$$I_2 = V\sqrt{\frac{L_1 C}{L_2(L_1 + L_2)}}.$$

P315

Rays making a single reflection from the cone travel as if they were emitted by a continuum of imaginary point sources located on a circle. Each of these imaginary sources is symmetrical to the point source P relative to the corresponding generatrix of the cone. The image of this circular source formed on the screen is a ring (figure 4). It's important that the bundle of rays traveling from each point source to the lens is planar: It doesn't pass through the entire surface of the lens, but rather along the corresponding diameter (for example, the rays emitted by point source P' travel in the vertical plane). Thus the attenuation of such a ray depends on the shape and orientation of the diaphragm.

Clearly the symmetrical diaphragm shown in figure 4a (page 11) will attenuate the rays emitted by all the imaginary sources equally. In this case, the brightness of the ring on the screen will be decreased uniformly.

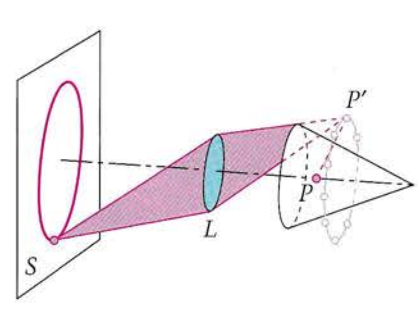


Figure 4

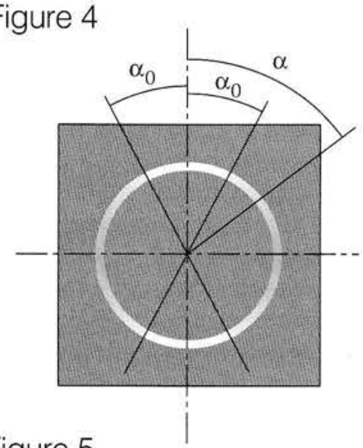


Figure 5

The diaphragm shown in figure 4b (page 11) will transmit without attenuation the rays whose planes form angles $\alpha < \alpha_0$ with the vertical plane. Thus the brightness of the corresponding upper and lower parts of the ring on the screen will remain the same. In contrast, the other rays will be partially blocked by the diaphragm, and attenuation of the rays traveling closer to the horizontal plane will be stronger. Therefore, the brightness of the side areas of the ring will decrease as the angle α increases from $\alpha = \alpha_0$ to $\alpha = \pi/2$. The image will look like that in figure 5. It's clear that the image is symmetrical relative to the vertical and horizontal axes.

Math

M311

Let a be the first term of the progression and d be its common difference. We can assume that $d > 0$. Then all terms of the form $a + 10^n d$, where n is a sufficiently large number (such that $10^n > a$), have the same sum of digits equal to the sum of digits of the numbers a and d.

A similar argument will show the proposition to be true not just for decimal notation, but no matter what base of notation we use for writing the terms of the progression.

M312

Points M and N are the feet of the altitudes of triangle ABC drawn from vertices B and A, respectively. Therefore, the third altitude of this triangle passes through their intersection point H. In addition, points C, M, N, and H lie on a circle δ with the diameter CH, since $\angle CMH = \angle CNH = 90°$. Let P be the center of this circle. Notice that as the diameter AB rotates, the measure of angle C remains the same. Indeed, it is equal to half the difference of the arcs MN and AB, which are constant in measure (see figure 6). Since the chord MN does not change, the circle δ (along which point C and the diametrically opposite point H move) and its center P also remain the same. The diameter CH, which is a part of the altitude under consideration, just rotates about point P.

M313

First, consider the extreme case when the given numbers take two different values: k of them are equal to $a < 0$, and the remaining $(n - k)$ are equal to $b > 0$. Then $ka + (n - k)b = 0$. Thus $ka = -(n - k)b$ and $ka^2 + (n - k)b^2 = -(n - k)ab - kab = 1$, from which we get $ab = -1/n$. Choosing $x_i = a$, $x_j = b$ satisfies the requirements of the problem.

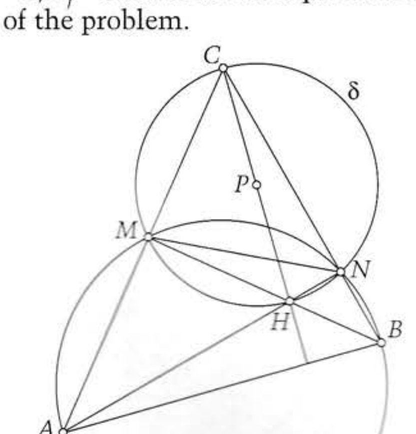


Figure 6

We try to reduce the general case to this one by "moving the variables apart." Notice that if we replace the pair of numbers u, v $(u < v)$ by $u - t$ and $v + t$ $(t > 0)$, their sum does not change and the sum of squares increases:

$$(u - t)^2 + (v + t)^2 = u^2 + v^2 + 2t(v - u) + t^2.$$

Let $a < 0$ be the minimum and $b > 0$ be the maximum of the given numbers $x_1, x_2, \ldots, x_n$. If this set contains two numbers different from a and b, move them apart without changing their sum so as to make one of them equal to one of the extreme numbers a or b. A little thought will show that after several such operations, we obtain a set of n numbers in which all but one number (call it c) equals a.

We have now replaced the original set $\{x_i\}$ with a new set of numbers, mostly equal to a or b. Assume that k of these new numbers are equal to a, and $m = n - k - 1$ of them are equal to b. The sum of these numbers is 0, and the sum of their squares is not less than 1:

$$ka + c + mb = 0,$$
$$ka^2 + c^2 + mb^2 = -(c + mb)a + c^2 - (ka + c)b \geq 1.$$

Since $a \leq c < b$, we have $(c - a)(c - b) = ca - c^2 + cb - ab > 0$.

Adding this inequality to the preceding one, we obtain the desired inequality

$$-(m + k + 1)ab = -nab > 1$$

or

$$ab < -1/n.$$

So choosing a and b again satisfies the requirements of our problem. This proof also shows that equality is achieved only for the extreme cases we started with.

This problem also has a direct, if slightly artificial, algebraic solution. For any i $(1 \leq i \leq n)$, we have

$$(x_i - a)(x_i - b) \leq 0.$$

Adding up all these inequalities, we obtain

$$\sum_{i\leq 1}^{n} x_i^2 - (a+b)\sum_{i=1}^{n} x_i + nab \leq 0. \quad (1)$$

But

$$\sum_{i\leq 1}^{n} x_i^2 = 1 \text{ and } \sum_{i=1}^{n} x_i = 0;$$

thus, inequality (1) turns into

$$nab \leq -1.$$

M314

This problem has many different solutions based on the following simple inequalities for two positive numbers:

$$x^2 + y^2 \geq 2xy,$$
$$\frac{x+y}{2} \geq \sqrt{xy} \geq \frac{2xy}{x+y}.$$

For proofs of these basic inequalities, see for example Beckenbach and Bellman, *An Introduction to Inequaliites*, Washington, DC: MAA, New Mathematical Library, 1961.

We present four such solutions; the reader can choose the most attractive one from his or her point of view.

1. The condition $ab > a + b$ can be rewritten as

$$(a - 1)(b - 1) > 1.$$

The expressions in both parentheses must be positive (indeed, if $0 < a < 1$ and $0 < b < 1$, then $(a - 1)(b - 1) < 1$). Then, by the inequality between the arithmetic and geometric means for the numbers $a - 1$ and $b - 1$, we have

$$a-1+b-1 \geq 2\sqrt{(a-1)(b-1)} > 2,$$

from which we get $a + b > 4$.

2. The given condition is equivalent to the condition that the harmonic mean of a and b is greater than 2:

$$\left(\frac{a^{-1}+b^{-1}}{2}\right)^{-1} = \frac{2ab}{a+b} > 2.$$

But the arithmetic mean is greater than or equal to the harmonic mean:

$$\frac{a+b}{2} \geq \frac{2ab}{a+b},$$

since $(a - b)^2 \geq 0$. Therefore,

$$a + b > 4.$$

3. Dividing the given condition by a and b, we obtain

$$a > \frac{a}{b} + 1, b > \frac{b}{a} + 1,$$

from which we get

$$a + b > \frac{a}{b} + \frac{b}{a} + 2 \geq 4.$$

4. We can add one more inequality on the left of the given inequality:

$$\frac{(a+b)^2}{4} \geq ab > a + b.$$

Then, for $S = a + b$ we have $S^2/4 > S$. Thus, $S^2 > 4S$, which (since $S > 0$) yields $S > 4$.

M315

Using the fact that two tangents drawn to any circle from the same point are equal, we denote the various line segments as shown in figure 7. (We also use the fact that the tangents drawn from point D to each of the two congruent circles form congruent angles; thus, all these tangents are equal.) Equating the two tangents drawn from point B to the farther circle and taking into account that $AB = BC = x + y$, we obtain

$$y + 2z = y + x + u,$$

from which we obtain

$$x + u = 2z. \qquad (1)$$

Now the solution can be completed by standard manipulations involving the formula that expresses the area of the triangle in terms of its semiperimeter s: $K = -rs$, and in terms of the radius r_a of the escribed circle tangent to side a: $K = r_a(s - a)$. If h denotes the altitude drawn to a leg of triangle ABC and r denotes the radius of the equal circles, then the area S_{ABC} is equal to

$$h \cdot \frac{x+y}{2} = S_{ABC} = S_{ABD} + S_{DBC}$$

$$= r(x + y + z) + r\left(y + z + \frac{x+u}{2} - z - u\right)$$

$$= r\left(x + y + \frac{x+u}{2} + y + \frac{x+u}{2} - u\right)$$

$$= r(2x + 2y).$$

It follows that $r = h/4$.

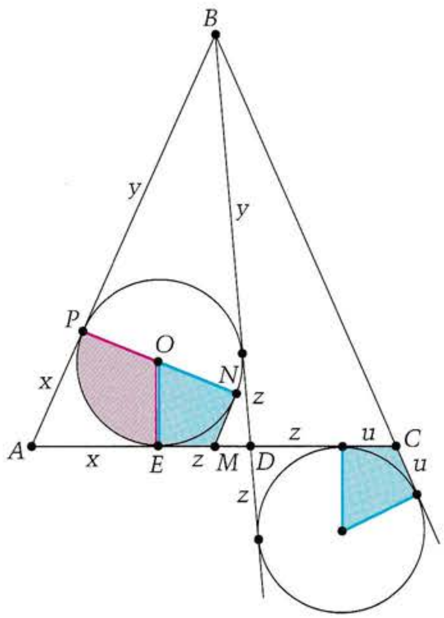


Figure 7

This result can also can be obtained geometrically, using equation (1). The blue quadrilateral with vertex C (see figure 7) can be flipped over and placed so that C coincides with E, and so that the sides formed by radii of the two circles also coincide. Then $EM = u$, and $AM = x + u$. But $MC = (z - u) + z + u = 2z = x + u$, so M is in fact the midpoint of AC. By construction, $\angle EMN$ is supplementary to $\angle BAC$, so MN is parallel to AB. Since MN is also perpendicular to radius ON (again by construction), MN is tangent to the circle, and it is not hard to see that N, O and P are collinear. Now it is wellknown that MN, the line through a midpoint of triangle ABC and parallel to side AB, bisects any line from C to AB. So it bisects the altitude from C to AB, and PZ is equal to half this altitude. It follows that $r = h/4$, as before.

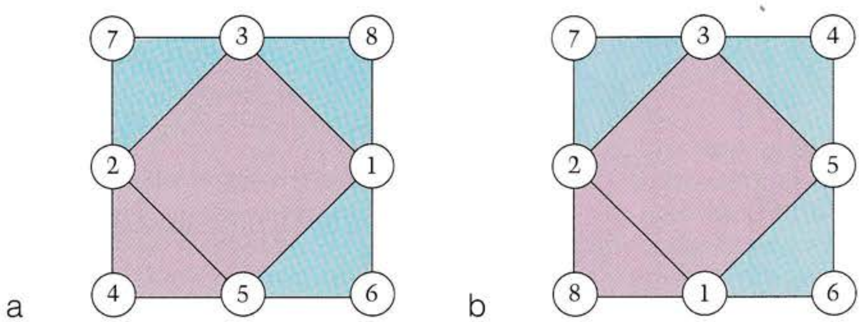


Figure 8

Brainteasers

B311

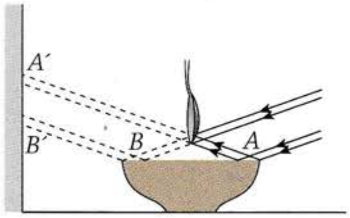
Let us suppose that the suitors didn't do too well with this problem, and were all spurned. We line them up, and ask them to give back their plums, in the reverse order in which they took them out. The third suitor must first give back his three "extra" plums. Then there are three plums in the basket, which must have been half of what remained when the second suitor was done. The third suitor had been given the other half (three more plums), and so now must put them back as well.

Now there are six plums in the basket, and the second suitor must first give back his one "extra" plum. This puts seven plums in the basket, which is half of what the second suitor must have seen at the beginning of his turn. He must now put back the other half, which is also seven plums.

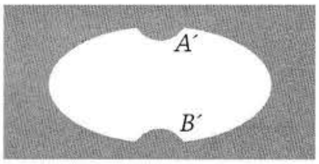
Now there are 14 plums in the basket. The first suitor must follow his colleagues, by putting his "extra" plum back, so there are 15 plums in the basket. Then he puts back an equal number, and there are 30 plums. This must have been the state of the basket at the beginning of the story.

B312

The first condition implies that Aramis placed fourth in the competition. From the second condition it follows that Porthos was second, and from the final condition it follows that D'Artagnan was first and Athos was third.



a



b

Figure 9

B313

Two solutions are possible (see figure 8).

B314

Note that $D \leq 3$, and D is even. Thus, $D = 2$. Then R is either 3 or 8. But R cannot be eight, since in this case $4I$ would have ended in the digit 5, which is impossible. Thus $R = 3$. Now we have $I = 8$, which then yields $B = 6$ and $\ddot{O} = 7$.

B315

The paths of the sunbeams are shown in figure 9a. The spoon blocks some rays falling on one side of the surface and some rays reflected from the opposite side of the surface. So two shadows will simultaneously appear on the light spot projected on the wall—one on the upper edge and another on the bottom edge (figure 9b). As the spoon comes nearer to the surface, the two shadows will grow and eventually connect.

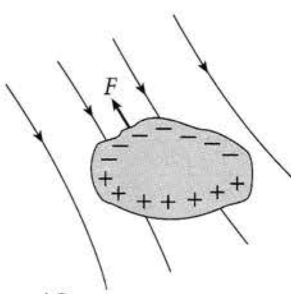


Figure 10

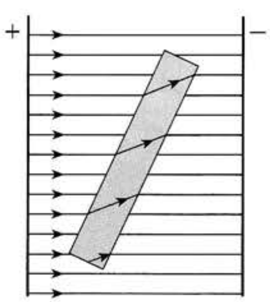


Figure 11

Kaleidoscope

1. Polarized charges appear on the piece of paper, which is a dielectric. The electric field is stronger near the comb, so the attraction toward the comb is stronger than the repulsion from it (figure 10). In contrast, equal and opposite forces act on a polarized dielectric placed in the homogeneous field of a parallel-plate capacitor.

2. In contrast to the free electrons in metals, the charges in a dielectric are bound.

3. The intensity of the fields in the air and the dielectric are equal. Attenuation of the field inside the dielectric due to its polarization is compensated by an increase in charge density in the lower part of the plates.

4. The film of dielectric oxide formed at the surface of the plate is very thin.

5. The charges will not change.

6. Both the direction and the density of the lines of force will change inside the bar due to polarization of the dielectric (figure 11).

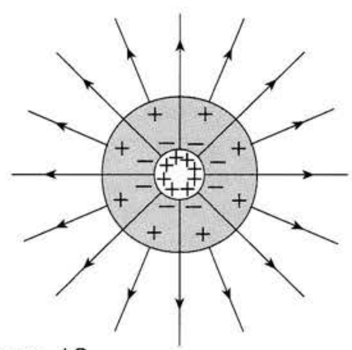


Figure 12

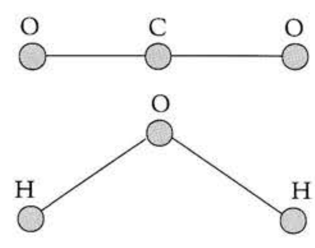


Figure 13

7. See figure 12. When the boundary of the dielectric is crossed, the number of lines of force changes abruptly due to the presence of polarized charges.

8. When the metal balls are immersed in the dielectric, the voltage difference and, consequently, the intensity of the electric field between them do not change. This is achieved by increasing the charge of the balls by a factor of ε, where ε is the relative dielectric constant of the liquid. Therefore, the attractive force between the balls will increase by a factor of ε.

9. The charge on each plate will increase by a factor of ε (see, for example, the solution to the previous problem). The intensity of the field generated in the air by each plate will increase by the same factor. Therefore, the attractive force between the plates will increase by a factor of ε^2.

10. The higher voltage requires a thicker layer of dielectric, which decreases its capacitance. To have the same capacity, it is necessary to increase the area of the plates. Both reasons lead to an increase in the capacitor's size.

11. $\varepsilon = 2$.

12. The energy will decrease by a factor of ε.

13. The dipole moments in a dielectric are oriented in the electric field with a certain time lag.

14. Metals reflects electromagnetic waves better than dielectrics.

15. Yes, it will, because both the atomic nucleus and the electronic shell change their mutual location in the field generated by an electron. Mutually repulsing electrons are shifted in the direction opposite to the electric field, while the attracted nucleus is displaced in the direction of the external field. As a result, the atom acquires a dipole moment.

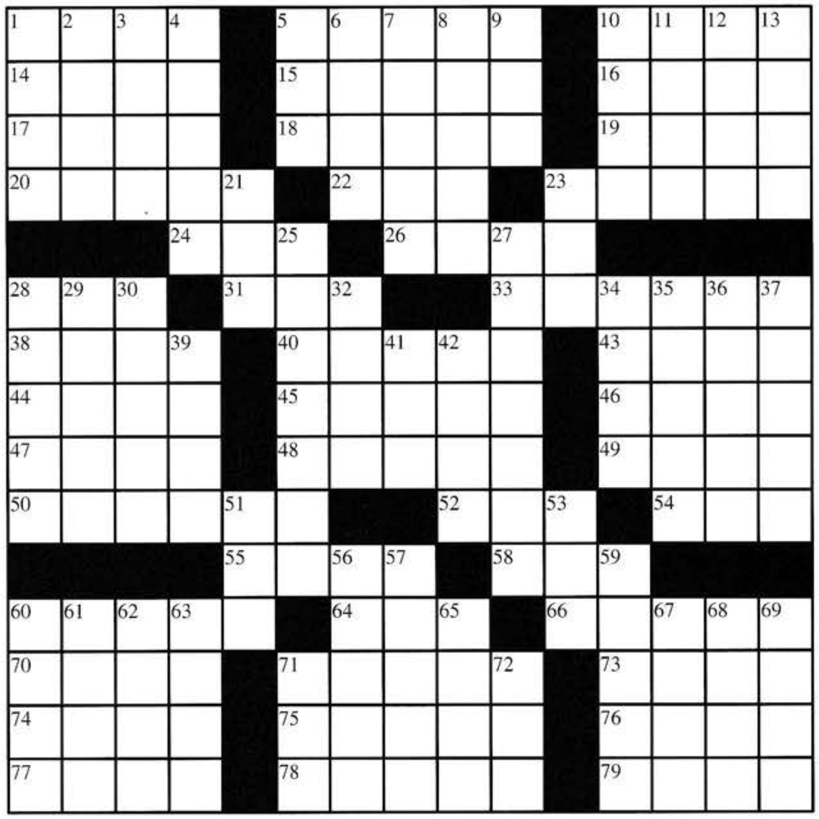
16. See figure 13.

Microexperiment

The electric field generated near the terminals of a battery is too weak to produce any discernible polarization of dielectrics.

Criss × cross science

by David R. Martin



ACROSS

1 Average
5 Crystal reflection physicist
10 Vaporize
14 Of aircraft
15 Ancient Greek colony
16 ___ de choc
17 ___ shift (of H spectrum)
18 Add
19 Close attention
20 Unit of magnetic flux density
22 Common logarithm base
23 Units of mass
24 10^{18}: pref.
26 ___ function
28 Trig. function
31 Guido's high note
33 Monochromatic light sources
38 An asteroid
40 Filth
43 Thermal energy
44 Togo capital
45 Insect: comb. form
46 TV's Alan
47 An asteroid
48 ___'s law (PV=constant)
49 Low dam
50 Type of shower
52 Part of SASE: abbr.
54 PC environment
55 ___-life
58 Trig. function
60 Quark property
64 Energy unit
66 Plane detector
70 Chaplin's widow
71 Moslem ruler
73 Prefix for 22 across
74 Skidded
75 Struck
76 Bakery worker
77 1936 Physics Nobelist
78 Aqueduct
79 Astronaut Sally

DOWN

1 Brewer grain
2 61,102 (in base 16)
3 Supplies weapons
4 Like neon
5 Binary digit
6 Square ___
7 Poker stakes
8 Red ___ star
9 Liquid meas.
10 Danish physicist Niels ___ (1885-1962)
11 Draft class
12 Footnote word
13 Semiconductor devices: abbr.
21 Cutting tool
23 Student's concern
25 Math course
27 Periodic table entry
28 Glashow-Weinberg-___ model
29 Be theatrical
30 Halley's ___
32 Italian river
34 George Bernard ___
35 Fished
36 Type of galaxy
37 Antares and Vega
39 ___ Descartes
41 Quality: suff.
42 Chemical measure
51 Unit of resistance
53 Reactive power unit
56 Math fact
57 Refrigerant
59 Zenith's opposite
60 Hyperbolic function
61 Vacant electron site
62 Southern blackbirds
63 Absorbed dose units
65 Elliott's ___ *Life*
67 One tenth: pref.
68 Scored perfectly
69 ___ earth (scarce metal oxide)
71 Keyboard letters
72 Reluctance unit

SOLUTION IN THE NEXT ISSUE

SOLUTION TO THE NOVEMBER/DECEMBER PUZZLE

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

IOI 2000

by Don Piele

YOUNG CHILDREN IN THE WESTERN WORLD are sometimes told by their parents, "If you dig long and deep enough you can dig a hole to China." I never made it to China that way, but I finally did make it to Beijing by plane on the occasion of the 12th International Olympiad in Informatics, the first IOI to be held on the continent of Asia. The weeklong event, September 23–30, 2000, was packed with excursions, entertainment, competitions, friendship, awards and, of course, abundant Chinese food.

We arrived in Beijing with the Canadian team after a 12-hour flight from Vancouver, British Columbia. Our delegation had gathered in Vancouver from all parts of the United States. Team leader Rob Kolstad from Colorado Springs, coach Hal Burch from Pittsburgh, coach Greg Galperin from Boston, and myself, USACO director, from Wisconsin. Team members Percy Liang and John Danaher interrupted their freshman year at the Massachusetts Institute of Technology for the trip. Greg Price from Thomas Jefferson HS of Science and Technology in Alexandria, Virginia, and Reid Barton from Arlington, Massachusetts, rounded out the USA team of four.

IOI guides

Shortly after arriving in Beijing, we walked through the sliding glass doors into the welcoming signs of our Chinese IOI guides. In their red and black vests displaying the IOI 2000 logo, they were easy to spot. Their first job was to greet us at the airport and get us safely transported by bus to our hotels. The team members were dropped off at the CATIC Grand Hotel and the delegation leaders went to the four star Continental Grand Hotel. We had arrived at 3 pm in the afternoon a bit tired from the long flight. It was very tempting to want to lie down for a short nap, but we all resisted knowing it is best to suffer the sleep loss early in order to adjust quickly to the 12-hour jet lag.

Summer Palace excursion

We awoke the next day ready for our first excursion. As is customary at IOI events, several days of excursions had been planned for the week. Our first trip was to the Summer Palace, a huge park in the northern part of Beijing that was the summer home for royalty for nearly 800 years. It had been restored after being destroyed in 1900 during the Boxer Rebellion. Our tour guide led us through this crowded park, which is both a favorite tourist attraction and a popular retreat for Beijingers. The huge Kunming Lake within the park was built completely by manual labor. A popular activity is to take a dragon boat ride on the lake. Another attraction is the world famous 728-meter-long corridor that is decorated with 8,000 paintings.

Opening ceremonies

The opening ceremonies were held in the Beijing International Convention Hall. Mrs. Chen Zhili from the Ministry of Education, delivered the opening address in Mandarin Chinese followed by an English translation. For entertainment we were treated to a number of outstanding performances featuring acrobats, musicians, and a troop from the Beijing Opera. Afterwards, all 512 guests including 276 contestants from 75 countries filed into a reception hall to try their hand with chopsticks as they sampled dished from an elaborate Chinese buffet. Zide Du, chairman of the organizing committee from China, was happy to see that his long hours of hard work had finally paid off and IOI 2000 was officially underway. His daughter was on hand to join in the festivities.

General assembly

That evening the delegation leaders met to choose the problems for the first competition day. The problems were presented by the scientific committee and accepted unanimously on the first vote. This happens so infrequently that the general assembly gave the scientific committee a round of applause. It is not easy to get approval of the first try from all countries.

Now the only thing left for many countries was translating the problems into their native language. Our delegation was fortunate and skipped this time consuming task since the official language of the IOI is English and all contestants receive a copy in English.

First competition

Early the next morning, the contestants began the first of two five-hour competitions. This gave the delegation leaders a break, and many headed back to their room for some sleep after an all-night translation session. After the contestants had finished programming their solutions to the three problems, it was time to begin the grading process. Using an automated system, the work of grading the contestants' programs was dramatically reduced. Differences in program performance were detected by running a series of test cases against each program and checking for speed and accuracy. After all the programs were tested, the results were made available to each contestant, along with the test cases. This gave the contestants the opportunity to double-check the grading process using their own computers. Our team was pleased with the first day's results.

Second excursion

Between competition days we were treated to an excursion to the heart of Beijing, beginning with the Forbidden City. This was the seat of imperial power during the Ming and Qing dynasties (1368–1911). It has been written, "Without seeing the magnificence of the royal palace, one can never sense the dignity of the emperor." We were given lots of opportunity to wander around the beautiful courtyards and inside the buildings. It reminded me of scenes from the motion picture The Last Emperor.

Our next stop was Tienanman Square, which was decorated with imported flowers for the October 1 celebration of the People's Revolution in China. Our team got in the spirit of the occasion by waving the red flag and Mao's Little Red Book, which were always available from obliging vendors. Kim Schrijvers, the team leader from the Netherlands, was able to quickly round up over 100 unsuspecting IOI participants for his "spontaneous" group photo on the square.

Second competition day

The second day was pretty much a carbon copy of the first competition day with the exception that the problems presented were a bit harder and arriving at an agreement on them took a bit longer. All in all it appeared that the creation of the International Scientific Committee had been a good idea since their work was very helpful to the Chinese Scientific Committee in selecting and testing out the competition problems. The final distribution of scores in the competition also confirmed that the problems were at the proper level of difficulty for a good distribution of scores. Now the guessing game began as the delegations wondered if their scores were high enough to get bronze, silver, or gold medals.

Following the last day of competition, we headed out in buses to the Chinese acrobatic show. This was a highly entertaining show featuring young children who could balance on just about anything and make it look easy. One boy was able to do a one-handed handstand on the head of another boy who was standing up. Another young girl constructed a tower of furniture supported below by a companion and managed to balance on the very top of a stack at least 30 feet high.

The Great Wall

The highlight of this day was a trip to the Great Wall of China, one of the great wonders of the world. This 2,000 year old structure, stretching for 4,500 miles stands today as a symbol of Chinese ingenuity and willpower. At one time every fourth person in China worked on the wall. It's a steep climb in some places to even walk on the wall since it goes straight up a mountain instead of taking the more gentle contour. After we had walked and climbed about as far as our tired legs could carry us, we returned to an outpost tower on the wall for a fully catered banquet. Chairs and tables had been hand carried onto the wall along with all of the dishes, glasses, and food for this spectacular occasion. As we sat together eating our meal, watching the sun set in the west and the lights come on illuminating the Great Wall, it was hard to believe this was really happening. A few of those at our table looked at each other and shook our heads, "This is truly unforgettable."

Closing ceremonies

Based on everything that I had already seen this week, and everything I had been told about Chinese ceremonies, I was expecting a grand finale. I was not disappointed. It began with video highlights of the

Left to right: John Danaher, Percy Liang, Gregory Price, Reid Barton week's activities, featuring scenes projected on large overhead screens within the convention hall. Official dignitaries from China occupied a special position in the first row. After a series of elaborate stage performances and speeches, it was time for the medals to be awarded. As is customary, only half of the participating students were awarded medals. Sixty-nine bronze medals were handed out individually to the winners. Gregory Price from our team won the bronze medal. Forty-seven silver medals were awarded, and two of them went to team members Percy Liang and John Danaher. The coveted gold medal was reserved for the top twenty-three participants, and Reid Barton got one of them. This was the second gold medal this year for Reid at an International Olympiad. He was awarded a gold medal at the Mathematics Olympiad held in Korea earlier in July.

Special recognition went to Jing Xu of China for being the best female contestant at the Olympiad. Only six of the 276 participants were women. A perfect score was recorded by one contestant, Mikhail Baoutine, from the Russian Federation. He was awarded the winner's trophy, a gold medal, and a laptop computer. All three of his teammates also won gold medals, making this the first time in IOI history that one country has won four gold medals.

Final banquet

That evening we were bused to the Red Rooster Theater and Restaurant in downtown Beijing where our party of over 500 was served a marvelous Chinese meal by waiters and waitresses on roller skates. It always amazed me how quickly every sit-down dinner in China was served, no matter the size of the party. The number of good service people in restaurants and shopping centers was one of the many things that astounded me about China. And the tall, slender, dancing and singing models decked out in alluring attire who entertained us that night was another surprise. It was the Beijing version of a Las Vegas chorus line.

New environments

Starting in 2001, the computing environment will include LINUX with the GNU C/C++ and the Free Pascal compilers. This will allow for more interesting problems and really speed up the grading process. This was adopted by the General Assembly on the recommendation made by our head coach, Rob Kolstad, in a presentation to the group.

Palindrome problem

Of the six problems used at IOI 2000, one of the easiest ones to solve was Palindrome. A palindrome is a symmetrical string that reads identically from left to right as from right to left. The problem was to write a program which, given a string, determines the minimal number of characters that needs to be inserted into the string to make a palindrome.

As an example, by inserting two characters, the string "*Ab3bd*" can be transformed into the palindromes "*dAb3bAd*" or "*Adb3bdA*." However, inserting fewer than two characters does not produce a palindrome.

INPUT

The input file name is PALIN.IN. The first line contains one integer, the length of the input string N, $3 \leq N \leq 5000$. The second line contains one string of length N.

OUTPUT

The output file name is PALIN.OUT. The first line contains one integer, which is the desired minimal number.

SAMPLE INPUT
5
Ab3bd

SAMPLE OUTPUT
2

Memorable IOI

After such an elaborate IOI in China, our delegation agreed that we had underestimated the amount of work that it takes to put on an event of this magnitude. This is a concern, because we are hosting the IOI in 2003. Hopefully, memories will have faded by then and our guests will forgive us for not having a branch of the Great Wall anywhere near Chicago. For now, the torch has been passed to Finland, who will host the 2001 IOI in Tampere. Zide Du and the entire Chinese organization did a wonderful job making the first IOI of the 21st century such a memorable one.

Photos

Being a member of the International Committee of the IOI has afforded me the opportunity to visit area of the world I have never been. As a result I often take many digital photos. A complete set of 240 photos in China can be viewed at www.zing.com. Search under Albums for IOI 2000.

Our sponsor

All expenses for training and travel for the USA team was paid for by our sponsor, USENIX. USENIX is the Advanced Computing Systems Association, which brings together engineers, system administrators, scientists, and technicians working on the cutting edge of computer science.

Finally

To find out more about the USACO or IOI 2000, go to our web site at www.usaco.org and click on 2000, then click on the International Olympiad in Informatics. To test the waters, try out our training materials at ace.delos.com/usacogate. When you are ready for a stimulating informatics challenge, join our USA 2001 competition through our USACO web site.