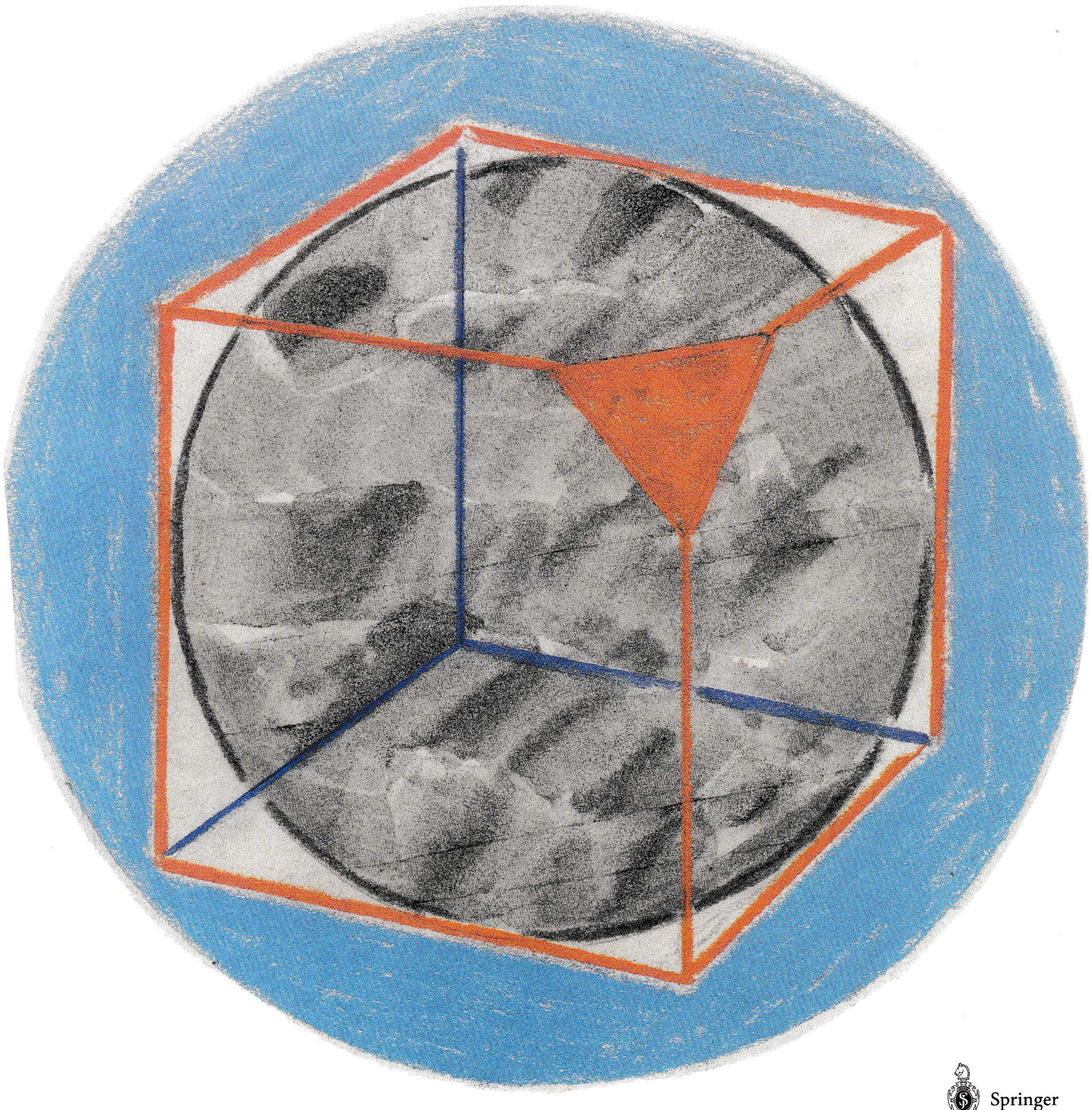
QUANTUM

MAY/JUNE 2001

US $7.50/CAN $12.00



Springer

NSTA

GALLERY Q

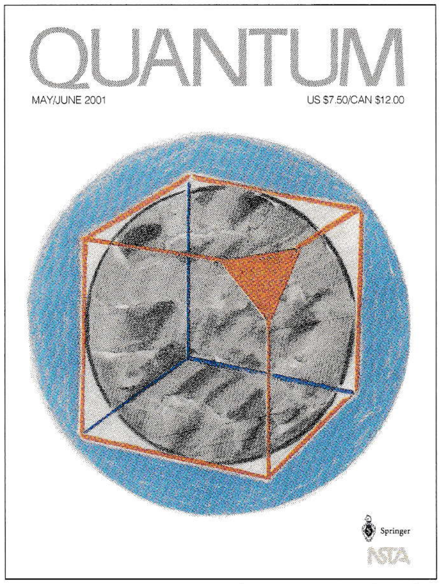
Oil on canvas, 25 5/8 × 21 3/8, Gift of Joseph H. Hazen,

Shooting Star (1936) by Joan Miró

METEORS ENTER OUR ATMOSPHERE EVERY DAY, but the lifespan of the resulting "shooting star" is so brief that catching a glimpse of one is still considered a rare treat. Rarer still is the opportunity to watch a comet move through the night sky, but the infrequency of their visits is more to blame than the shortness of their lifespan. Though long-lived, comets eventually burn out as they exhaust the fuel that creates their fiery glow. But how long can we expect a comet such as Halley's to light up the sky as it passes through our Solar System? To find out what the brightest minds have to say about this burning question, turn to page 4.

QUANTUM

MAY/JUNE 2001 VOLUME 11, NUMBER 5



Cover art by Yuri Vaschenko

In this issue we take a look at the possibility of inscribing a "chip off the old block" (a polyhedron carved from the corner of a cube) within a sphere. With a little help from Steinitz Thereom, you'll be in good shape to tackle this geometrical challenge. Turn to page 18 to start pondering these polyhedrons.

NOTICE TO SUBSCRIBERS

We regret to inform our loyal readership that the National Science Teachers Association will cease publication of *Quantum* after the July/August 2001 issue.

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

DEPARTMENTS

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.

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.

Just for the fun of it!

B321

Eating into profits. After selling his last peach for \$2.30, a merchant calculated that the average price of his peaches was \$2.45. However, a buyer returned a peach because it had a worm hole. The buyer agreed to pay only \$1.58 for this peach. The merchant recalculated the average price, which became \$2.42. How many peaches did the merchant sell?

B322

Cutting corners. A corner of a square is snipped off to form a right triangle. It turns out that the sum of the legs of the triangle is equal to a side of the square. Prove that the sum of the angles subtended by the hypotenuse at each of the three remaining vertices of the square is 90°.

B323

Checkmate math. At the end of a chess tournament, each participant won the same number of games playing white as all the other players, taken together, won playing black. Prove that all the participants won the same number of games.

Art by Pavel Chernusky

B324

Pricey address. A building has four apartments on each floor, and the apartments are numbered consecutively. The residents of one of the floors decided to place new apartment numbers on the doors. This required seven digits, which they ordered from a firm that charged n dollars for the digit n (for example, the digit 0 was free). The residents collected 3 dollars from each apartment on the floor, which exactly covered the cost of the new digits. Which digits were ordered?

ANSWERS, HINTS & SOLUTIONS ON PAGE 52

How long does a comet live?

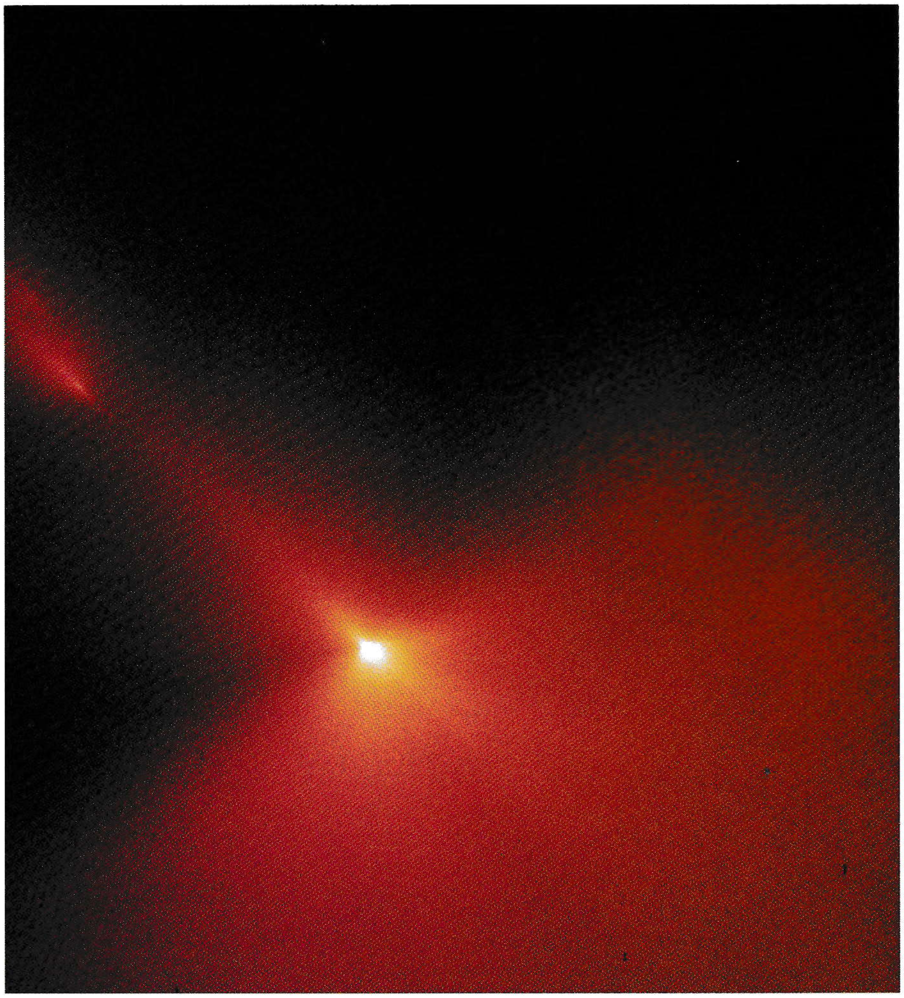
Until it runs out of gas

by S. Varlamov

THE PLANETS IN OUR SOLAR System are usually divided into two groups. The planets nearest the Sun (Mercury, Venus, Earth, and Mars) are called the terrestrial (earthlike) planets—their surfaces have approximately the same chemical composition. In contrast, the Jovian (Jupiter-like) planets (Jupiter, Saturn, Neptune, and Uranus) have much more helium and hydrogen in their outer layers. Many minor planets (asteroids) also orbit the Sun, and their chemical composition is similar to the terrestrial planets.

Sometimes the region near the Sun is visited by comets, which are chemically different from both the Jovian and the terrestrial planets. It has been hypothesized that a vast number of small heavenly bodies exist in the Solar System beyond Pluto's orbit, in the so-called Oort cloud. As a rule, the solid bodies in this cloud, together with gases and intergalactic dust, revolve around the Sun in the same direction as the planets. They move very slowly, but sometimes, when they pass very near each other, the magnitudes and directions of their velocities can change considerably.

If, as a result of such an encounter, one body gives a significant portion of its momentum to another



Photos courtesy of NASA

Comet Hyakutake. This image captures an area 3,340 km (2,070 miles) *across. It shows that most of the dust from the comet is produced on the side facing the Sun. At the upper left are three fragments from the comet that have produced their own tails.*

body, its orbital path changes, taking it closer to the Sun. Astronomers say that such a body falls from the Oort cloud into the solar region. Depending on its velocity at the outskirts of the Solar System, this body can became a "one-time visitor" or a comet with a long period. Passing near one of the large planets, such a body may again change its velocity (that is, perform a gravitational maneuver) and become a comet with a rather short period of revolution, like Halley's comet.

Astronomical observations of the emission spectra of comets' tails have shown that the comet cores consist of volatile substances such as water, methane, and ammonia. There is still much to be learned about the differences in the chemical composition of the Jovian planets, the terrestrial planets, and the comets. In this article we'll try to answer a different question: How long does the icy core of a comet live? Is its lifetime long or short compared to the lifetime of the Solar System?

Imagine that a spherical comet with an initial radius of 1 km and an initial temperature of 0 K has entered the Solar System and has begun to revolve along a circular orbit with a radius equal to half the distance between the Sun and Earth—that is, 0.5 astronomical unit (a.u.). Let our comet rotate rather rapidly about an axis normal to the orbital plane. We'll also assume that the greater part (75%) of the solar radiation is reflected from the comet's surface. These data are sufficient to evaluate the lifetime of such a comet, provided it isn't destroyed by a catastrophic collision.

The water molecules that break away from the icy surface will be returned to the surface by the comet's gravitational field only if the speed of their thermal motion is many times smaller than the escape speed

$$v_{\text{II}} = \sqrt{2GM/R}.$$

For our comet this speed is about 0.7 m/sec, so it's clear that the comet's own gravitation cannot hold the water molecules leaving its surface. In other words, there will be no atmosphere around the comet. The molecules that leave the comet will never come back.

Gaspra asteroid. The nuclei of comets probably look like this after their volatile substances have evaporated.

It's also clear that the largest flux of solar radiation will strike a comet's surface near its equator. The intensity of the sunlight falling perpendicularly on a plate of area 1 m^2 on the Earth (the so-called solar constant) is I = 1.36 kW. The same power will be absorbed by each square meter of ice near the comet's equator. Although the incident flux of the solar radiation is increased by a factor of four because the comet travels closer to the Sun, most of the energy (75%) is reflected by the comet's surface.

The power incident on 1 m^2 of equatorial surface and averaged over a large period of time is W/π = 433 W. This estimate can be obtained rather easily. Let's take a band of width h = 1 m that circles the entire equator. This band collects sunlight from an area equal to $h \cdot 2R$, while the collected energy is distributed over the band's entire area $h \cdot 2\pi R$.

Consider the following situation. As soon as the comet is placed in orbit, it's illuminated by the Sun, so the temperature of its surface rises. The external ice layers are gradually warmed and the heat is transferred to the interior of the comet. As the surface is heated, a greater role is played by the dissipation of heat to the environment. The surface ice loses heat in several ways. First, its evaporation requires energy; second, thermal radiation carries away some energy; and third, some thermal energy is spent on warming the inner layers of the comet (this heat will eventually be spent on the evaporation of surface ice or thermal radiation into space).

Now let's estimate the mean temperature at the surface of a comet orbiting for a long time about the Sun. As a way of orienting ourselves, we'll note that the mean temperature at the Earth's surface is about 290 K. Our planet dissipates energy into space mainly due to thermal radiation. Taking into account that more than 70% of Earth's surface is covered with water, and that the Earth absorbs the same energy per unit area as our model comet, the mean temperature of the comet's surface cannot be higher than 290 K.

Ice is a very poor conductor of heat (its thermal conductivity is only 2.2 W/(m · K), so the equatorial surface will be warmed rapidly. In order to remove all the heat released at the comet's surface in the equatorial zone, the slope of the temperature dependence on depth (the *temperature gradient*) must be 200 K/m.

Let's assume that only an ice layer with a thickness of 290/200 m ≅ 1.5 m is warmed, and estimate the time needed to warm this layer to a temperature of 290/2 K. Consider an ice cube with an edge length A = 1.5 m located at the equator. For the given temperature gradient (200 K/m), radiation with a power of about $A^2 \cdot 430$ W enters this cube perpendicularly to its face. The amount of heat necessary to warm the cube is $Q = cM\Delta T$. A reference book gives the value c = 2,100 J/(kg · K) for the specific heat of ice. The temperature difference of the opposite faces of the cube is somewhere between 0 and 290 K (we es-

T(K)	143	152	161	171	183	197
$\log(P/P_{atm})$	−11	−10	−9	−8	−7	−6
T(K)	212	231	253	281	319	373
$\log(P/P_{atm})$	−5	−4	−3	−2	−1	0

Table 1

timate it as $\Delta T \cong 150$ K). Finally, the mass of the ice cube is $M \cong 3{,}000$ kg. The time needed to warm such a cube is about 10^6 sec, which is slightly more than 10 days. During this time the temperature gradient decreases considerably, and the heat flowing into the comet will no longer be counterbalanced by the heat arriving at the comet's surface. We can be reasonably sure that the lifetime of a comet is far greater than 10 days, so the thermal flow into the interior of the comet can be neglected when we assess its surface temperature.

The power of the thermal energy radiated from a surface area S is $\alpha S \sigma T^4$, where $\sigma = 5.67 \cdot 10^{-8}$ W/(m$^2 \cdot$ K^4) is Boltzmann's constant, T is the surface temperature, and α is a coefficient characterizing the difference between a real radiating body and an ideal black body. In our case, this coefficient is about 1: A comet absorbs solar radiation in the visible range, while it radiates electromagnetic energy in the infrared range. Thus it follows that its temperature should be exactly 290 K!

In our reasoning we didn't take into account the heat loss caused by evaporation at the surface. The number of molecules evaporated from a unit area per unit time can be estimated from the pressure of the saturated gas of a substance at a given temperature: $P = nkT$. To an order of magnitude the number of molecules leaving the comet's surface, which is in contact with a layer of saturated vapor, is equal to the number of collisions of the vapor molecules with the surface. Imagine that each molecule striking the surface during some period of time adheres to it. During this time the same number of molecules must be vaporized. In reality, the number of molecules leaving the surface is several times less, because not every incident molecule adheres to the surface: most of the molecules recoil elastically from it.

Every vaporized molecule carries away the energy needed to escape from its neighbors, as well as the mean thermal energy corresponding to the temperature of the surface. Since there is no atmosphere around the comet, the vapor doesn't perform the work needed to expand in an atmosphere. Each unit of surface area loses the following amount of heat per second:

$$W \approx \left(\frac{\beta}{N_A}\right)\frac{P}{kT}\sqrt{\frac{3kT}{m}} = \frac{P\beta\sqrt{3}}{\sqrt{RTM}}$$

$$\approx 1.6 \cdot 10^5 \frac{P}{\sqrt{T}} \text{ (SI units)},$$

where N_A is Avogadro's number, $\beta = 3.6 \cdot 10^4$ J/mol is the molar latent heat of evaporation of water, and M is the molar mass of the ice.

Let's use experimental data describing the dependence of the saturated vapor pressure on temperature (see table 1), and calculate the dependence of the heat loss caused by evaporation on temperature (table 2).

T(K)	143	152	161	171
I(W/m^2)	0.013	0.13	1.3	12.3
T(K)	183	197	212	
I(W/m^2)	120	1150	11,000	

Table 2

Our estimates show that the equatorial surface of the comet will be heated to a temperature between 183 K and 197 K. (By the way, how will this temperature vary over the course of the comet's "day"?) At such temperatures the thermal loss to radiation per unit area will be 70 W/m^2 and 92 W/m^2 at temperatures of 183 K and 197 K, respectively. Therefore, we conclude that evaporation is the major mechanism of heat flow from the surface. The second table shows that it's responsible for about 80% of all the comet's heat loss.

Now let's estimate the lifetime of a comet if its heat loss is caused entirely by evaporation of ice on its surface. Clearly the evaporation proceeds more rapidly in the equatorial region than in the polar area. As a result of such nonuniform evaporation, the comet assumes a shape that is elongated along its axis of rotation. The rate of equatorial "thinning" of the comet doesn't depend on the comet's size, since it's determined by the balance of energy gained from solar radiation and lost by evaporation:

$$IM/(\pi\beta\rho) \cong 2 \cdot 10^{-7} \text{ m/sec}.$$

In other words, the radius of the comet's equatorial region decreases (on average) by $2 \cdot 10^{-7}$ m every second. Therefore, it only takes $2.5 \cdot 10^9$ s to completely evaporate the comet. This means that the lifetime of our comet is only about 80 years! This is a trifle compared to the age of the Solar System.

Of course, there are no comets wandering along circular orbits about the Sun. The periodic comets spend most of their lives far from the Sun. However, they lose mass most rapidly during the short periods when they travel near the Sun at comparatively small distances. If Halley's comet had an ice nucleus with a diameter of 10 km and passed near the Sun at a distance of 0.5 a.u. (the latter figure corresponding to an 80-year period), it would be completely evaporated after only 1,300 revolutions—that is, after 100,000 years, which is a

CONTINUED ON PAGE 17

Challenges

Physics

P321

Chalk line. A piece of chalk rests on a horizontal board with a coefficient of friction μ. Suddenly the board starts to move horizontally at a speed v_0, and after a time τ it stops abruptly. Find the length of the line drawn by the chalk on the board.

(A. Zilberman)

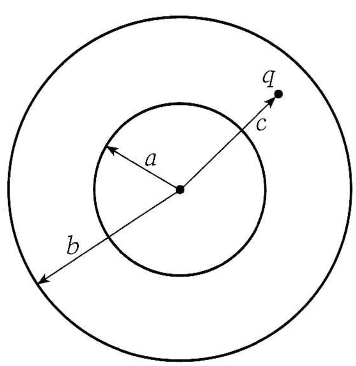


Figure 1

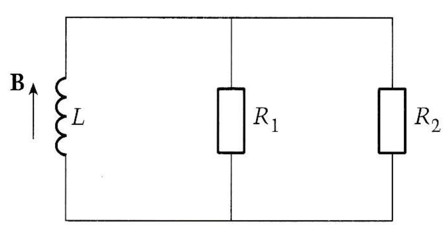


Figure 2

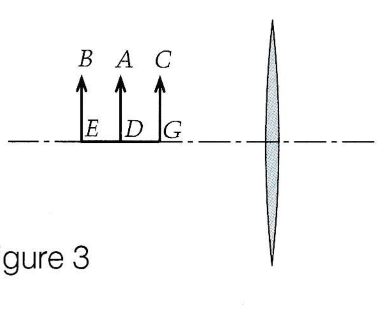


Figure 3

P322

Density at a distance. Astronomers discovered a very dense planet in the system $\tau_{\text{Lynx Major}}$. The period of the planet's rotation about its axis is only T = 6 min. Calculate the planet's density.

P323

Charges and spheres. A point electric charge q is located between two uncharged metallic, concentric spheres of radii a and b at a distance c from the center of the system (figure 1). What charge will flow through a thin wire when it touches both spheres? (V. Komov)

P324

Calculate what dissipates. A coil with inductance L (figure 2) is situated in a magnetic field. The field is abruptly turned off. Immediately thereafter, an electric current I flows through the resistor R_1. Neglecting the resistance of the coil, find the amount of heat dissipated by each resistor (R_1 and R_2). The direction of the magnetic field was perpendicular to the plane of the turns of the coil. (V. Mozhayev)

P325

Tri to solve this. The image of a trident $BACGDE$ is formed by a thin converging lens (figure 3). The base EDG of the trident coincides with the principle optical axis, and AB = AC (that is, the trident is symmetrical). The lens magnifies the segments DG and ED by factors of β_1 and β_2, respectively. By how much is the segment AD magnified?

(Y. Cheshev)

Math

M320

54-gon conclusions. Prove that in (a) a regular 12-gon and (b) a regular 54-gon there exist four diagonals that meet at a point and do not pass through the center of the polygon.

(S. Tokarev)

M321

Budget items. A chamber consisting of 2,000 deputies must pass a state budget that consists of 200 expense items. Every deputy prepared a draft budget, where the maximum expenditure for every item is specified and the total expenditures do not exceed a given value S. For each item the chamber passes a maximum expenditure that is approved by no fewer than k deputies. What is the minimum value of k for which the total expenditure can be guaranteed to be not greater than S?

(I. Sergeyev)

M322

Find the sum. The numbers α and β satisfy the equations

$$\alpha^3 - 3\alpha^2 + 5\alpha = 1,$$
$$\beta^3 - 3\beta^2 + 5\beta = 5.$$

Find $\alpha + \beta$. (V. Kukushkin)

M323

Clean slate. Written on a blackboard are n numbers. The following operation is performed on these numbers: two numbers, a and b, are erased, and the number $(a + b)/4$ is

CONTINUED ON PAGE 11

In search of perfection

What's beyond 6 and 28?

by I. Depman

THE FAMOUS GREEK PHIlosopher and mathematician Nicomachus of Gerasa wrote: "Perfect numbers are beautiful. It is well known that beautiful things are rare, whereas ugly ones are numerous. Almost all numbers are either abundant or deficient, but perfect numbers are extremely rare."

How numerous are perfect numbers? Nicomachus, who lived in the 1st century A.D., didn't know the answer.

The first perfect number that was known in ancient Greece was 6. The most honored guest at a banquet occupied the sixth place. According to Pythagorean doctrine (and Nicomachus was a Pythagorean), the number 6 possessed various mystical properties.

Plato devotes particular attention to this number in his *Dialogues*. It is not without reason that the biblical tradition states that God created the world in six days; indeed, 6 is the first of the perfect numbers.

In 1961, the Mathematical education *published an article by professor I. Depman on perfect numbers, which are equal to the sum of their own divisors. Here we print a revised version of this article prepared for publication by Academician V. V. Petryanov-Sokolov. The supplement on recently discovered perfect numbers was written by Professor A. A. Bukhshtab.*

The next perfect number known in ancient times was 28. In 1917, a strange ancient structure was discovered in Rome: 28 cells were situated around a large central hall. This was the site of the Neo-Pythagorean Academy of Sciences, which included 28 members. Until recently, many scientific societies traditionally consisted of 28 members, although the reason for such a number had long been forgotten.

Ancient mathematicians were impressed by the peculiar property of these two numbers—they are equal to the sum of their own divisors:

$$6 = 1 + 2 + 3;$$
$$28 = 1 + 2 + 4 + 7 + 14.$$

Only two perfect numbers had been known before Euclid, and nobody knew if other such numbers existed or how numerous they were. The great founder of geometry focused much effort on properties of numbers, and he was certainly interested in perfect numbers. Euclid proved that any number that can be represented as a product of the factors 2^{p-1} and $2^p - 1$, where $2^p - 1$ is prime, is perfect. For $p = 2$, Euclid's formula $2^{p-1}(2^p - 1)$ yields $2^{2-1}(2^2 - 1) = 6$, which is the first perfect number. For $p = 3$, we obtain the second perfect number:

$$2^{3-1}(2^3 - 1) = 28.$$

Using his formula, Euclid found two more perfect numbers—for $p = 5$ and for $p = 7$. They are

$$2^{5-1}(2^5 - 1) = 2^4(2^5 - 1) = 16 \cdot 31 = 496$$

and

$$2^{7-1}(2^7 - 1) = 2^6(2^7 - 1) = 64 \cdot 127 = 8128.$$

The reader may enjoy verifying that they are indeed perfect.

For almost fifteen hundred years these four numbers were the only perfect numbers known. Nobody knew if additional "Euclidean" perfect numbers, or perfect numbers found in other ways, existed.

The intractable nature of this problem and the difficulty of working with perfect numbers led some to think that they were divine. One of the most famous scientists of the Middle Ages, a friend and teacher of Charlemagne, the Abbot Alcuin, wrote textbooks on arithmetic and organized schools. He was convinced that humanity is imperfect

Art by Vasily Vlasov


2305 843 009 213693951
2305 843 008 139 952 128
B.B.
p=3217

and evil, simply because humans descended from eight persons who escaped on Noah's ark—and eight is an imperfect number. Before the flood, human beings were better; indeed, it was supposed we descended from a single person—Adam, and unity can be considered a perfect number (it's equal to its single divisor). Alcuin lived in the 8th century, but even in the 12th century the Church taught that it was quite sufficient to study perfect numbers to save one's soul, and that a person who found a new perfect number could look forward to eternal bliss.

However, even the promise of this reward didn't help medieval mathematicians. The fifth perfect number wasn't discovered until the 15th century. This number also satisfies Euclid's formula.

Actually, it's surprising that this number was discovered as early as the 15th century. It's equal to

$$33{,}550{,}336$$

and corresponds to $p = 13$ in Euclid's formula.

Two hundred years later, the Frenchman Marin Mersenne, a mathematician and musician, one of the founders of the French Academy of Sciences, and a friend of Descartes and Fermat, declared without proof that the next six perfect numbers that satisfy Euclid's formula correspond to values of p equal to 17, 19, 31, 67, 127, and 257.

It was clear to Mersenne's contemporaries that he was unable to verify his assertion by direct calculations. Indeed, to do so, he would have had to prove that the numbers $2^p - 1$ were prime for the corresponding values of p. It's not difficult to calculate those numbers, but it was almost impossible at that time to find out whether or not they're prime. Thus it remained unknown whether or not Mersenne was right.

It was discovered later that the Italian Cataldi, who was a professor of mathematics in Florence and Bologna and the first to describe a method for calculating square roots, also studied perfect numbers (apparently to save his soul). In his notes, the values of the sixth and seventh perfect numbers were given—almost a hundred years before Mersenne. These were

8,589,869,056 (the sixth perfect number),

137,438,691,328 (the seventh perfect number).

Both numbers were identical to those given by Mersenne:

$$2^{16}(2^{17} - 1) \text{ and } 2^{18}(2^{19} - 1).$$

However, the fact that these numbers were perfect remained unproved. It was necessary to verify that $2^{17} - 1$ and $2^{19} - 1$ were prime.

An academician from St. Petersburg, a founder of modern mathematics, the great Leonhard Euler, was a human calculator par excellence. He proved a new theorem on these mysterious and enigmatic perfect numbers. He proved that all even perfect numbers have the form indicated by Euclid. As for the form taken by odd perfect numbers, or whether they even exist—that question persists to the present day.

Euler proved that the first three numbers of those indicated by Mersenne—namely, $2^{17} - 1$, $2^{19} - 1$, and $2^{31} - 1$—are prime. Thus the sixth and seventh perfect numbers found by Cataldi and then Mersenne turned out to be correct. We don't know, and probably never will know, how they were found. The only "explanation" available, which was given by contemporaries of those scientists, is that they were aided by Providence.

The eighth perfect number, corresponding to $p = 31$ in Euclid's formula, is

$$2{,}305{,}843{,}008{,}139{,}952{,}128.$$

For an entire century this number remained the largest perfect number known. In this period, a new method was found for verifying whether or not the number $2^p - 1$ is prime without performing direct calculations. It turned out that not all numbers indicated by Mersenne were perfect. He predicted the value $p = 127$ correctly, but the numbers with $p = 67$ and $p = 257$ are not perfect, contrary to Mersenne's expectations. On the other hand, the numbers corresponding to $p = 61$, $p = 89$, and $p = 107$ turned out to be perfect.

The ninth perfect number wasn't calculated until 1883. It was found by the Russian priest I. M. Pervushin. He calculated the largest prime number (of his time) of the form $2^p - 1$ for $p = 61$—namely,

$$2{,}305{,}843{,}009{,}213{,}693{,}951$$

—and the corresponding perfect number

$$2{,}305{,}843{,}009{,}213{,}693{,}951 \cdot 2^{60}.$$

This calculation was a heroic deed. Mersenne said that even eternity was not long enough to check if a number consisting of 15 to 20 digits is prime. Pervushin didn't use any technology in his work, and his number consisted of 37 digits.

At the beginning of the 20th century, the first mechanical calculating devices appeared, which facilitated the search for new perfect numbers.

The tenth perfect number was found in 1911, consisting of 54 digits:

$$618{,}970{,}019{,}642{,}690{,}137{,}449{,}562{,}111 \cdot 2^{88}.$$

The eleventh number, consisting of 65 digits, was discovered in 1914:

$$162{,}259{,}276{,}829{,}213{,}363{,}391{,}578{,}010{,}288{,}127 \cdot 2^{106}.$$

The twelfth perfect number was also discovered in 1914. It consists of 77 digits and is written as

$$2^{126}(2^{127} - 1).$$

In 1932, Lemere set himself the goal of finding the thirteenth perfect number. For this purpose, he decided to check if the last number $2^p - 1$ (for $p = 257$) indicated by Mersenne was prime. Using mechanical calculators available at that time, he spent a year verifying that this number is composite. Thus, the twelfth perfect number remained the largest one until 1952.

* * *

The thirteenth perfect number was found with the help of a computer. On January 30, 1952, the American mathematician Robinson used a computer to check whether numbers of the form $2^p - 1$ are prime. To begin, Robinson checked the number $2^{257} - 1$. He invited Lemere, who had spent a year on this work 20 years earlier, to witness the calculation. Lemere was glad to see that the computer needed only 18 seconds to obtain the result. To find a new perfect number, it was necessary to find a new prime satisfying Euclid's formula. The computer continued its calculations. In two hours it checked 42 numbers, the smallest of which consisted of 80 digits. However, all these numbers turned out to be composite. A new perfect number was found that evening. It was

$$2^{520}(2^{521} - 1) \qquad (p = 521).$$

The thirteenth perfect number consists of 314 digits.

The fourteenth perfect number was found the same day toward midnight. After checking thirteen more Euclidean numbers, it found the prime $2^{607} - 1$, which is written with 183 digits. The corresponding perfect number is

$$2^{606}(2^{607} - 1) \qquad (p = 607).$$

The fourteenth perfect number has 366 digits.

The fifteenth perfect number was found in June 1952. Computers were scarce at that time, and the machine could study the problem of perfect numbers only "at leisure."

Further work brought the prime number $2^{1279} - 1$ and the corresponding perfect number consisting of 770 digits:

$$2^{1278}(2^{1279} - 1) \qquad (p = 1279).$$

The sixteenth and seventeenth perfect numbers were discovered in October 1952. By that time, the computer found two more Euclidean primes: $2^{2203} - 1$ and $2^{2281} - 1$. The corresponding perfect numbers are

$$2^{2202}(2^{2203} - 1) \qquad (p = 2203),$$

consisting of 1327 digits, and

$$2^{2280}(2^{2281} - 1) \qquad (p = 2281),$$

consisting of 1373 digits.

The eighteenth perfect number was found in September 1957 by the Swedish mathematician G. Riesel. Using a computer, he spent five and a half hours verifying that the number $2^{3217} - 1$ is prime and obtaining the eighteenth perfect number:

$$2^{3216}(2^{3217} - 1) \qquad (p = 3217),$$

consisting of about 2000 digits.

The search for larger perfect numbers required more and more computational effort. But computers were becoming more and more powerful, so in 1962 two new perfect numbers were discovered, and in 1965 three more. In the Euclidean formula, these numbers correspond to p = 4253, 4423, 9689, 9941, and 11213. The perfect number $2^{11212}(2^{11213} - 1)$ consists of 3376 digits. Obviously such numbers could never be found and verified without the help of powerful computing technology.

And there we have it: everything the human race has learned about perfect numbers over the past two thousand years.[1]

The history of searching for perfect numbers shows how computers can increase our capabilities. However, in the words of Edmund Landau, the great expert in number theory, "*Two problems remain open:*

Are there infinitely many even perfect numbers?—I don't know.

Are there infinitely many odd perfect numbers or does at least one such number exist?—I don't even know if one such number exists."

What is there to add?

Exercises

1. Prove that the number $2^{k-1}(2^k - 1)$, where $2^k - 1$ is a prime, is perfect.

2. Denote by $\sigma(n)$, where n is a natural number, the sum of all divisors of n. Prove that if n_1 and n_2 are coprime, then $\sigma(n_1 \cdot n_2) = \sigma(n_1) \cdot \sigma(n_2)$.

3. Let n be an even perfect number. Then $\sigma(n) = 2n$. Represent n as $2^{k-1}b$, where $k \geq 2$ and b is odd, and prove that $b = (2^k - 1)c$. Then prove that $c = 1$ and $2^k - 1$ is prime.

[1]Actually, we do know about more than 30 prime Mersenne numbers, the largest of which corresponds to the prime number p = 216091.

CONTINUED FROM PAGE 7

written in their place. This operation is repeated $n - 1$ times. As a result, a single number remains on the blackboard. Prove that if all the original numbers are equal to one, the resulting number is not less than $1/n$.

(B. Berlov)

M324

Crime solver. An investigator devised a plan for interrogating a witness that guarantees that a crime would be solved. He is going to ask questions that assume only a yes or no answer. The next question may depend on the answers obtained to the preceding questions. The investigator assumed that all the answers would be correct. He calculated that not more than 91 questions had been asked regardless of the answers. Prove that the investigator can devise a plan consisting of not more than 105 questions that guarantees the crime will be solved even if one of the answers may be false (however, all the answers are allowed to be correct as well).

Note: If you can devise only plans consisting of more than 105 questions, give the best of them.

(A. Andzhans, I. Solovyov, and V. Slitinsky)

ANSWERS, HINTS & SOLUTIONS ON PAGE 50

Peering into potential wells

A potential well is a hole that hasn't been dug yet.
—Scientific folklore

by K. Kikoin

IN PHYSICS WE OFTEN COME across sentences like: "A particle (or a system) is situated in a potential well." What is this well, and why must any physical object eventually get trapped in it? Let's take a close look at this potential pitfall. We'll examine several concrete examples, starting with the simplest one.

Ball and box

Let's drop a metal ball into a box with an uneven bottom, one having hills and valleys. After bouncing and rolling around, the ball will come to rest at the bottom of a valley. Why does the ball stop at the bottom and not, say, on a slope of the uneven surface? The obvious answer is the following. Two forces act on the ball: the force of gravity (its weight) $\mathbf{F}_g$ and the normal force $\mathbf{F}_n$. When the ball is located on the slope of a hill or valley (figure 1), the net force $\mathbf{F}$ is directed downward along the slope, so according to Newton's second law the ball accelerates and moves to the bottom. At the bottom, the two forces ($\mathbf{F}_g$ and $\mathbf{F}_n'$) cancel. So, eventually, the ball will come to rest at this spot.

Now let's consider the same problem from another perspective. A ball resting at the bottom of a valley has no kinetic energy. In other words, all the ball's energy is potential (E_p). Let's imagine moving the ball from this equilibrium position. Obviously, work must be performed against the force of gravity during any motion of the ball up the hill. This work will increase the potential energy of the ball. However, if the potential energy increases during any displacement r relative to the equilibrium position, it is minimal at the equilibrium point ($r = 0$, figure 2).

Now let's follow the conversion of energy of a ball thrown into the box. At the outset the ball had an initial amount of kinetic and potential energy. The up and down motion of the ball along the hills and valleys on the bottom is accompanied by the transformation of kinetic to potential energy and back to kinetic energy. In addition, a very important process occurs continuously: The frictional forces perform work (eventually this energy is dissipated as heat to the surroundings). At long last, the total amount of mechanical (kinetic plus potential) energy will become so

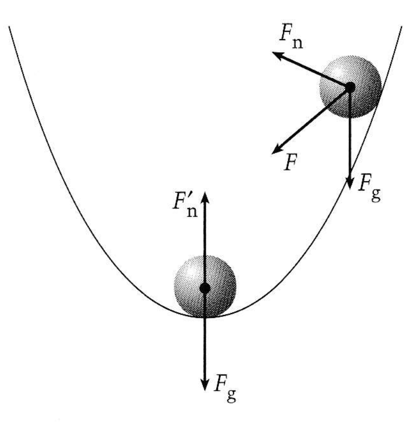


Figure 1

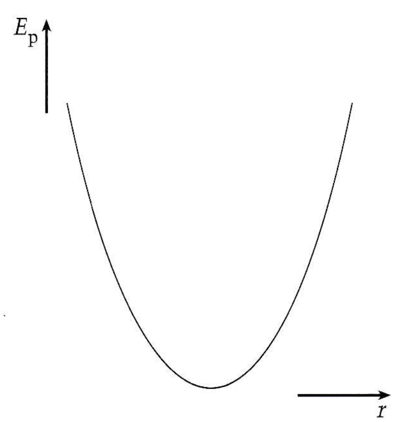


Figure 2

Art by Leonid Tishkov


Tishkov 2000

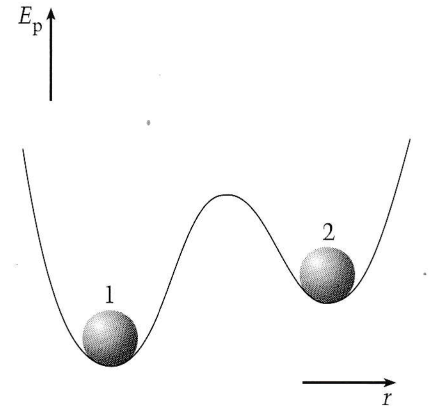


Figure 3

small that the ball will not get out of the valley. After many oscillations, the ball will come to rest at the bottom of the valley. In this position, all the residual energy of the ball is potential. Moreover, the value of this energy is minimal within the valley. For example, figure 3 shows two possible equilibrium positions of a ball in adjacent valleys of various depths. Each equilibrium position is characterized by its own minimum of the potential energy, and $E_{p\,min\,1} < E_{p\,min\,2}$ (this means that positive mechanical work must be performed to shift the ball from position 1 to position 2).

And so we arrive at an alternative definition of stable equilibrium: *The stable equilibrium of an object always corresponds to the minimum of its potential energy*. In other words, *when in stable equilibrium, an object is situated in a potential well.*

The example we looked at is very simple and clear: For a ball, a potential well is literally a depression or a low spot. Next we'll move on to an object for which no "well" in the usual sense of the word can be imagined, but the potential well exists nevertheless and determines its equilibrium state.

Soap bubble

Every child knows that a soap bubble is always round. But does every child know why?

To answer this question, we need to recall the phenomenon of surface tension. From the physical point of view, a soap bubble is a liquid film that delimits a volume filled with gas. Looking at this film at a greater magnification, one can see that it consists of two surface layers (inner and outer) and the liquid between them. The molecules on a surface are subject to quite different conditions from those affecting molecules between the surfaces. Every molecule inside the liquid is surrounded by similar molecules, whose effects on a given molecule is therefore zero. In contrast, a surface molecule is affected by significantly different forces from the adjacent liquid and the gas. Since the density of a gas is far less than that of a liquid, the net force acting on a surface molecule is always directed inward, into the liquid. Therefore, in order to leave the surface, the molecules must perform a fixed amount of work against these attractive forces. In other words, the surface molecules have a greater potential energy than the inner molecules. This extra potential energy is referred to as the surface energy E_s:

$$E_s \sim S,$$

or

$$E_s = \sigma S.$$

The coefficient of proportionality σ is the surface tension. It's equal to the ratio of the work W needed to increase the surface area by ΔS to ΔS itself. Every liquid is characterized by its own coefficient of surface tension.

When we inflate a soap bubble, its volume (and the surface area of the film) increase, although the amount of liquid in the surface remains constant. Clearly, an increase in surface area is possible only because new molecules move to the surface from the interior of the liquid. This means that some work must be performed to drive this process and increase the surface energy of the film.

Now we can explain why a soap bubble is always round and not, say, an ellipsoid or polyhedron. The sphere has this wonderful property: Among all geometrical objects of the same volume, it has the smallest surface area. Therefore, if we try to deform a soap bubble to make its shape ellipsoidal, we must perform a certain amount of work to increase its surface area and surface energy. In other words, *the spherical shape of a soap bubble corresponds to the minimal potential energy of the film that forms its shell.*

So, you may ask, what lies at the bottom of the potential well in this example? Obviously it's not an individual particle, since the surface tension results not from the behavior of individual molecules but from the interaction of a vast number of molecules. We can't even say that the potential well contains all the surface molecules, since a change in the bubble's radius is accompanied by movement of molecules from the surface to the interior of the liquid and vice versa. In addition, the continuous exchange of molecules between the surface and deep layers occurs due to thermal motion. Therefore, *the potential well contains all the liquid that forms the soap film.*

We could apply the concept of potential energy (and the potential well) to this problem only because there is a universal parameter that describes the energy of the system—the surface area of the bubble, which is simply proportional to its surface energy.

Strictly speaking, we can describe in this way only half of the potential well corresponding to the spherical form of a bubble (figure 4). We can decrease the area only by compress-

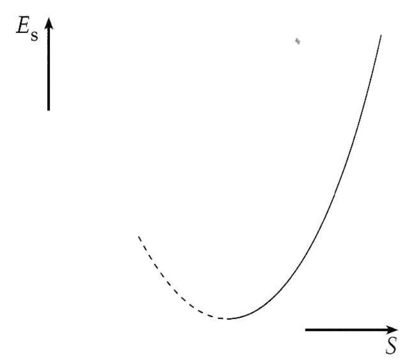


Figure 4

ing the bubble—that is, by performing work in compressing the gas inside the bubble. The corresponding increase in the system's energy in this case in shown by the broken line in figure 4. So in reality the potential well of a soap bubble contains the complex system "surface molecules of the liquid + interior liquid molecules + molecules of gas inside the bubble."

Ionic crystal

Now we come to our last example of a system situated in a potential well. The shape of the potential well for such a system can be found by means of precise but very cumbersome calculations. However, an approximate result can be obtained with very simple reasoning. We'll take this efficient route, in which we sacrifice a little precision for the sake of much greater clarity.

This system is the ionic crystal, a typical example of which is ordinary table salt (NaCl). It's known that atoms (or ions) in crystal bodies form regular geometric lattices. The crystal of table salt has a simple cubic lattice, the nodes of which contain positive sodium (Na^+) and negative chloride (Cl^-) ions arranged in a three-dimensional checkerboard pattern (figure 5). Every Na^+ ion has six Cl^- ions as its nearest neighbors, and every Cl^- ion has six neighboring Na^+ ions. The distance between the two nearest identical ions on the face of the cube is called the lattice constant a.

To determine the shape of the potential well of a particular crystal, we need to find its total potential energy, which is the energy of interaction between all the ions in the crystal. As a first step, we consider the interaction between the closest neighboring Na^+ and Cl^- ions. This interaction is composed of the coulomb attraction and the quantum mechanical repulsion, which prevents the ions from getting too close to one another.

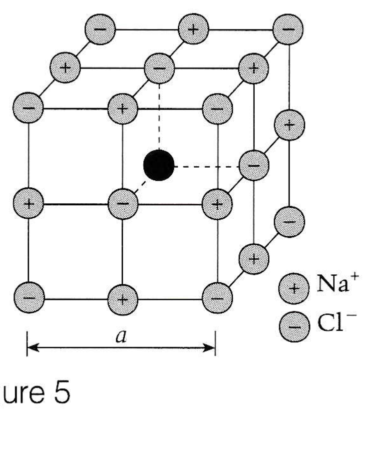


Figure 5

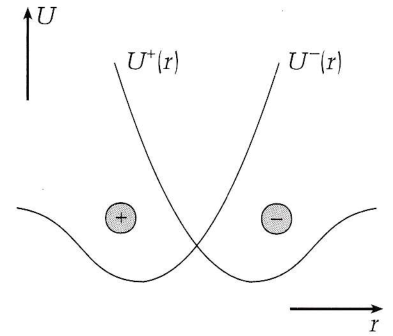


Figure 6

We can describe this interaction using the notion of potential. Imagine a sodium ion situated in the attractive electric and repulsive quantum fields generated by a chloride ion. According to electrostatics, the potential of the attractive field is $-ke/r$. Here k is Coulomb's constant, e is the charge of a monovalent ion equal to the charge of an electron, and r is the distance between the sodium and chloride ions. As a rule, precise expressions of the repulsive potential for various ions are not known, but as an approximation we can describe it by the rapidly decreasing power function b/r^n, where b is a certain constant. The exponent n for crystals like NaCl is assumed to be 9.

Thus a positive Na^+ ion is situated in a compound field of negative Cl^- ion, whose potential is described by the following formula:

$$U^-(r) = -\frac{ke}{r} + \frac{b}{r^n}.$$

Since n is large, at small distances the repulsion potential increases very rapidly and becomes substantially larger than the attraction potential. At large distances, the repulsion potential is virtually zero, so the total potential coincides with the coulomb potential. As a result, $U^-(r)$ takes the form of a curve with a minimum and forms the potential well for the sodium ion (figure 6). In turn, a sodium ion generates a similar well to trap the neighboring chloride atom.

Therefore, two adjacent ions create potential wells for each other by repulsion at small distances and attraction at large distances. In these wells both ions keep each other in the stable equilibrium state.

Of course, an entire crystal is a far more complex system than the pair of unlike ions we just considered. In addition to several neighboring ions with opposite charges, any ion is surrounded by ever more distant ions that also affect the given ion and therefore contribute to the formation of the total potential function. Although the size of a crystal is limited, the number of neighbors of any ion can be considered infinite. Indeed, consider the most finely ground table salt. Its crystals are about 0.1 mm across, while the lattice constant is $a \sim 10^{-8}$ cm, so the edge of a crystal of table salt is composed of $\sim 10^6$ atoms. Correspondingly, the area of a face and the volume of a crystal of table salt have 10^{12} and 10^{18} identical ions.

A calculation of the total energy of interaction of all the ions in the crystal would seem to be insanely difficult at first glance. However, using a clever trick based on certain fundamental physical properties, we can solve the problem in a few strokes. Details are provided in Supplement II, so here we'll show only the final result. The total potential energy of all the ions in the crystal is described by the formula

$$U_{cr}(a) = -\frac{Ae^2}{a} + \frac{B}{a^n},$$

where A and B are constants. The plot of this function is a curve with a minimum (that is, with a potential well). What's at the bottom of this well? The entire crystal, of course! This means that the lattice constant assumes a value a_0 at equilibrium, corresponding to the minimum of the potential energy.

To sum up, we've looked at three objects that are completely different at first glance. However, a common feature emerged—the potential energy of each took the shape of a curve with a minimum. In the general case, the shape of a potential well may be more complex (in particular, the potential energy may depend on many parameters rather than a single one). Still, the general principle is always the same: *At stable equilibrium any system has minimal potential energy—in other words, it is situated in a potential well.*

Supplement I

Let's calculate the potential ϕ_1 created at point 1 with coordinate r_1 by an electric charge $+e$ located at the origin of the coordinate system (figure 7). This potential is equal to the work of the electric field needed to move a unit positive charge from the given point to infinity.

Let's place a test charge q at point 1. According to Coulomb's law, it experiences a force $F = kqe/r_1^2$ due to the presence of charge e. Now we calculate the work W_{1-2} needed to transfer charge q from point 1 to the nearest point 2. To simplify our calculations, we'll choose a distance $r_2 - r_1$ so small that we can consider the force to be constant over the entire interval 1–2: $F_1 = kqe/r^2 = kqe/(r_1 r_2)$. Thus

$$W_{1-2} = F_1(r_2 - r_1) = \frac{kqe}{r_1 r_2}(r_2 - r_1) = kqe\left(\frac{1}{r_1} - \frac{1}{r_2}\right).$$

Similarly, we can calculate the work W_{2-3} needed to transfer the charge q from point 2 to the neighboring point 3:

$$W_{2-3} = F_2(r_3 - r_2) = \frac{kqe}{r_2 r_3}(r_3 - r_2) = kqe\left(\frac{1}{r_2} - \frac{1}{r_3}\right).$$

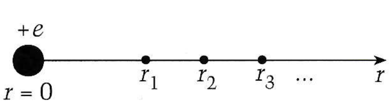


Figure 7

Clearly the work of the field in the interval 1–3 equals the sum of the work performed in the intervals 1–2 and 2–3:

$$W_{1-3} = W_{1-2} + W_{2-3} = kqe\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + kqe\left(\frac{1}{r_2} - \frac{1}{r_3}\right) = kqe\left(\frac{1}{r_1} - \frac{1}{r_3}\right).$$

Continuing this calculation, we see that the intermediary points play no role, and the resulting expression for the work contains only the coordinates of the initial and final points. Therefore, the work of the electric field to transfer the test charge q from point 1 to infinity ($1/r_\infty = 0$) is

$$W_1 = kqe\left(\frac{1}{r_1} - \frac{1}{r_\infty}\right) = \frac{kqe}{r_1}.$$

Finally, the potential at point 1 is determined by the formula

$$\varphi_1 = \frac{W_1}{q} = \frac{ke}{r_1}.$$

Those who know how to integrate could have obtained this result in one line (if you ever needed an incentive to learn calculus).

Supplement II

In order to find the total potential energy of a crystal, let's start with a computation of its electrostatic energy. Consider a sodium ion Na^+ located at the center of the cubic lattice shown in figure 5. Its closest neighbors are six chloride ions located at distance $r = a/2$ from it. Each of these chloride ions generates a potential $-ke/r = -2ke/a$. According to the superposition principle, the potential of several charges is equal to the potentials generated individually by each charge. Therefore, the six chloride ions generate a potential of $-12ke/a$. The next set of neighboring charges are 12 sodium ions located at a distance $a/\sqrt{2}$ from the chosen (central) sodium ion. These sodium ions generate a total potential $+12\sqrt{2}ke/a = +16.97ke/a$. Continuing this calculation, we find that the third set of charges produces a potential $-9.24ke/a$, the fourth set a potential of $+6ke/a$, and so forth.

There is no simple regularity in this sequence (for example, the nineteenth charge set contributes $-3.08ke/a$ to the total potential), so it's useless to compute the infinite sum by adding the successive members of the series. However, there's something interesting about this series: The sign of its members alternates, and this implies the possibility of some rearrangement of the addends to compute the total sum. Indeed, such a clever rearrangement does exist, and the idea is based on the notion of electric neutrality.

Let's calculate the total charge of all the ions of the cube shown in figure 5. We'll assume that only a half of each ion located at a face belongs to the cube, a quarter of the ions located on an edge, and one eighth of the ions located at a corner. This is by no means a casual assumption: If we subdivide the entire crystal into cubes similar to the one we're considering, we'll see that an ion on a face belongs to two cubes, an ion on an edge belongs to four cubes, and an ion at a corner belongs to eight cubes. Our great achievement is this: The total charge of such a cube composed of partial charges is exactly zero! In other words, we found a way to divide the crystal into electrically neutral cells. The coulomb energy of the neutral cube is

$$-\frac{12}{2}\frac{ke^2}{a} + \frac{16.97}{4}\frac{ke^2}{a} - \frac{9.24}{8}\frac{ke^2}{a} = -2.91\frac{ke^2}{a}.$$

Now let's consider a more complex set of ions. It will be composed of the initial cube (with whole ions) enveloped by a single layer of ion fragments. The edge of this new cube is $2a$; the ions of the inner cube belong entirely to the new construction, while only a half, a quarter, or an eighth of the enveloping ions will belong to the system, depending on their location. This is a more complex division of the entire crystal into an electrically neutral cell. We would expect that the resulting potential energy will be more precise. Indeed, the electrostatic energy of the new electric cell is $-3.50\ ke^2/a$. If we take an electrically neutral cube with a side length of $3a$, its energy will be $-3.49\ ke^2/a$. Further increases in the volume of the cube have practically no effect, and eventually we obtain a value for its energy equal to approximately $-3.495\ ke^2/a$.

The same idea of dividing the entire crystal into elementary cells suggests another way of approximating the crystal energy. Take an ion with charge +1 and surround it with a sphere of radius a filled with a uniformly distributed negative charge whose total value is –1. Calculating the electrostatic energy of this neutral sphere yields the value $-3.6\ ke^2/a$, which is very close to the correct value.

Thus a correct model that takes into account the most essential features of a physical process greatly simplifies our calculations.

Paradoxically, the energy associated with the quantum repulsion can be calculated in a far simpler way, although the nature of this repulsion is far from trivial. The repulsion potential decreases very rapidly with distance. Thus we can take into consideration the repulsion that acts only between neighboring ions at a distance $a/2$ from one another. In this approximation, the total energy of repulsion is just the sum of the energies of repulsion ("pairwise interaction") counted for all ion pairs. We immediately get $+B/a^n$, where $B = Nb$, N being the number of pairs in the cell. The electrostatic energy corresponding to individual ions calculated above must also be multiplied by this factor N. As a result, we get the formula for the total energy of the cubic crystal used in the main body of this article.

Quantum articles on potential and crystals:

A. Eisenkraft and L. D. Kirkpatrick, "Electricity in the Air," November/December 1992, pp. 46–48.

Y. Bruk and A. Stasenko, "Hard-Core Heavenly Bodies," July/August 1993, pp. 34–38.

A. Stasenko, "From the Edge of the Universe to Tartarus," March/April 1996, pp. 4–8.

V. Meshcheryakov, "Planetary Building Blocks," July/August 1998, pp. 4–10.

A. Leonovich, "Do You Have Potential?" November/December 1998, pp. 28–29.

A. Dozorov, "Electrical Multipoles," September/October 1999, pp. 4–8.

G. Myakishev, "Molecular Interaction Up Close," May/June 2000, pp. 8–13.

CONTINUED FROM PAGE 6

negligible period compared to the age of the Solar System (5 billion years). This explains why comets are such rare guests in the firmament and do not visit us every other night—they simply cannot collect in any significant numbers near the Sun.

The appearance of a comet with a period comparable to the period of revolution of the large planets about the Sun is a rare event because it requires a combination of circumstances whereby the comet, passing near the Sun, must be subjected to a strong interaction with one of the large planets. The planet must modify the comet's trajectory such that its long period (or "nonperiod," if it would have been a one-time visitor) turns into a short period. What's the probability of such an event? Perhaps a *Quantum* reader will be the one to answer this question.

Imagine that every ten years (on average) a comet appears in the night sky whose characteristics are similar to those of Halley's comet. This would mean that about 10 large objects periodically appear near the Sun and are visible as comets. Since the lifetime of such a comet is about 100,000 years, one would conclude that the frequency with which such comets emerge from the Oort cloud is about once every 10,000 years, and that approximately a half-million comets like Halley's have appeared and vaporized near the Sun while the Earth has been in existence.

Quantum on comets and the Universe:

A. Leonovich, "What's New in the Solar System?" November/December 1990, pp. 32–33.

Y. Osipov, "Catch as Catch Can," January/February 1992, pp. 38–43.

Y. Zeldovich, "A Universe of Questions," January/February 1992, pp. 6–11.

Y. Solovyov, "The Universe Discovered," May/June 1992, pp. 12–18.

W. A. Hiscock, "The Inevitability of Black Holes," March/April 1993, pp. 26–29.

A. G. W. Cameron, "When a Body Meets a Body," January/February 1995, pp. 16–20.

S. Silich, "Interstellar Bubbles," November/December 1997, pp. 14–19.

I. D. Novikov, "The Thermodynamic Universe," March/April 1998, pp. 10–14.

V. Surdin, "A Star Is Born," March/April 2000, pp. 12–17.

Uninscribable polyhedrons?

Proving Steinitz Theorem

by E. Andreev

CONSIDER A CUBE WITH ONE of its vertices cut off by a plane (figure 1). It would be interesting to determine whether this polyhedron can be inscribed in a sphere. Does the answer depend on the cutting plane? This article is dedicated to solving this and similar problems.

Let a convex bounded polyhedron be given. More precisely, we'll deal with a three dimensional body that is bounded on all sides by plane polygons (called faces) and lies on one side of the plane of each face. We want to find out whether this polyhedron can be inscribed in a sphere.

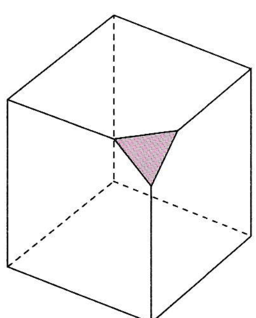


Figure 1

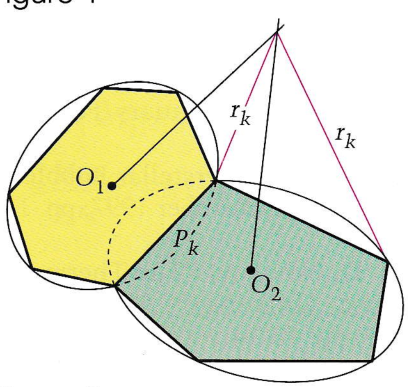


Figure 2

We'll designate the given polyhedron by M, and number (separately) its faces, edges, and vertices, calling the ith face F_i, the ith edge E_i, and the ith vertex V_i. Two faces are called adjacent if they have a common edge, and two vertices are considered neighboring if they are connected by an edge.

If such a polyhedron can be inscribed in a sphere, then, looking along the plane of each face, it is clear that any face can be inscribed in a circle. To solve our problem, we must first check whether all polygons F_i are inscribed.

Next, we choose some edge E_k, and consider the two faces F_i and F_j that are bounded by this edge. Let O_1 and O_2 be the circumcenters of these two faces (figure 2). Then the lines through O_1, O_2 perpendicular to the planes of the two faces both lie on the perpendicular bisector plane of E_k, and so intersect. (Indeed, what would the two faces look like if they were parallel?) Designate by r_k the distance from the intersection point of these perpendiculars to one of the vertices of F_i or F_j. It's clear that the length of r_k does not depend on the choice of this vertex.

Problem 1. *Prove that polyhedron M can be inscribed if and only if all the polygons F_i are inscribed and $r_1 = r_2 = \dots = r_m$, where m is the number of edges of the polyhedron.*

This and similar tests have been known for a long time. In the beginning of the 20th century, it turned out that it's sometimes possible to prove that M cannot be inscribed in a sphere even without doing any computations.

The German mathematician E. Steinitz was the first to notice this fact. In 1927 he published an article in which the following theorem was proved.

Steinitz Theorem. *Assume that all the vertices of polyhedron M can be colored black and white so that*

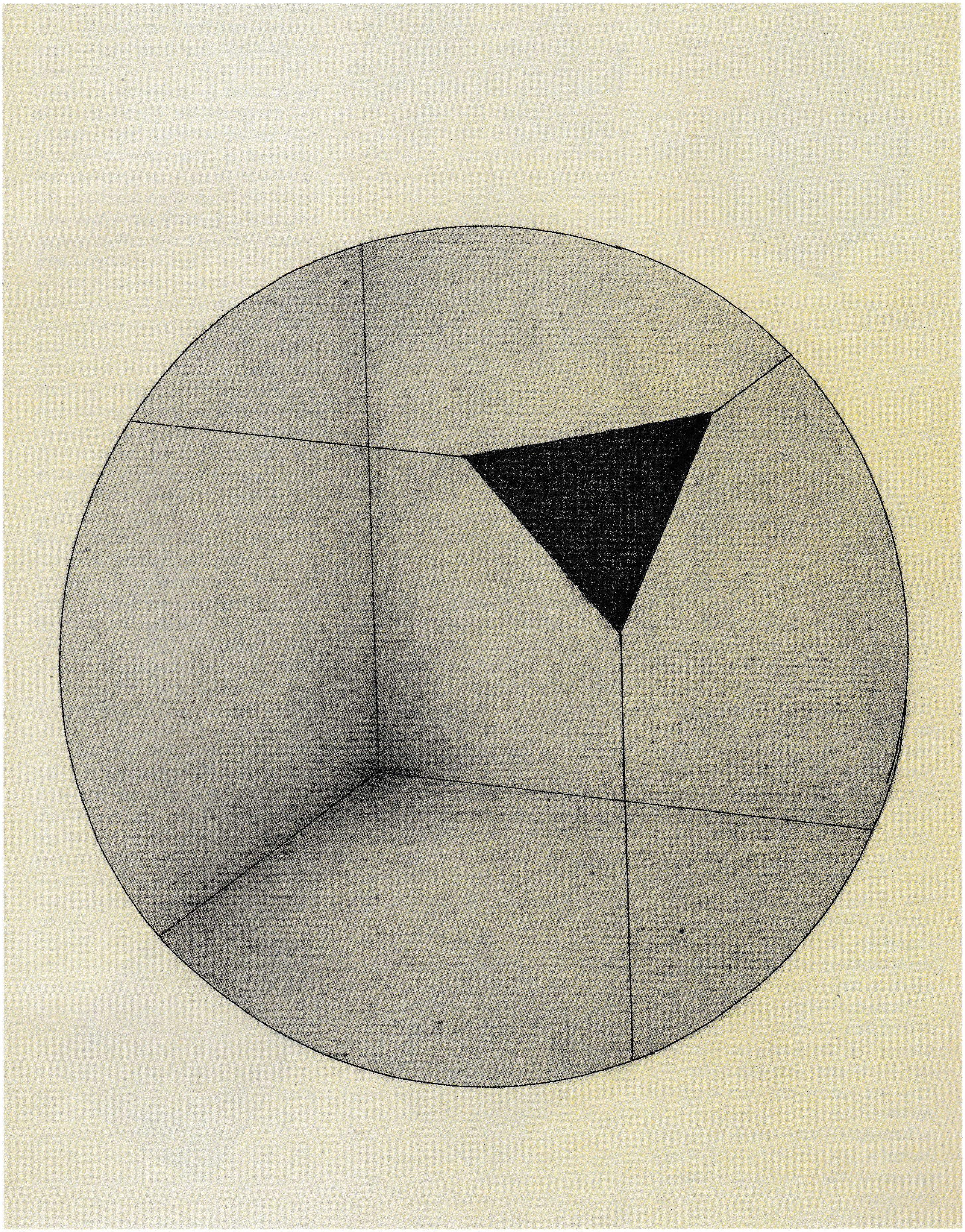
(1) no two black vertices are neighboring;

(2) the number of black vertices is greater than the number of the white ones.

Given these conditions, M cannot be inscribed in a sphere.

Before proving the theorem, we'll make a few remarks. Assume that

Art by Yuri Vaschenko



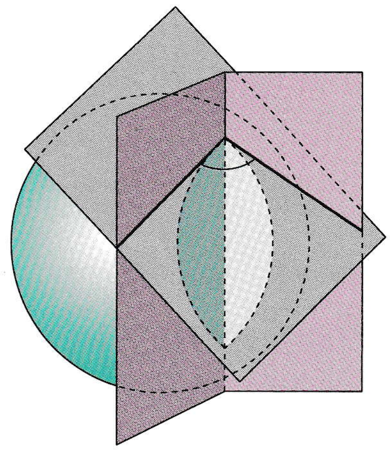


Figure 3

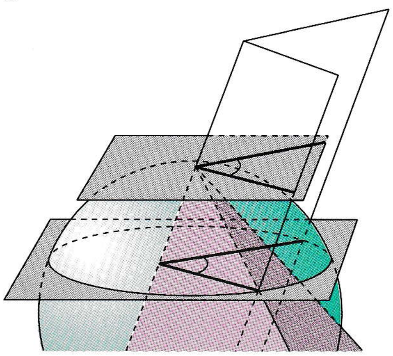


Figure 4

we are given a fixed sphere and a dihedral angle whose edge intersects this sphere. Draw the tangent plane to the sphere through one of the two intersection points (figure 3). The intersection of this tangent plane with the dihedral angle forms a plane angle. We will call this angle the plane angle of the given dihedral angle with respect to the given sphere, or simply the *relative angle of the given dihedral angle.* It's clear that the relative angle is independent of the choice of one of the two intersection points of the edge with the sphere. If the edge is tangent to the sphere, we set the relative angle equal to zero.

Now consider a convex polyhedral angle such that all its edges intersect the given sphere. The relative angles of its dihedral angles will be called *relative angles of the given polyhedral angle.*

Lemma 1. *If the vertex of a polyhedral angle with n faces lies on a sphere, and all its edges intersect the sphere, then the sum of its relative angles is* $\pi(n-2)$.

Proof. Draw the tangent plane through the polyhedral angle's vertex, and then draw a plane parallel to this tangent plane and such that it intersects the sphere and all edges of the given polyhedral angle (this is possible since all edges of the angle intersect the sphere). The intersection of the polyhedral angle with this plane is a polygon, and pairs of sides of this polygon are parallel to the sides of the relative angles we want to sum. So each relative angle is equal to an angle of the polygon, which makes their sum equal to $\pi(n-2)$.

Proof of Steinitz Theorem. Assume that polyhedron M is inscribed in a sphere and that its faces are colored as required by the theorem. We will show that this implies a contradiction involving certain relative angles. Designate the relative angle of the dihedral angle with edge E_i by γ_i. Let n_k be the number of edges that meet at vertex V_k. By Lemma 1, the sum of the relative angles at vertex V_k is $\pi(n_k-2)$. Set $\beta_i = \pi - \gamma_i$ and call β_i the exterior relative angle. Then $\gamma_i = \pi - \beta_i$. Since the sum of the interior relative angles of any polyhedral angle is $\pi(n_k-2)$, the sum of the exterior relative angles is 2π. Thus if the edges $P_i, P_j, \ldots, P_l$ meet at vertex V_k, then

$$\beta_i + \beta_j + \ldots + \beta_l = 2\pi.$$

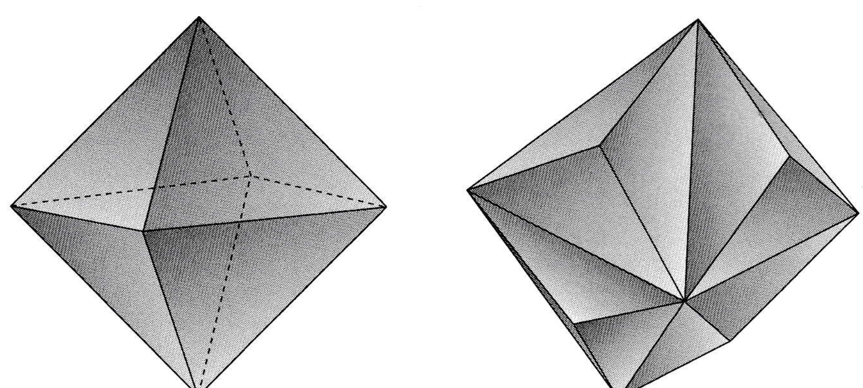
Let's write similar equations for each vertex of the polyhedron, multiply the equations for black vertices by -1, and sum all the equations obtained. Black vertices are more numerous; therefore, the right-hand side will be negative.

Consider the sum on the left-hand side. If the ith edge connects a black vertex with a white one, then the number β_i enters the left-hand side twice—once with a positive sign, and once with a negative sign, which gives a net contribution of 0 to the sum. If the edge connects two white vertices, then β_i enters the left-hand side with a positive sign both times. By our assumption, there are no edges with two black vertices; therefore, the sum on the left-hand side is not less than zero. Thus we have arrived at a contradiction, which proves that polyhedron M cannot be inscribed in a sphere.

Before analyzing the theorem proved—that is, before revealing its meaning—we want to demonstrate that there exist polyhedrons satisfying the conditions of the theorem. Consider an octahedron (figure 5a) and construct a regular triangular pyramid on each of its faces, using the face as a base. The altitude of the pyramid must be small enough that the dihedral angles at the base will be less than 25°. Consider the polyhedron M glued together from the original octahedron and eight newly constructed pyramids (figure 5b).

Problem 2. *Prove that polyhedron M is convex.*

We color all the vertices of M that are vertices of the original octahedron white, and all the other vertices black. It's clear that the conditions of the Steinitz Theorem are satisfied. Thus we have constructed a polyhedron that cannot be inscribed in a sphere, and this fact can



a b

Figure 5

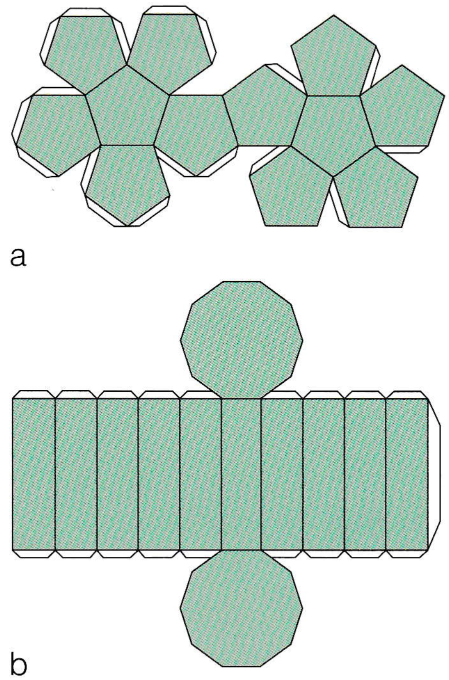
be verified without knowing the size and angles of the polyhedron; all we need is the *structure* of the polyhedron.

We now consider in more detail the concept of structure. In the plane, everything is very simple. To describe the structure of a convex polyhedron, it's sufficient to say that it has as many sides as it has vertices; every side is adjacent to two others, and every vertex is adjacent to two sides.

The situation in space is much more complicated. For example, the dodecahedron (figure 6a) and the decagonal prism (figure 6b) have the same number of faces (12), the same number of edges (30), and the same number of vertices (20), but their structure is different. To describe the structure of a polyhedron, we must specify not only the number of edges, faces, and vertices, but also how the faces are connected; that is, which faces are adjacent, which vertices are neighboring, and which faces meet at each vertex. Two polyhedrons have the same structure if they have the same number of vertices, edges, and faces and are composed of these elements identically.

Suppose that a convex polyhedron M is such that neither it nor any polyhedron of the same structure can be inscribed in a sphere. We call such a polyhedron *absolutely uninscribable*. The Steinitz theorem implies that any polyhedron satisfying conditions (1) and (2) is not only uninscribed, but also absolutely uninscribable. The polyhedron in figure 5b provides such an example. We can compare this result to the situation on a plane. Here, all polygons with n sides have the same structure, and there exist an inscribable polygon of that structure for any n. For this reason, it was believed, before Steinitz proved his theorem, that absolutely uninscribable polyhedrons do not exist. There were even quite plausible, though not quite complete, proofs of this "fact."

Now another problem presents itself: Find all absolutely uninscribable polyhedrons. More specifically, find necessary and sufficient condi-



a b

Figure 6

tions for a polyhedron to be absolutely uninscribable. We have the Steinitz Theorem, which gives sufficient conditions. Are these conditions also necessary? The following problems show that this is not the case.

Problem 3. *Let all the vertices of a polyhedron M be colored black and white so that*

(1) the number of white vertices is not greater than the number of black ones;

(2) no two black vertices are neighboring, and there exist two neighboring white vertices.

Prove that polyhedron M cannot be inscribed in a sphere.

Problem 4. *Construct a polyhedron satisfying the conditions of problem 3, but not satisfying the conditions of the Steinitz Theorem.*

It's important to note another fact.

Lemma 2. *If the same number of edges (say k) meet at every vertex of a polyhedron, then this polyhedron satisfies neither the conditions of the Steinitz Theorem nor the conditions of problem 3.*

Proof. Assume the opposite. Let m be the total number of edges of the polyhedron, p the number of black vertices, and g the number of white ones, and let $p \ge g$ (in the Steinitz Theorem, this inequality is strict). Then $2m = k(p + g)$; indeed, every edge has two endpoints, and k edges meet at every vertex. At least one of the endpoints of every edge is white—that is, $m \le kg$ (in the case of problem 3, this inequality is strict). On the other hand, $p + g \ge 2g$, and this inequality is strict under the conditions of the Steinitz theorem. Thus we obtain the inequalities

$$kg \le m \le kg.$$

If M satisfies the conditions of the Steinitz theorem or the conditions of problem 3, then one of these inequalities is strict, which is impossible. This contradiction proves the assertion of the lemma.

Now consider polyhedrons with three faces concurring at every vertex. In a certain sense, such polyhedrons are the most typical, but we know nothing about whether they can be inscribed.

Consider three planes that intersect in pairs. They either form a trihedral angle or the lines of their intersection are parallel. In the latter case, we say that the planes form an infinite trihedral angle. Let's introduce a sphere. Assume that all the edges of the trihedral angle intersect this sphere, but the vertex doesn't lie on it, or maybe the angle has no vertex at all. It turns out that if the vertex lies inside the sphere, the sum of the relative angles is greater than π; if it lies outside the sphere or the trihedral angle is infinite, the sum is less than π.

Indeed, consider the circles that are cut out by the faces of the trihedral angle on the sphere; the angles between these circles are exactly the relative angles of the trihedral angle. Take a point on the sphere that doesn't belong to any of these circles and construct stereographic projections of the circles from this point onto a plane. Three cases are possible (see figure 7a, 7b, and 7c). In the first case, the vertex is outside the sphere; in the second case it lies on the sphere; and in the third case it is

CONTINUED ON PAGE 46

What is thought?

"...thought is a change in the body"

—Leucippus and Democritus

by V. Meshcheryakov

IMAGINE YOURSELF INSIDE A CAT scan machine, a sophisticated device for studying the molecular structure of living creatures. You would be subjected to electromagnetic radiation with a frequency of about 10^7 Hz and a magnetic field of about 1 T (by way of comparison, the magnetic field of planet Earth is less than 10^{-4} T). And yet our brain doesn't feel a thing! The electromagnetic radiation from a personal computer may interfere with TV reception, yet it doesn't disturb you or your brain at all! The environment around us is jammed with electromagnetic waves transmitting radio and TV programs. Our towns are enmeshed by current-carrying (and field-generating!) wires; at home we're never more than a step away from an electrical appliance. Your brain is oblivious to it all.

These are all contemporary examples. Let's remember, however, plain old lightning, which is an electric discharge known to humans from time immemorial. Lightning generates electromagnetic waves that can be detected by a receiver in any frequency range. What about our thoughts? Are they affected? Well, aside from the fear we may experience, throughout all of recorded history lightning has not been found to affect us at all.

We may also recall visible light, as well as infrared and ultraviolet radiation. They are also electromagnetic waves. Most of this radiation reaches the brain—it doesn't just reach it, it penetrates, passes through, and permeates!

The interaction of an electromagnetic field with biological systems is determined primarily by two factors: the intensity and the frequency of the electromagnetic radiation. The search for the parameters of electromagnetic waves capable of controlling but not disturbing the state of cells in a biological system (or, as they say, modifying the informational structure of a living organism) led to low-intensity waves of the millimeter range. However, even in this case no evidence was found that electromagnetic fields affected the thought process.

Perhaps it would be better to try to explain the influence of thought on the electromagnetic field? There are indeed devices (one of them is called "SQUID") that measure weak, brain-generated magnetic fields with strengths of only 10^{-13} T, but they don't have any connection with thoughts.

What can we conclude from these negative results? Maybe electromagnetic interaction is not the process responsible for thought generation. There are arguments in favor of this hypothesis. It's corroborated not only by a huge amount of experimental data but by common sense as well. Judge for yourself: Would the Creator have allowed thought to be controlled by an electromagnetic mechanism, knowing that an immature humanity would extract the secrets of electromagnetism in the 20th century? Imagine radio frequency-control of the brain as a weapon at a dictator's disposal! This possibility is much more dangerous than radiation damage to an organism characterized by a breaking of intermolecular bonds and ionization of atoms. Keep in mind that the rate of radioactive decay decreases with time and that sources of intense radiation are rare, expensive, and hard to get a hold of, while radio waves are cheap and readily available everywhere.

Well, at present only two interdependent states of the human mind and an electromagnetic field seem to be possible. Either there is no interaction for some range of electromagnetic parameters, or this interaction

Art by Pavel Chernusky

exists for other parameters and its effect is destructive to the brain. This hypothesis doesn't contradict modern views of the interaction between fields and matter, even though no theory of brain–field interaction exists. Although the theory of the electromagnetic field was elaborated long ago, in the 19th century, we still don't know what "mind" is.

Maybe this isn't a pressing problem? Well, history can show us many scientists who have worked on seemingly "nonpressing" problems. Examples are the search for extraterrestrial intelligence; parapsychology; telekinesis; travel into the past, future, and even "parallel" time; teleportation; and much more.

We might recall the experiments of the dye maker Steven Gray and the priest Granville Wheeler. At the beginning of the 18th century they constructed current-carrying lines by rubbing glass rods and channeling the charge through threads. Was this research a pressing matter in the period preceding the development of the theory of electricity and its practical application by more than 100 years? Perhaps not. But among all of Gray's papers, including those devoted to contemporary problems in optics, astronomy, and meteorology, only the paper on electricity was published in the *Proceedings of the Royal Society* during the very time when the great Isaac Newton was its president.

How far could this outstanding personality see, this man who declared that he did not invent hypotheses! Perhaps that's why he framed his view of biological structures in the form of a question: Does not Living Motion occur by means of the oscillation of the Living Medium excited in the Brain by the force of Will and transmitted through continuous, transparent, and homogeneous nerve capillaries to the muscles, thereby contracting and distending them?

So, is the problem of elucidating the mechanism of the brain's activity important to humanity? The United States Congress declared the 1990s "The Decade of the Brain," because "fundamental discoveries about the organization of the brain at the molecular and cellular levels have laid the foundation for understanding the mechanisms of the mental activity of human beings."

Virtual reality

At present, one of the main indications of the importance of brain research is the development of new technologies. It's confirmed by the appearance of the notion of virtual reality.

As early as 1959, the renowned American physicist Richard Phillips Feynman (1918–1988) declared that in the future, after mastering operations with individual atoms, humanity will have the technology to synthesize everything. An unexpected start in the practical realization of this fantastic idea occurred in 1981, when the physicists Gerd Binning and Heinrich Rohrer of the Swiss branch of IBM constructed a device that incorporated a controllable single-atom contact between two solid objects. The first application of this device was scanning the surfaces of solid objects to examine their structure. Improvements made in the following decade led to the appearance of devices that could build molecular aggregates from individual atoms according to a chosen plan. In their Nobel Prize lecture (in 1987) Binning and Rohrer proclaimed: "At last it has become possible to affect individual atoms and to modify individual molecules."

This was the birth of a new technology known now as *nanotechnology*. Why "nano"? The typical distance between atoms in the condensed state of matter is about 10^{-10} m, while the characteristic linear size of a typical molecular structure composed of hundreds and thousands of atoms is about 10^{-9} m, or 1 nanometer.

The importance of Binning and Rohrer's work was immediately recognized. Not only did they receive the Nobel Prize, but an avalanche of scientific research followed. The number of fields affected was unprecedented: physics, electronics, biology, medicine, gerontology, cybernetics, ecology, design, art, religion, philosophy, science fiction... We're now living in a period many call the "nano-industrial revolution."

Given the prospects for the future development of nanotechnology, Feynman's published remarks in 1984 were timely. He was speculating about the possibility of creating a quantum-mechanical computer in which the role of the logic elements would be played by individual atoms. Here's how things actually played out.

In the 1980s computers could write data with an information density of about 1 bit per 10^{11} atoms. For a typical atomic volume of 10^{-29} m^3, a one-bit memory element occupies 10^{-18} m^3, which corresponds to a linear size per bit of about 10^{-6} m = 1 μ. This is why computer chips are considered microelements.

The size of the logic elements in processors has approached the nanometer range, but hasn't quite reached it yet. Specialists say that production of nanoelements—that is, chips with an element density of about 1 bit per 10^2–10^3 atoms—will begin in the next decade, despite the fact there are problems yet to be solved.

One of the problems is that modern circuit design is planar (the elements are located on a plane), and switching to three-dimensional structures will require industrial development and customization of nanotechnology methods. The prospects of developing chips with an element density of 1 atom per bit are still hazy, but the general tendency is quite clear: continual miniaturization of the elements (memory chips, processors, controllers) in order to develop a technology of so-called *terabit crystals*. Even now numerous projects are under way in the US and Japan, costing hundreds of millions of dollars or more, involving dozens of private firms and government organizations in developing these nanotechnologies.

It was on this road that the concept of "downloading" was born. We can readily grasp its essential meaning by way of nanotechnology. Consider the human brain, which according to neurophysiologists contains about 10^5 neurons per cubic millimeter of brain tissue. Taking into consideration that the cerebral cortex occupies a volume of about 10^{-3} m^3, we get the total number of neurons. This is an amazingly large number: 10^{11}.

Supposing that 1 bit of human memory requires at least 1 neuron, the total informational content of the brain can be described by the number 10^{11}. Imagine we could scan the neuron state one cell at a time to obtain the physical characteristics of every neuron in the brain. Then we "download" this "file" into terabit chips with a memory capacity of about 10^{12} bits. This "smart" chip could be incorporated in a personal computer or the controller of a robotic policeman. We could also send it off to wander the World Wide Web. The scanned information carries some personal features with it, so it's not baseless to think we could equip this individual information flow with the ability to make its own decisions—that is, to generate new information, which means to think.

So our hypothetical "something" obtained by scanning the brain and downloading the information into a computer system is known as the state of virtual reality and determines the image that arises from the achievements of nanotechnology.

Negative reality

Now let's consider the problem from another direction, leaving aside for now the puzzle of brain scanning.

Imagine you're a design engineer with all possible combinations of atoms at your disposal. You could construct a ball, a cube, a ring. These shapes you could pull out of your own memory. However, if you tried to reproduce the shape of a snowflake, you'd at least need a detailed photo, and a diagram showing the arrangement of all the atoms would be even better. These imaginary experiments deal with what we take to be nonliving matter. Let's move on and try to construct a DNA molecule, which is only one of the numerous components of the neurons in the brain. This is a very simple problem—just get the map of the molecular structure. Unpack the set of atoms you bought at the atom store and start building!

But first let's look at the maps or handbooks to find out which sort of molecule we'll end up with—living or nonliving. Unfortunately, the research journals don't have the answer to this one, nor to the question of how to synthesize a neuron or, taking it several steps further, a human brain. At present, science doesn't know whether the result will be living or nonliving, a thinking object or just a pile of atoms.

> *"... there is no reasonable basis to consider that intelligence is coupled with the body."*
> —Aristotle

We've reached the point where so-called "traditional science" lags behind progressive trends. Its experimental methods still can't obtain data on the effects of external forces on human intellectual activity—or vice versa, on the effects of human thoughts on the measuring devices. The analytical part of science still can't tell the difference between live and dead groups of molecules, between thinking and unthinking molecular aggregates. And the understanding we seek isn't verbal—who can talk about it longest and cleverest; we're looking for a constructive understanding—an understanding that could provide the basis for building an artificial intellect. Here the criterion is straightforward. If you know how electromagnetic waves or thoughts are generated, you should also know how to make a radio or an artificial brain.

At the same time, experimental physics and biology can manipulate individual atoms and molecules and construct atomic clusters that are the elementary building blocks of living and dead matter, of thinking and nonthinking organisms. (For now this is possible only in laboratories. So far only nanotechnological inspection of CD and DVD disks has reached the factory floor.) So the search goes on. Who's taking part? Theologians, biologists, physicists, and, of course, science fiction writers.

On this thorny road, we should note that no highly refined reproduction of an original thinking specimen produces a perfect identical copy or even explains why the copies may be different. The physical characteristics of the atoms are scanned by measuring devices. In other words, only measurable parameters can be scanned. However, we have not yet invented a device capable of reading thoughts. Perhaps the main reason is that a thought cannot be expressed by a static set of atoms, ions, and molecules. A thought is a process, a movement. We know something is moving, but we don't know how.

So, our attempt to create an artificial intellect must necessarily be related to discovering how the brain functions. This wonderful "device" was given to us by nature, or by God. But our search must be based on modern views of the world, on the thousands of years of human experience that speaks of the continuity of knowledge and rejects, if not the most intellectual and experimental attempts to "jump" to metaphysical reality, then at least the expectation that they will lead to any constructive consequences—that is, to actual objects of widespread use.

However, orthodox science sticks to its principles and keeps mum. The reason is simple. On the one hand, we have an excellent model of the brain based on a complex integral system composed of a finite number of elements working on the basis of electric charge transfer or excitation in the electron subsystem, which results in the transmission and reception of information coded in terms of electromagnetic field parameters. On the other hand, more than a hundred years of practical experience in using such systems has not demonstrated, either accidentally or intentionally, the effect of electromagnetic radiation on mental processes.

In other words, on the one hand we have a computer as the most perfect model of the brain, and on the other hand it's always possible to find a set of parameters of an external electromagnetic field that will not destroy but will modify, in a given way, the functioning of the resistors, capacitors, transistors, and all the other components of this computer.

So the question is this: Can a computer, now or in the future—even a supercomputer using terachips—model the human brain even approximately? Most likely, no.

Prospective reality

"So," you may ask, "are there really no other models of mental activity?" It's a complicated situation. Imagine someone who doesn't know physics but wants to understand how a car works. What should that person examine? The wheels, carburetor, motor, or gasoline tank? After all, one might guess that it is the burning of gasoline that moves the car and declare that combustion is the principle of car performance one is looking for. Although some people will be satisfied with this explanation, there is a yawning chasm between this conclusion and, say, the Carnot cycle, which provides data about the engine's efficiency. The same is true with respect to the activity of a neuron. Many features of neurons are known: structure, chemical composition, the character of diffusion fluxes, thermal conditions, distribution of electric and magnetic fields, electrical and thermal conductance, and so on. Nevertheless, there is no answer to the question: What physical process underlies mental activity?

Maybe we need to feel our way to a qualitatively new, non-electromagnetic model of the brain, a model based not on the transfer of electrons or other charged particles, or on the excited states of such particles?

From the outset we reject any attempts to bring in supernatural concepts such as a "biological field." Recall that the transmission of information (to say nothing of its generation) requires a material carrier. We don't have a lot of choices here. Only two suitable substances are known—the electromagnetic field and a substance composed of atoms, which in turn consist of protons and neutrons (forming the nuclei) and electrons located near the nuclei; the electrons actually occupy almost the entire volume of the atom. The protons and electrons are electrically charged, so an arbitrary piece of a substance is united by the forces of interatomic bonds, which are electromagnetic in nature. However, this piece of matter is "electromagnetically neutral" in equilibrium.

Now let's think about the following problem: How can we, using a set of neutral atoms involved in a complex electromagnetic interaction, produce and transmit information without disturbing the electric neutrality of this set? The answer is simple. We'll push the nearest group of atoms. Since the atoms are held together by interatomic forces, and the piece of matter as a whole is a more or less elastic medium, the perturbation applied to the atomic equilibrium position will spread along the piece. If we act on one side of the piece with a particular force pattern, we can produce a certain signal of the same nature on the opposite side.

As we said earlier, we can place ourselves in a CAT scan machine to examine the effect of the electromagnetic field on the brain's functioning. But where should we place the head to detect the possible effects of mechanical forces? The simplest experiment requires no devices at all, because elastic waves exist everywhere. Mechanical watches are ticking, voices are heard from the TV, the telephone is ringing—all these events occur in your home. In the open air we hear the rustle of foliage, the slamming of doors, the squeal of brakes, the roar of airplane engines. In addition, there are natural events that are quite noisy: thunderstorms, thundering surf, the rumble of an earthquake. Do these events affect mental activity? Basically, no. Indeed, how could stable thought formation occur in a brain that functions by means of elastic waves under conditions of complete sonic anarchy, created by both Nature and bustling Humanity? The Creator took care to screen the brain from its noisy surroundings.

We can do our own experiments. Take a sound generator (usually available in a school lab), connect it to headphones, and try to determine the upper and lower boundaries of perceptible sound frequencies. You'll have no problem with the lower boundary. In the frequency range below 20 Hz the membranes of the headphones produce clicks instead of a clear tone. The value 20 ± 10 Hz is quite adequate for our purpose.

In contrast, it's not easy to estimate the upper boundary. When the frequency of sound exceeds 18 kHz (the region of the upper limit of human hearing), it's impossible to detect changes in the sound intensity with the accuracy achieved for the lower boundary. It looks like we still "hear" (perceive) the sound, but this is only an auditory illusion. In reality, it's impossible to fix the exact frequency of the attenuation of auditory sensitivity. Such experiments should not be repeated too frequently, lest the experimentalist get a headache. It's quite sufficient to see for yourself that such an auditory illusion does exist.

It's tempting, using the results of such psychophysical experiments, to admit the existence of nondestructive interaction of elastic waves with the brain. However, a new problem arises at this point. Sound waves of about 10^4 Hz have a wavelength of about one centimeter. When spreading within the brain, such a wave induces oscillations of huge atomic aggregates composed of 10^{23} atoms. This means that if the element of human memory is not an atom or nanochip, but a rather "large" microchip containing 10^{11} atoms, the elastic waves simultaneously displace 10^{12} "memory cells." It's clear that individual control of any of these 10^{12} micrometer-sized cells by an elastic wave with a wavelength of one centimeter is impossible, because such a long wave cannot change the state of any cell without disturbing its near and distant neighbors. In the same way, the surf simultaneously drenches all objects located at equal distances from its crest. In other words, if elastic waves really participate in mental activity, their wavelengths must not be greater than one nanometer, which is a typical size of the molecular constructions in the brain (for example, the DNA molecule is 2 nm in diameter). To produce elastic oscillations with such wavelengths, one needs frequencies higher than a few hundred gigahertz.

Eureka! Something interesting is lurking here. It's noteworthy that this frequency range of elastic oscillations is not reached in experimental physics. We have no method of regularly producing hypersound of ultrahigh frequencies with controllable parameters, which is a prerequisite for conducting a physical experiment. However, the lack of experimental data has never been an invincible barrier for inquisitive humanity in its drive to understand nature. For example, the physics of atomic groups was understood at the time of Maxwell and Boltzmann, although experimental proof of the existence of atoms was obtained only after the lifetimes of these scientific giants.

At present, dozens of mechanisms have been invented for storing and transmitting data in biological and inorganic molecular structures. These include quasi-elastic models. Let's briefly consider the incorrect (according to the consensus) but still interesting and widely discussed model, suggested in the 1970s by the outstanding British physicist Herbert Froelich. To explain the mode of data transmission in biological systems, he suggested a mechanism whereby elastic oscillations at about 10^{12} Hz are generated in a medium consisting of charged particles. However, in this frequency range the oscillations have a quantum nature. This is an unfavorable feature of the model, because in order to control the transition of a neuron or its parts from one state to another, they must remain in the initial state for a sufficiently long period. This condition is not met for oscillatory states with a lifetime of about 10^{-12} s. Therefore, Froelich advanced the concept of so-called *coherent excitation*, which is a classical wave train composed of quantum states. This wave train propagates with the speed of ordinary sound and (more importantly) can change the elastic state of a neuron for a long time. The details of Froelich's model were not elaborated, because it was soon established that the formation of such coherent trains is impossible. Nevertheless, this model is very attractive conceptually, because it replaced conventional electromagnetic memory cells with one based on deformation.

> *"... thinking is a most mysterious process."*
> —Sextus Empiricus

In this context, the problem of the deformation of biological objects is particularly interesting. As early as 1678, the Dutch scientist Anton van Leuwenhoek (1632–1723) noted that when he was seriously ill, the erythrocytes of his blood looked rigid and nonpliable, and that they became soft and elastic when he got better. Three centuries have passed, but we know only that elasticity (that is, the capacity to reversibly change shape under the action of external forces) is an intrinsic property of all biological objects, neurons included, and not a feature peculiar to erythrocytes. Experimental and theoretical attempts to reveal the mechanism of this elasticity remain unsuccessful.

Summing up our consideration of prospective reality, let's consider a question: Could it be that the lack of both an experimental technology for studying super-extra-ultrasound and a theory of the deformation of biological objects is the "yellow brick road" showing us the way to understanding how the human brain functions? In addition, current research into the response of nanometer molecular aggregates to the action of extraneous forces—that is, the study of the elastic properties of nanocrystals—is one of the most intensively developed branches of nanotechnology. Perhaps it will be the analysis of force effects on molecules, atomic clusters, or nanocrystals that provides the answer to the question "What is thought?" and consequently leads to the creation of artificial intelligence?

By way of conclusion (or, if you like, as a guide to further exploration of the mystery of the brain), let's recall what one of the founders of biophysics, Emile Du Bois-Reymond (1818–1896), said: "There are no forces affecting particles in a living organism that do not operate outside the organism as well." ◘

Quantum on the philosophy of science and mental activity:

M. Berkinblit and E. Glagoleva, "Mathematics in Living Organisms," November/December 1992, pp. 34–39.

A. Sakharov, "Does Elementary Length Exist?" May/June 1997, pp. 14–20.

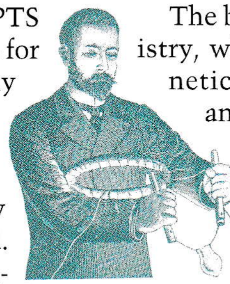
A. Kuzin, "The Anthropic Principle," January/February 1999, pp. 4–9.

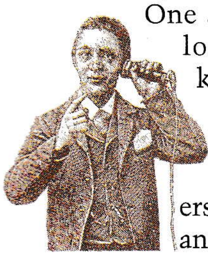
J. Raskin, "The Toy That Drove the Universe," November/December 1999, pp. 49–50.

V. F. Yakovlev, "Physics for Fools," November/December 1990, pp. 17–19.

Matter and m

FROM THE FIRST ATTEMPTS centuries ago to use magnets for our own purposes, primarily navigation, we have made great progress in understanding the subtle processes occurring in the objects around us when they are placed in a magnetic field. One amazing discovery followed another, and this knowledge was applied in creating electrical machines and transformers, audio recorders and players, communication devices, and computer technology.





"Magnetic phenomena are produced exclusively by electricity, and there is no other difference between two magnetic poles than their position relative to the electric currents that constitute the magnet."
—Andre Marie Ampere

The birth of magnetochemistry, which studies the magnetic properties of matter and their relation to molecular structure; the application of magnetostriction (the capacity of bodies to change their shape and size when magnetized) for generating ultrasound waves; the use of the ferromagnetic effect as a tool for assessing the quality of semiconductors; the production of ferrites as an alternative to metal magnets; the discovery that superconductivity and magnetism are interrelated—these were the glorious achievements of the 20th century in the study of magnetic phenomena.

However, that century did not provide solutions to all the problems involving magnetism—much is left for this new century! You may find yourself investigating substances with as yet unknown magnetic properties and using them to create high-capacity information storage devices that are very small and very reliable. And how many interesting and even mysterious magnetic effects can be observed in living organisms!

Don't forget that even poets have been enchanted by the phenomenon of magnetism, and in the French language the very word "magnet" derives from the verb "to love." And if this theme attracts you like a magnet—success is assured!

Problems and questions

1. Why do vertical steel security bars on windows eventually become magnetized? Which end of the vertical bar is the north pole and which one the south pole?

2. Are permanent magnets really permanent?

3. Can an iron sphere be magnetized?

4. Why isn't a magnetic crane used to transfer the red-hot items on a rolling mill at a steel plant?

5. A cuvette containing a copper sulfate solution is placed between the poles of a strong electromagnet (the surface of the solution is perpendicular to the magnetic field). A copper electrode is immersed into the solution at the center of the cuvette and connected to the positive terminal of a battery, while the negative terminal is connected to a copper ring immersed into the solution along the perimeter of the cuvette. What will happen when the circuit is closed?

6. A long, thin uncharged bar made of a nonmagnetic substance moves with a constant speed perpendicular to the lines of a magnetic field, as shown in figure 1. A potential difference appears between the faces of the bar. Why?

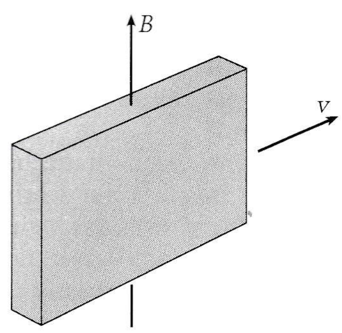


Figure 1

7. Is it possible to block an external magnetic field by means of a ferromagnetic screen, just as one blocks an electrostatic field?

magnetism

8. Why are oscillations of a compass needle damped more rapidly if the case of the device is made of brass rather than plastic?

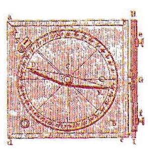
9. Why is the core of a transformer composed of several individual plates?

10. Does the inductance in a coil with an iron core depend on the current flowing in it?

11. How will the magnetic field generated by a current-carrying coil vary if a core is introduced into it? Consider cores made of (a) iron, (b) aluminum, (c) copper.

12. The magnetic permeability of several liquids was being investigated. The liquids were poured one after the other into connected vessels, one of which was placed between the poles of a strong electromagnet. Why do some liquids rise in this vessel while others sink?

13. Why is the flame of a candle placed between the poles of a magnet expelled outward?



14. What will occur in a ring when a magnet is introduced into it if the ring is made of (a) a dielectric, (b) a conductor, (c) a superconductor?

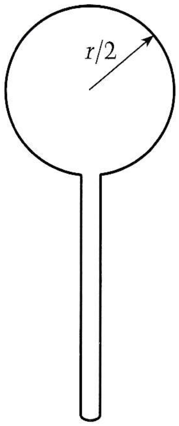
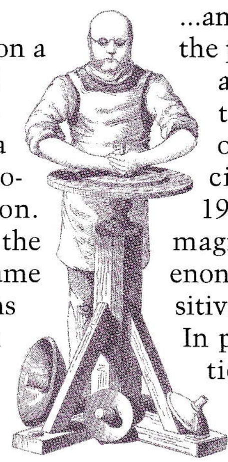


Figure 2

15. An electric current flows in a superconducting ring of radius r. The ring is deformed as shown in figure 2. How will the magnetic field change at the center of the small ring as compared to the field at the center of the original ring?

Microexperiment

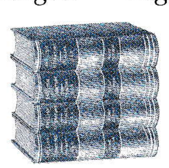
Suspend a thin iron nail on a light incombustible thread such that the nail is deflected into the flame of a burner when a strong electromagnet nearby is turned on. You'll see that after a while, the nail jumps away from the flame as if it got "burnt" and returns to its original position. Then the sequence begins again and keeps repeating itself. What causes the nail's periodic motion?



It's interesting that...

...the first comprehensive work on the properties and practical application of magnets, in which the magnetic stone was described and instructions were given on how to find the magnetic poles and magnetize iron needles, was the manuscript "Message on the Magnet, from Pierre de Maricourt, known as Peregrine, to the knight Siguerre de Foucaucourt," which appeared in France in 1269.

...as early as the 16th century, William Gilbert hypothesized that there should be north-seeking and south-seeking "magnetic charges." This idea was developed by Charles Coulomb (1736–1806), who suggested the law of interaction of these "charges," which coincided exactly with the known law of electrostatic interaction. Eventually Andre Marie Ampere (1775–1836) made the hypothesis of specific magnetic charges superfluous by explaining all electromagnetic phenomena on the basis of elementary electric currents.



...any rotating body, including the planets, must at least have a weak magnetization. Attempts were made by the outstanding Russian physicist Pyotr Lebedev (1866–1912) to detect this rotation magnetization. The phenomenon was later observed by sensitive and sophisticated devices. In particular, the magnetization of a rod rotating about its longitudinal axis was measured.

...the total magnetic permeability of an alloy of diamagnetic gold and paramagnetic platinum is smaller by two orders of magnitude compared to ordinary nonferromagnetic substances.

...some alloys of paramagnetic and diamagnetic metals—for example, the so-called Heusler alloy, composed of copper, manganese, and aluminum—are almost equal to iron in their magnetic properties. At present, useful magnets are even obtained from organic materials.

...new discoveries in magnetism make it possible to produce memory chips with superdense data recording capabilities, where an area the size of a thumbnail can hold tens of thousands of copies of Homer's *Odyssey*.

ANSWERS & SOLUTIONS ON PAGE 53

Where is last year's winter?

by A. Stasenko

"*Large ceramic vessels for food storage were buried at a depth greater than a person's height to keep them cool (found at Knossos, Troy, and Tiryns).*"

—Encyclopedia of Antiquity

THE GREAT LUCRETIUS (98–55 B.C.) offered a curious explanation for the half-year periodicity of temperature oscillations at some depth below the surface of the Earth: "Let us now consider *why it is that well water is warmer in winter and cooler in summer*. This happens because in summer the earth is relaxed by the warmth and any particles it may contain of its own heat are dispersed into the air. ...Conversely, when all the earth is compressed by cold and contracts and virtually congeals, it naturally happens that in contracting it squeezes out any heat it may contain into the wells."[1]

Since that time physics has developed concepts more rigorous than "heat particles" and "cold pressure." Physicists prefer to speak in terms of density ρ and specific heat c. These concepts will come in hand here. There's another physical quantity that we'll use as well: thermal conductivity. Let's take a closer look at this useful notion.

Imagine that you need to know how much heat is dissipated to the environment every second through each square meter of the walls in your house. This is an important number that will help you calculate how much coal, oil, or electricity you'll need to keep your house warm. Let the temperature of the

Art by Yuri Vaschenko

[1]Lucretius, *The Nature of the Universe* (trans. Ronald Latham), Baltimore: Penguin Books, 1962.

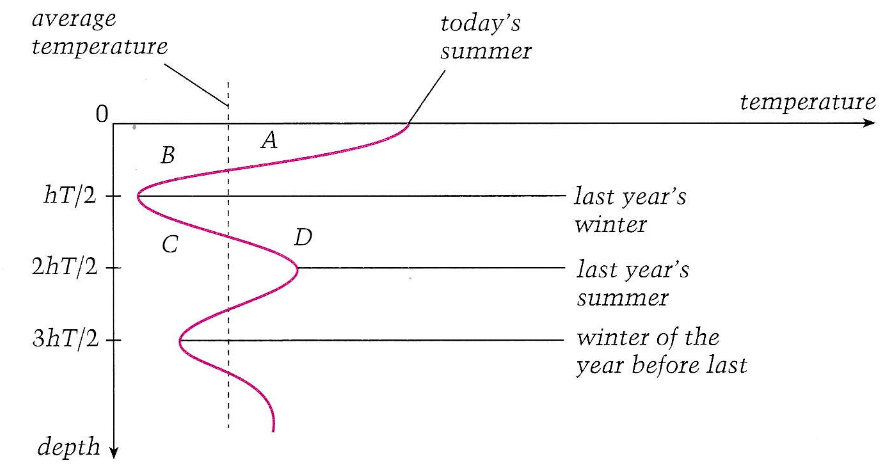


Figure 1

internal and external surfaces of the walls be T_{int} and T_{ext}, and let the thickness of the walls be h. The density of heat flow q (in units of $J/(m^2 \cdot s)$) is given by

$$q = \lambda \frac{T_{int} - T_{ext}}{h}. \qquad (1)$$

This formula introduces the coefficient of thermal conductivity λ, which depends on neither the temperature of the wall's sides nor the wall's thickness. The coefficient characterizes only the physical properties of the material the wall is made of, so a builder can look it up in a reference book.

It's easy to find the units used to express the coefficient of thermal conductivity from equation (1): It is $[\lambda] = J/(m \cdot s \cdot K)$. It's important that the time unit be incorporated in this expression—this makes it possible to write a combination of ρ, c, h, and λ that is expressed in units of time:

$$\tau \sim \frac{\rho c}{\lambda} h^2.$$

What can we do with this value? Well, it helps to know the depth $h_{T/2}$ reached by the temperature that was at the surface a half-year ago. From this we get

$$h_{T/2} \sim \sqrt{\frac{\lambda}{\rho c} \frac{T}{2}}. \qquad (2)$$

Of course, there are many varieties of "earth" or "soil." For example, loam, sandy soil, and granite differ widely in their densities, specific heats, and thermal conductivities. However, an "average" value like $2 \cdot 10^{-7}$ m²/s can be assumed for the expression

$$a = \frac{\lambda}{\rho c},$$

which is known as the *temperature conductivity coefficient*.

How many seconds are there in a year? Let's do the arithmetic:

$$T = 3{,}600 \text{ s/hour} \cdot 24 \text{ hours/day} \cdot 365 \text{ days} \cong 3 \cdot 10^7 \text{ s}.$$

Plugging this numerical value into the formula for $h_{T/2}$, we get

$$h_{T/2} \sim \sqrt{\frac{aT}{2}} = \sqrt{\frac{2 \cdot 10^{-7} \text{ m}^2/\text{s} \cdot 3 \cdot 10^7 \text{ s}}{2}} \sim 2 \text{ m}.$$

Of course, this is only a rough, order-of-magnitude estimate. But even this approximation explains why the ancient Greeks buried their amphoras deeper than the height of a person. This just happens to be the depth reached by the thermal wave of the previous winter at the time the soil is heated by the summer Sun.

Figure 1 qualitatively shows the "instantaneous" distribution of temperature at different depths. Why do the temperature oscillations fade at the deeper layers? This is due to the same phenomenon of thermal conductivity described above. On the one hand, it allows the thermal energy to reach into the ground; on the other hand, it helps disperse and flatten the crests and depressions in the temperature curve. In particular, equation (1) says that the thermal energy flows downward along the "slope" AB and rises upward along the slope DC.

One can plot a similar instantaneous distribution of temperature for sound waves in the air. Fortunately, the thermal conductivity of the air doesn't play a significant role for the frequencies used when we talk or listen to music, because the successive compressions and relaxations occur so quickly that thermal conductivity has no time to damp the crests and depressions of the temperature. As physicists would say, the *dispersion* and *attenuation* (fading) of the acoustic wave are insignificant. But it's unlikely the ancient Greeks thought about these problems when they buried their amphoras. ■

Quantum on heat transfer and heat exchange:

A. Kikoyin, "Temperature, Heat, and Thermometers," May 1990, pp. 16–21.

I. I. Mazin, "An Invitation to the Bathhouse," September/October 1990, pp. 20–22.

A. V. Byalko, "Atmospherics," March/April 1991, pp. 12–17.

A. Stasenko, "Baby, It's Cold There!" March/April 1992, pp. 12–19.

Y. Amstislavsky, "Less Heat and More Light," November/December 1995, pp. 4–8.

V. Lange and T. Lange, "The Power of the Sun and You," July/August 1996, pp. 16–18.

A. Byalko, "Hands-on (or -off?) Science," November/December 1997, pp. 4–8.

A. Eisenkraft and L. D. Kirkpatrick, "Elephant Ears," November/December 1997, pp. 30–32, and May/June 1998, p. 30.

V. Milman, "Why Is a Burnt Match Bent?" November/December 1998, pp. 40–42.

V. Novozhilov, "Fertilizer with a Bang," September/October 2000, pp. 8–12.

Exploring remainders and congruences

by A. Yegorov

ON THE THIRD OF SEPTEMber in the year 2000, I wanted to find out which day of the week December 20, 2001, would be. I had no calendar at hand, so I had to do some calculations. I knew that September 3 was a Sunday. There are $27 + 31 + 30 + 20 + 365 = 473$ days from September 3, 2000, to December 20, 2001, which makes 67 full weeks and 4 days ($473 = 67 \cdot 7 + 4$). Therefore, December 20, 2001, must be a Thursday.

A student squared a multidigit number and obtained 46,991,075. The teacher looked at his answer and immediately said that the answer was incorrect. How did the teacher know?

Exercise 1. Can the square of an integer number end in the digits 75?

We will see further in this article that the solution to this simple problem and many others is based on divisibility considerations. First, recall the concept of division with remainder.

Division with remainder

Definition. *To divide a natural number a by a natural number b means to write a as $a = qb + r$, where q and r are nonnegative integers and $r < b$. The number q is called the quotient and r the remainder upon division of a by b.*

In practice, division with remainder can be done using the algorithm called *long division*. For example,

$$\begin{array}{r} 12 \\ 14\overline{)179} \\ \underline{14} \\ 39 \\ \underline{28} \\ 11 \end{array}$$

So $179 = 12 \cdot 14 + 11$. Here the remainder is 11, and the quotient is $12 \div 179 = 12 \cdot 14 + 11$.

Notice that we don't require in this definition that a be greater than b. For example, we can divide 5 by 7 by writing $5 = 0 \cdot 7 + 5$. Generally, if $a < b$, then $a = 0 \cdot b + a$. Thus, in this case, $q = 0$ and $r = a$.

Note. We can define division with remainder for any integer a by any integer $b \neq 0$ as follows. To divide a by b with a remainder means to represent a in the form $a = qb + r$, where q is an integer and $0 \le r < |b|$. For example, for $a = -15$ and $b = 7$, we have $-15 = (-3) \cdot 7 + 6$. For $a = -224$ and $b = -9$, we have $-224 = 25 \cdot (-9) + 1$.

If the remainder is zero, we say that a is *divisible* by b.

The following simple fact is extremely important: *If a and b are divisible by c, then the number ka + lb is divisible by c for any integers k and l.*

Problem 1. Find the integer $d > 1$, such that for any integer n, the numbers $7n + 1$ and $8n + 3$ are both divisible by d.

Solution. Since $7(8n + 3) - 8(7n + 1) = 13$, the number 13 is divisible by d. Now, since $d \neq 1$ and 13 is a prime, we have $d = 13$.

Exercises

2. Divide the following with a remainder: (i) 1931 by 17; (ii) -295 by 31; (iii) -1005 by -98.

3. If the number $17x + 3y$ is divisible by 61, prove that $8x + 5y$ is also divisible by 61 (x and y are integers).

4. Determine the remainders upon division of (i) n by $n - 1$ and by $n - 2$; (ii) $n^2 + n + 1$ by $n + 1$ and by $n + 2$; (iii) $n^4 + 1$ by $n + 3$ (where $n \ge 180$).

5. Find all integers n for which the following numbers are integers: (i) $(n^2 + 1)/(n - 1)$; (ii) $(n^5 + 3)/(n^2 + 1)$.

Congruences

From this point on, we'll assume that all the numbers we're dealing with are integers. Consider one more problem.

Problem 2. Determine the last digit of the number 2^{999}.

Solution. Let's write out the sequence of powers of two:

$$2, 4, 8, 16, 32, 64, \ldots$$

We see that the last digits of the numbers in this sequence repeat in a cycle of 4. Thus the last digit of the number 2^n depends only on the remainder upon division of n by 4. Since $999 = 996 + 3 = 4 \cdot 249 + 3$, the answer is 8.

In this example, the set of exponents of the powers of two is divided

Art by Yuri Vaschenko

into four classes consisting of the numbers n of the form

$$4k,\ 4k + 1,\ 4k + 2,\ 4k + 3.$$

In general, for any natural number m, all the integers (not just the positive ones) can be divided into m classes, each class containing all the numbers that give the same remainder upon division by m. There is a description of these classes:

(0) the numbers a of the form
$$a = km;$$

(1) the numbers a of the form
$$a = km + 1;$$

(2) the numbers a of the form
$$a = km + 2;$$

. . . .

$(m - 1)$ the numbers a of the form
$$a = km + m - 1.$$

It's clear that every number belongs to one of these classes. The difference of two numbers from the same class is divisible by m, and the difference of two numbers from different classes is not divisible by m.

Definition. *If the difference of the integers a and b is divisible by the integer m, then a and b are said to be congruent modulo m.*

Congruence modulo m is written as

$$a \equiv b \pmod{m}.$$

The numbers a and b are congruent modulo m if and only if they belong to the same class—that is, when they give the same remainder upon division by m. In other words, $a \equiv b \pmod{m}$ means that $a = b + km$, where k is an integer. For example, $27 \equiv 7 \pmod{10}$, $78 \equiv 6 \pmod{24}$, $6 \equiv 0 \pmod{3}$, and $25 \equiv -4 \pmod{29}$.

Exercises

6. For any integer a, prove that (i) $a^3 \equiv a \pmod{6}$; (ii) $a^5 \equiv a \pmod{5}$.

7. Prove that the number

$$\frac{a(a+1)\ldots(a+k-1)}{k!}$$

is an integer.

8. Prove that $2^{100} \equiv 3^{100}$ for (i) modulo 5; (ii) modulo 13; (iii) modulo 211.

9. Prove that $11^{10} - 1$ is divisible by 100.

10. Let $S(N)$ be the sum of all digits of the number N. Prove that $N \equiv S(N)$ modulo 3 and 9.

11. Let $S(A) = S(5A)$. Prove that $A \equiv 0 \pmod{9}$.

12. The decimal notation of a number involves 1991 ones and a certain number of zeros. Can this number be a perfect square?

13. Use the fact that $10 \equiv -1 \pmod{11}$ to prove the following test for divisibility by 11: $a = \overline{a_n a_{n-1} \ldots a_0} \equiv 0 \pmod{11}$ if and only if $(-1)^n a_n + (-1)^{n-1} a_{n-1} + \ldots + a_0$ is divisible by 11.

Properties of congruences

Properties of congruences are similar to the properties of equalities.

1. If $a \equiv b \pmod m$ and $b \equiv c \pmod m$, then $a \equiv c \pmod m$.

If $a \equiv b \pmod m$ and $c \equiv d \pmod m$, then

2. $a + c \equiv b + d \pmod m$;
3. $a - c \equiv b - d \pmod m$;
4. $ac \equiv bd \pmod m$.

Thus congruences, just like equations, can be added, subtracted, and multiplied by each other.

Here, by way of example, we prove property (4). Since $a \equiv b$ and $c \equiv d$, then $a - b$ and $c - d$ are divisible by m. Now, the equality $ac - bd = a(c - d) + d(a - b)$ implies that $ac - bd$ is divisible by m; thus

$$ac \equiv bd \pmod m.$$

Exercise 14. Prove properties 1–3.

Let $a \equiv b \pmod m$. Properties 1–4 imply that, for any natural k,

5. $a^k \equiv b^k \pmod m$.

In addition, sometimes congruences can be reduced by a factor common to the left- and right-hand side.

6. If $ac \equiv bc \pmod m$ and the numbers c and m are coprime, then

$$a \equiv b \pmod m.$$

7. If $a \equiv b \pmod m$, k is an integer, and $a = ka_1$, $b = kb_1$, and $m = km_1$, then

$$a_1 \equiv b_1 \pmod{m_1}.$$

In other words, both sides of a congruence and the modulus can be divided by their common divisor.

Let's prove property 6. The number $c(a - b)$ is divisible by m. Since c and m are coprime, $a - b$ is divisible by m; therefore,

$$a \equiv b \pmod m.$$

Exercise 15. Prove property 7.

These properties imply that the remainder upon division of any algebraic expression involving integers $a, b, c, \ldots, z$ by a number m doesn't change when those integers are replaced by their remainders upon division by m.

Problem 3. Find the remainder upon division by 3 of the number

$$N = (1^2 + 1)(2^2 + 1)(3^2 + 1) \ldots (1000^2 + 1)$$

Solution. The remark above implies that

$$N \equiv (1^2 + 1)^{333} \cdot (2^2 + 1)^{333} \cdot (3^2 + 1)^{333} \equiv 2^{333} \cdot 2^{333} \cdot 1^{333} \equiv 2^{666} \equiv (2^2)^{333} \equiv 1^{333} \equiv 1.$$

Problem 4. Find all natural numbers n for which $8n + 3$ is divisible by 13 (see problem 1).

Solution. Using the properties of congruences, we can write the following chain of relations.

Let $8n + 3 \equiv 0 \pmod{13}$. Then
$8n \equiv -3 \pmod{13}$,
$8 \times 8n \equiv -24 \pmod{13}$,
$64n \equiv -n \equiv -24 \equiv -11 \pmod{13}$.

Therefore, $n \equiv 11 \pmod{13}$. Finally, $8n + 3$ is divisible by 13 if and only if $n = 13k + 11$.

Exercises

16. Find the remainders upon division of (i) $2^{1991} + 1$ by 17; (ii) $(3^{20} + 11)^{55}$ by 13.

17. Prove that (i) $2^{50} + 1$ is divisible by 125; (ii) $2^{48} - 1$ is divisible by 105; (iii) $2^{3^n} + 1$ is divisible by 3^{n+1} and is not divisible by 3^{n+2}.

18. Find all primes p for which $20p^2 + 1$ is a prime.

19. Prove that (i) $1^{1991} + 2^{1991} + \ldots + 30^{1991}$ is divisible by 31; (ii) $1^m + 2^m + \ldots + (n-1)^m$ is divisible by n for any odd m and n.

20. Find all natural n for which the number $20^4 + 16^4 - 3^4 - 1$ is divisible by 323.

21. Prove that the number $5^{2n+1} + 3^{n+2} \cdot 2^{n-1}$ is divisible by 19 for any natural n.

22. Find all n for which the fraction

$$\frac{15n+2}{14n+3}$$

is not in lowest terms.

Chinese remainder theorem

Consider m members of the arithmetic progression

$$a,\ a + d,\ \ldots,\ a + (m - 1)d, \qquad (*)$$

where a is an integer and d and m are coprime. The following theorem is often useful.

Theorem 1. *There exists exactly one member of progression (*) that is divisible by m.*

Proof. The difference of the kth and lth members of (*), which is equal to $d(k - l)$, is not divisible by m. Indeed, otherwise, $k - l$ would be divisible by m, which is impossible since $|k - l| < m$.

Therefore, no two of the numbers (*) are congruent modulo m, and they give different remainders upon division by m.

Therefore, all the congruence classes modulo m are represented among the numbers (*)—that is, for each of the remainders $0, 1, 2, \ldots, m - 1$, there exists exactly one member of sequence (*) that gives this remainder upon division by m.

This proves Theorem 1, and even a bit more.

Exercises

23. Find all triples of prime numbers of the form

$$p,\ p + 2,\ p + 4.$$

24. Find the longest finite arithmetic progression with the difference 6 consisting of prime numbers.

25. Fifteen prime numbers form an arithmetic progression with a difference d. Prove that $d > 30{,}000$.

We now use theorem 1 to prove the so-called *Chinese remainder theorem*. It was known in ancient China as far back as 2000 years ago.

Theorem 2. *Suppose n numbers $m_1, m_2, \ldots, m_n$ are coprime in pairs, and let the n numbers $r_1, r_2, \ldots, r_n$ be given such that $0 \le r_i \le m_i - 1$ ($i = 1, 2, \ldots, n$). Then there exists a number N that gives the remainder r_i upon division by m_i.*

In other words, $N \equiv r_i \pmod{m_i}$ for all $i = 1, 2, \ldots, n$.

Proof. We'll construct the proof by induction on n. For $n = 1$, the assertion of the theorem is obvious. Assume that the assertion is true for $n = k - 1$. Then, there exists a number M such that

$$M \equiv r_i \pmod{m_i} \text{ for all } i = 1, 2, \ldots, k - 1.$$

Let $d = m_1 m_2 \ldots m_{k-1}$. Consider the numbers

$M, M + d, M + 2d, \ldots, M + (m_k - 1)d.$

Since d and m_k are coprime, the proof of theorem 1 implies that there exists a number N among the numbers in the above sequence that gives the remainder r_k upon division by m_k. At the same time, N gives the remainders $r_1, r_2, \ldots, r_{k-1}$ when divided by $m_1, m_2, \ldots, m_{k-1}$, respectively. Thus, the theorem is proved.

Finally, we prove one more theorem.

Theorem 3. *For any numbers $m_1, m_2, \ldots, m_n$ that are coprime in pairs, and for any remainders $r_1, r_2, \ldots, r_n$ upon division by $m_1, m_2, \ldots, m_n$, there exist n consecutive integers $a, a + 1, \ldots, a + n - 1$ such that $a \equiv r_1 \pmod{m_1}$, $a + 1 \equiv r_2 \pmod{m_2}$, ..., $a + n - 1 \equiv r_n \pmod{m_n}$.*

In other words, for any numbers $m_1, m_2, \ldots, m_n$ that are coprime in pairs, there exist n consecutive integers that give any predefined set of remainders when divided by $m_1, m_2, \ldots, m_n$.

Proof. By the Chinese remainder theorem, there exists a number a such that

$$a \equiv r_1 \pmod{m_1},$$
$$a \equiv r_2 - 1 \pmod{m_2},$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$a \equiv r_n - \mathrm{n} + 1 \pmod{m_n}.$$

Then $a, a + 1, \ldots, a + n - 1$ satisfy the conditions of the theorem.

Exercises

26. Prove that (i) among any 10 and (ii) among any 16 sequential natural numbers there exists one that is coprime to all others. (iii) Is this assertion true for any 17 sequential natural numbers?

27. Prove that for any n there exist n sequential natural numbers, each of which is divisible by the square of a natural number.

28. Does a moment exist when the hour, minute, and second hands of a properly adjusted clock form angles of 120° with each other?

29. Find the smallest natural number that gives the remainders 1, 2, 4, and 6 when divided by 2, 3, 5, and 7, respectively.

30. Find the smallest even number a such that $a + 1$ is divisible by 3, $a + 2$ is divisible by 5, $a + 3$ is divisible by 7, $a + 4$ is divisible by 11, and $a + 5$ is divisible by 13.

Solving congruences

In problem 4 we found all integer n for which $8n + 3$ is divisible by 13. In other words, we solved the congruence

$$8n + 3 \equiv 0 \pmod{13}.$$

Now we're able to solve this problem for the general case. Let coprime numbers a and m be given. We want to solve the congruence

$$ax \equiv b \pmod{m},$$

where b is an arbitrary number.

By theorem 1 there exists a k such that $ak \equiv 1 \pmod{m}$. Multiply both sides of our original congruence by k to obtain

$$k(ax) = (ak)x \equiv x \equiv bk \pmod{m},$$

from which we immediately get

$$x = bk + ml,$$

where l is an arbitrary integer.

Thus the problem of finding k arises. For rather small m, k can be found by a straightforward search; however, the general solution of this and many other problems deserves a separate article.

Problem 5. Solve the congruence

$$32n \equiv 7 \pmod{37}.$$

Solution. Since $32 \equiv -5 \pmod{37}$, we obtain the equivalent congruences

$$5n \equiv -7 \equiv 30 \pmod{37},$$

or

$$n \equiv 6 \pmod{37}.$$

The problem of solving linear equations with integer coefficients in integers can be reduced to solving congruences.

Problem 6. Find all pairs of integers x, y satisfying the equation $7x - 23y = 131$.

Solution. Since $23 \equiv 2 \pmod 7$, we obtain the congruences $2y \equiv -131 \pmod 7$ or $2y \equiv 9 \equiv 2 \pmod 7$, from which we get $y \equiv 1 \pmod 7$.

Thus $y = 7k + 1$, where k is an integer. Now we can easily find x:

$$7x - 23(7k + 1) = 131,$$

which can be rearranged as

$$7x = 154 + 23 \cdot 7k$$

and further reduced to

$$x = 22 + 23k.$$

In conclusion, we suggest that you solve the following problems.

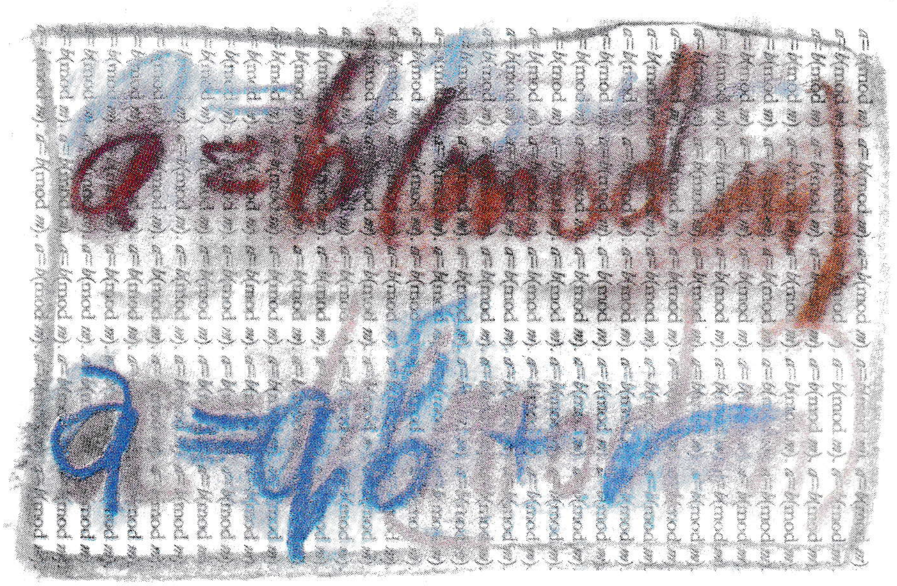
Exercises

31. Solve the congruences (i) $17x \equiv 19 \pmod{37}$; (ii) $147x \equiv 63 \pmod{29}$.

32. Find integer solutions to the equations (i) $7x + 8y = 1$; (ii) $13x - 15y = 16$; (iii) $257x + 18y = 175$.

33. Find integer solutions to the system of equations

$$\begin{cases} 32x + 5y - 7z = 1, \\ 4x + 9y + 11z = 2. \end{cases}$$



The wave mechanics of Erwin Schrödinger

by A. Vasilyev

THE NAME OF THE OUTstanding physicist Erwin Schrödinger is inseparably linked to the formation and development of quantum mechanics. His wave equation is the centerpiece in this theory and ensured Schrödinger a prominent place in the history of physics.

The range of Schrödinger's creativity is amazing. He made significant, and sometimes definitive, contributions to quantum theory and electrodynamics, the physics of elementary particles and cosmic rays, statistical mechanics and thermodynamics, the general theory of relativity, cosmology, and field theory. He also conducted pioneering interdisciplinary research in physics and biology and wrote papers on the philosophy of natural science.

The scope of Schrödinger's interests went far beyond the limits of physics and the natural sciences in general. He was an expert in ancient and oriental philosophy, read widely in world literature, mastered many languages, including Latin and ancient Greek, and preferred to read classic works of literature in their original language. And that's not all: Schrödinger sculpted in clay and wrote poetry—he even published a book of his verses. His contemporaries were astounded by his wide-ranging and encyclopedic mind.

The turbulent events of the 20th century (the two world wars in particular) had a major impact on Schrödinger's life. He was forced to move repeatedly from one European country to another, returning to his native Austria only in his declining years.

Erwin Schrödinger was born in 1887 in Vienna, where he graduated from an elite school. After passing his final exams with flying colors, he entered the University of Vienna, where he chose physics and mathematics as his specialty. As a pupil of Friedrich Hasenoehrl (1874–1917), an outstanding representative of the Vienna school of physics, Schrödinger learned thoroughly the mathematical methods of physics. Even in his student years Schrödinger combined a brilliant physical intuition with a masterly command of those methods. He began his scientific work at the University of Vienna by studying classical mechanics, Brownian motion, and the theory of errors. However, he was soon attracted by quantum theory, which at that time was already scoring great successes.

In 1920 Schrödinger moved to Germany, but soon thereafter he was invited to head the department of theoretical physics at the University of Zürich. Around this time the French physicist Louis de Broglie was developing the idea of expanding the wave–particle duality of light (postulated by Albert Einstein to explain the photoelectric effect) to material particles. According to de Broglie, every particle that has energy and momentum can be characterized by some oscillation frequency and wavelength. Schrödinger familiarized himself with this theory in 1925, and it inspired him to develop wave mechanics to describe the physical properties of atoms. The next year he began publishing a series of papers under the common title "Quantization as a Problem of Eigenvalues," which in time became a classic in the science literature but immediately put the heretofore mysterious theory of wave mechanics on a solid footing.

The concepts of quantum physics in use at that time were uncoordinated and contradictory in many respects. For example, in Bohr's atomic model the laws of classical mechanics and electrodynamics were used to calculate the orbits of electrons and the spectral lines of radiated (or absorbed) light, while quantum conditions were applied to explain the stability of electron orbits. An important step toward overcoming this contradiction was made in 1925 by Werner Heisenberg, whose work laid the foundations of matrix mechanics (created subsequently by Heinsenberg together with Max Born and P. Jordan).

Heisenberg began with the proposition that, in studying the physics of the microcosm, one should be interested not in quantities that cannot be observed (such as the orbits of electrons or their orbital periods), but in values that can be measured

Art by Vadim Ivanyuk

$\Delta\Psi + (2m/\hbar^2)(E + e^2/r)\Psi = 0$
$\Delta = d^2/dx^2 + d^2/dy^2 + d^2/dz^2$
Э. ШРЕДИНГЕР
ФИЗИКА
БИОЛОГИЯ
ЛИТЕРАТУРА
ЛИНГВИСТИКА
ЛАТЫНЬ
ВОЛНОВАЯ
МЕХАНИКА

experimentally—for instance, the frequencies of the emitted light and the intensities of spectral lines. Heisenberg's goal was to formulate a strict quantum theory that, by analogy with classical mechanics, would involve relationships between observable values only. His formal mathematical apparatus of matrix mechanics was soon refined and yielded precise solutions to many physics problems.

Schrödinger was aware of the fresh winds of quantum theory, but he disliked the complicated tricks of matrix mechanics and the fact that he could not visualize it. Working as he did in Zürich, he was far from the centers of atomic physics, and he had little personal contact with its leading exponents. In addition, Schrödinger's independent nature would not permit him to attach himself to one or another doctrine. As a result, Schrödinger took his own path to wave mechanics by considering an atom as an oscillating system and postulating the equivalence of all possible natural ("eigen") oscillations of this system with the stable energy states of the atom. His acquaintance with the theory of oscillations helped Schrödinger formulate the basic physical idea within the framework of the problem of the eigenvalues of differential equations.

In constructing a new atomic theory on the basis of de Broglie's hypothesis, Schrödinger leaned heavily on his knowledge of the analytical mechanics of William Hamilton. As early as the middle of the 19th century, this great Irish mathematician had given theoretical mechanics its final form and established formal relationships between classical mechanics and geometrical optics. The optical-mechanical analogy helped him formulate the basic laws of these seemingly different sciences through similar mathematical equations. Within the framework of this optical-mechanical analogy, the laws of motion of a material point with a given energy in a static field of forces looks like the propagation of a monochromatic light beam in a medium with a varying refractive index. Within this analogy, the constant value of energy of a material point corresponds to the constant oscillation frequency of the light, while the speed of the material point corresponds to the group speed of light propagation in the medium.

Schrödinger decided to extend the mathematical analogy between optics and mechanics to the wave properties of light and matter. In this way he overcame numerous problems and finally obtained the famous wave equation for hydrogen atoms:

$$\nabla\psi + (2m/h^2)(E + e^2/r)\psi = 0,$$

where ψ is the wave function, m is the electron's mass, e is the electron's charge, r is the distance between the electron and the nucleus, E is the total energy of the system, and h is Planck's constant. The symbol ∇ (the Laplace operator) means a special mathematical operation: the sum of second derivatives with respect to the space coordinates. In Cartesian coordinates this operator takes the form

$$\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

This equation is a generalization of de Broglie's hypothesis regarding the wave properties of matter. From the mathematical viewpoint it's just a linear differential equation whose solutions are standing waves. Now the stationary electron orbits in Bohr's atom model could be considered natural oscillations similar to the oscillations of a taut string, which oscillates only at a certain set of discrete frequencies determined by its length and boundary conditions.

Using his equation, Schrödinger calculated the energy levels of an atomic harmonic oscillator. Choosing the hydrogen atom as an example of such an oscillator, he showed that the theoretical energy levels either coincide with those obtained with Heisenberg's matrix mechanics or closely agree with experimental data. His application of well-known methods of mathematical physics made Schrödinger's theory more attractive to physicists than Heisenberg's matrix theory. Moreover, in his third paper on quantization Schrödinger demonstrated the complete mathematical equivalence of matrix and wave mechanics: Heisenberg's matrices could be calculated from his wave functions and vice versa.

Although the works of Heisenberg and Schrödinger completed the edifice of quantum theory, they did not end the discussion on its physical meaning. Now the nature of the wave function became the dominant problem. On this issue, physicists split into two camps. For Schrödinger, the authority of the classical concept of motion was unshakeable, and so he visualized the wave function and in this regard spoke of oscillatory motion in three-dimensional space. The quantum jump in the atom during the transition from one state to another was interpreted as a gradual transformation from the state corresponding to natural oscillations with energy E_m to the state with energy E_n, where the extra energy is radiated as an electromagnetic wave. Within this framework, the electron was considered a charged cloud enveloping the nucleus that could be transformed into a spatially distributed electromagnetic wave moving continuously without any quantum jumps. In this way quantum mechanics could coexist naturally with classical physics, and this attracted Schrödinger and the pleiad of outstanding physicists nurtured in classical physics: Louis de Broglie, Albert Einstein, Max von Laue, and Max Planck.

Their opponents were eminent physicists as well: Wolfgang Pauli, Werner Heisenberg, and Niels Bohr. Intensive work on these controversial theories showed that the semiclassical interpretation of wave mechanics was wrong, and it was impossible to construct a consistent quantum theory entirely on the basis of wave mathematics without the concept of the wave–particle duality.

A hint at a solution to this impasse was found in research on atomic collisions performed by Max Born in 1926. Analysis of electron and alpha-particle scattering on atomic nuclei provided the key to understanding the meaning of Schrödinger's wave function: The square of its amplitude meant the probability of finding the particle at a given point in space. Therefore, the wave function describes individual events (such as the emission of a single quantum of light) only as a probability of its occurrence. This interpretation put wave mechanics on a firm physical basis, and soon this novel theory became relatively complete and consistent. At present, the statistical interpretation of quantum theory is universally recognized.

Although Schrödinger's hopes of creating a kind of classical field theory for atomic phenomena were not realized, his wave mechanics was a giant step in the development of mathematical methods of quantum theory. Moreover, it helped many physicists understand and "feel" its essence. In later years Schrödinger worked intensively on many theoretical problems of wave mechanics and their numerous practical applications. In this period he wrote very important papers on perturbation theory.

As the creator of wave mechanics, Erwin Schrödinger advanced himself to the first rank of contemporary physicists. In 1933 he became a Nobel laureate along with Paul Dirac for their "discovery of new forms of atomic theory." ◘

Quantum on quantum theory:

A. Eisenkraft and L. D. Kirkpatrick, "The Nature of Light," November/December 1996, pp. 30–31, and May/June 1997, p. 39.

A. Stasenko, "Not All Is Revealed," November/December 1996, pp. 42–44.

A. Sakharov, "Does Elementary Length Exist?" May/June 1997, pp. 14–20.

D. Sviridov and R. Sviridova, "The Quantum Nature of Light," November/December 1998, pp. 28–31.

A. Korzhuyev, "Bohr's Quantum Leap," January/February 1999, pp. 42–46.

A. Semenov, "Vacuum," July/August 1999, pp. 12–15.

A. Levashov, "Experiments of Frank and Hertz," March/April 2000, pp. 38–41.

Self-propelled sprinkler systems

by A. Stasenko

ONCE UPON A TIME, A WOMan asked her grandson, who was a student at the Moscow Physics and Technology Institute, to water her vegetable garden. She had everything he needed: water faucet, hose... There was only one catch: her grandson was not only a clever kid, he liked to read. And somewhere he read that way back in 450 B.C. (or thereabouts) an ancient Greek (or somebody) expressed the view that manual labor should be replaced—gradually at first, and then completely—by automatons. The student's chin dropped onto his chest as he gave this a good hard think. He recalled Segner's wheel and imagined a pipe with ends curved in opposite directions, set on a pipe such that it could rotate freely, driven by the force of the running water. The student mentally added the necessary labels to his imaginary machine (figure 1) and thought some more.

If v_w is the speed of water in a pipe with cross-sectional area S, then $Q = \rho_w v_w S$ is the amount of water coming out of the end every second. Since water is incompressible (that is, its density ρ_w is constant), the water speed is identical for all sections of the pipe having the same cross section S. (Here the student felt he was applying the conservation of mass.)

If a pipe of length $2l$ acquires an angular speed ω due to the reactive force of the water, the tangential velocity at the ends of the pipe is ωl and is directed counter to the velocity $\mathbf{v}_w$ of the moving water. Therefore, in the reference frame of the garden, the speed of the ejected water is $v_0 = v_w - \omega l$. Thus the flow of momentum through the opening is $Q(v_w - \omega l)$, which has the dimensions (kg/s) · (m/s) ≡ N. Surprise! It's the same dimensions as force.

So we have a pair of forces F, equal in magnitude, parallel, and acting in opposite directions. Here l is the moment arm of each force relative to the axis of rotation. As a result, the pipe experiences a torque:

$$F \cdot 2l = 2Q(v_w - \omega l)l. \qquad (1)$$

What angular speed can the pipe achieve? Is there anything that might prevent unlimited acceleration of the spinning pipe? Of course there is. For instance, torque from

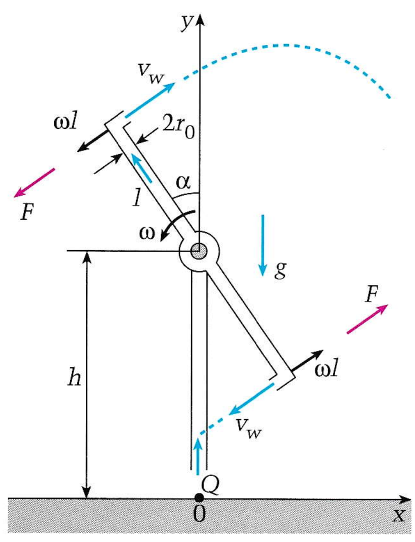


Figure 1

friction in the hub (can we ever escape from friction!), air resistance, and so on. Our student was tempted to utter the well-worn words, "Neglect air resistance," but as an honorable fellow he decided to attempt a numeric estimate.

He knew that the force of air resistance acting on a moving object is proportional to the square of the speed v of the object relative to the air, its perpendicular cross section $S_\perp$, and the density of air ρ. Before we can neglect the air resistance, we need to compare it with some other force that is considered essential—say, with the flow of momentum of the water (Qv_w), which in the reference system of the pipe is

$$\frac{\rho v^2 S_\perp}{Q v_w} \le \frac{\rho \cdot (\omega l)^2 \cdot 2r_0 l}{\rho_w \pi r_0^2 v_w \cdot v_w}$$

$$\sim \left(\frac{\rho}{\rho_w}\right)\left(\frac{\omega l}{v_w}\right)^2\left(\frac{l}{r_0}\right).$$

To make this inequality stronger, we use the largest change in speed, $v = \omega l$, which is achieved at the end of the pipe; and, of course, we neglect the fact that a rotating pipe pulls the surrounding air along with it. But let's not let that keep us from finishing our estimate. Clearly ωl cannot be greater than v_w—otherwise the pipe won't spin. The density ratio ρ/ρ_w is about 10^{-3}, so for a pipe of "reasonable" size (say, $l \sim 10$ cm and $2r_0 \sim 1$ cm) we get a ratio of the two forces of the order of 10^{-2}, or even less. This means that, with an accuracy of a few percent, we can indeed neglect the air resistance.

What else do we need to take into account? Friction in the hub, of course. For steady-state rotation the frictional braking torque τ_{fr} is equal to the accelerating reactive torque of the water jet:

$$2Q(v_w - \omega l)l = \tau_{fr},$$

from which we get the speed of the water ejected from the pipe in the garden's coordinate system:

$$v_0 = v_w - \omega l = \frac{\tau_{fr}}{2lQ}.$$

(Again we assume that ωl is not greater than the speed of the water relative to the pipe.)

Now what? Well, now we say, in our best professorial voice, that the problem "reduces to" a familiar textbook problem about the motion of an object thrown at an angle α to the horizon with an initial speed v_0 from a point with coordinates

$$x_0 = -l \sin \alpha, \quad y_0 = h + l \cos \alpha$$

(see figure 1). The solution to this standard problem yields

$$x = -l \sin \alpha + v_0 \cos \alpha \cdot t,$$

$$y = h + l \cos \alpha + v_0 \sin \alpha \cdot t - \frac{gt^2}{2},$$

where t is the time from the moment when the elementary mass of water is ejected from the pipe. This solution assumes that the elements of the water stream, or the drops that the stream breaks into, interact with neither the air nor one another.

When you water a garden, two things matter: *where* and *how much*. In our system of coordinates (the garden), the ordinate of the soil is $y = 0$. Denote the point on the soil where the water lands as $x_\downarrow$. Eliminating the time from the previous two equations, we get

$$x_\downarrow = -l \sin \alpha + \frac{v_0^2 \cos \alpha}{g} x \left(\sin \alpha + \sqrt{\sin^2 \alpha + \frac{2gh}{v_0^2}\left(1 + \frac{l}{h}\cos \alpha\right)} \right). \qquad (2)$$

Art by Tomas Bunk

In particular, this formula contains the known expression for the distance a thrown object travels from the origin ($h = 0$, $l = 0$):

$$x_{0\downarrow} = 2\frac{v_0^2}{g} \cos \alpha \sin \alpha.$$

As you probably know, in this case the largest distance flown is attained for $\alpha = 45°$:

$$x_{0\max} = \frac{v_0^2}{g}. \qquad (3)$$

Formula (2) shows that in the general case the solution depends on two parameters: the ratio of the initial values of the potential and kinetic energies

$$a = \frac{gh}{v_0^2/2}$$

and the geometric characteristics of our device l/h. No wonder our student turned to his computer for help in analyzing these relationships.

We'll restrict our analysis to a characteristic case. For example, assume $l/h << 1$ (which means that the rotator is small compared to its height above the ground). In addition, we'll assume that the "firing range" is also greater than l (in other words, even when ejected at zero altitude the water stream will travel a long way). In this case, the abscissa of the landing point on the soil, determined by the characteristic distance defined by formula (3), is

$$\bar{x} = \frac{x_\downarrow}{v_0^2/g} = \cos\alpha\left(\sin\alpha + \sqrt{\sin^2\alpha + a}\right).$$

We can also find the angle of ejection corresponding to the maximum "firing distance," and the distance itself, by equating the time-derivative to zero:

$$\frac{d\bar{x}}{d\alpha} = 1 - 2\sin^2\alpha + \frac{\sin\alpha}{\sqrt{\sin^2\alpha + a}}\left(1 - 2\sin^2\alpha - a\right) = 0,$$

from which we get

$$\sin^2\alpha_m = \frac{1}{a+2}, \quad \bar{x}_{\max} = \sqrt{a+1}.$$

We see that the higher the ejection point, the greater the range of the water. For example, if $a = 1$ (the initial values of the potential and kinetic energies are equal),

$$\alpha_m = \arcsin\frac{1}{\sqrt{3}} \approx 35°,$$

and the "watering range" is greater by a factor of $\sqrt{2}$ compared to the range when the water is ejected at ground level (formula (3)).

Figure 2 shows the qualitative dependence of the relative coordinate of the landing point on the angle α (or on time, because $\alpha = \omega t$). The positions of the rotating pipe corresponding to some characteristic angles are shown in the upper part of figure 2. Even this graph clarifies the irregular character of the watering process. The bottom part of this figure shows the graph of the function $|d\alpha/d\bar{x}|$, which describes the distribution of water on the soil. Indeed, if the pipe turns through an angle $d\alpha$, the water ejected during the corresponding time interval $dt = d\alpha/\omega$ will land in the region $d\bar{x}$ (we assume that the water is absorbed immediately by the soil). It's clear that at some moments in time the watering density tends to infinity. This is because at these points (where $\alpha = \alpha_m$) the stream of water landing there stops for a finite time in order to change direction. Infinity appears in our calculations because we considered the stream an infinitely thin thread carrying a nonzero water flow.

"However," the student said to himself, "the stream of water is not a line—it has a diameter of $2r_0$ even at the nozzle of the pipe, and as it travels it breaks up into drops, which are decelerated by the surrounding air. Also, the centrifugal inertial force can modify the pressure distribution along the stream's axis. Obviously this whole theory needs to be tested, and expanded," which our student continued to think about.

And what about Grandma's garden? (Let's just say there's more brown there than green.)

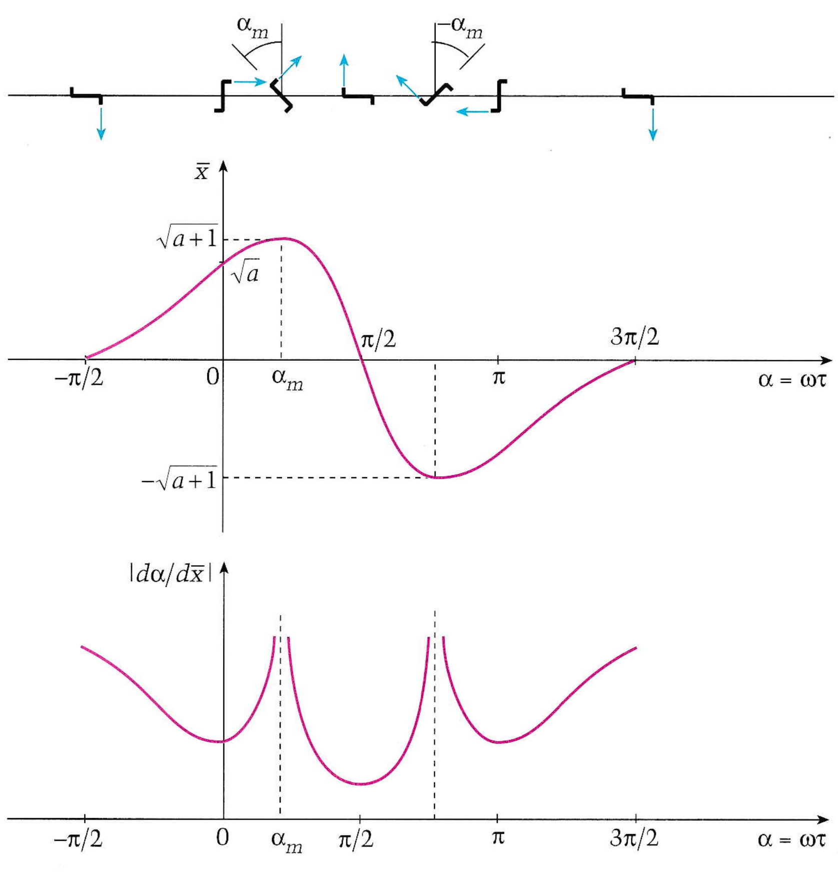
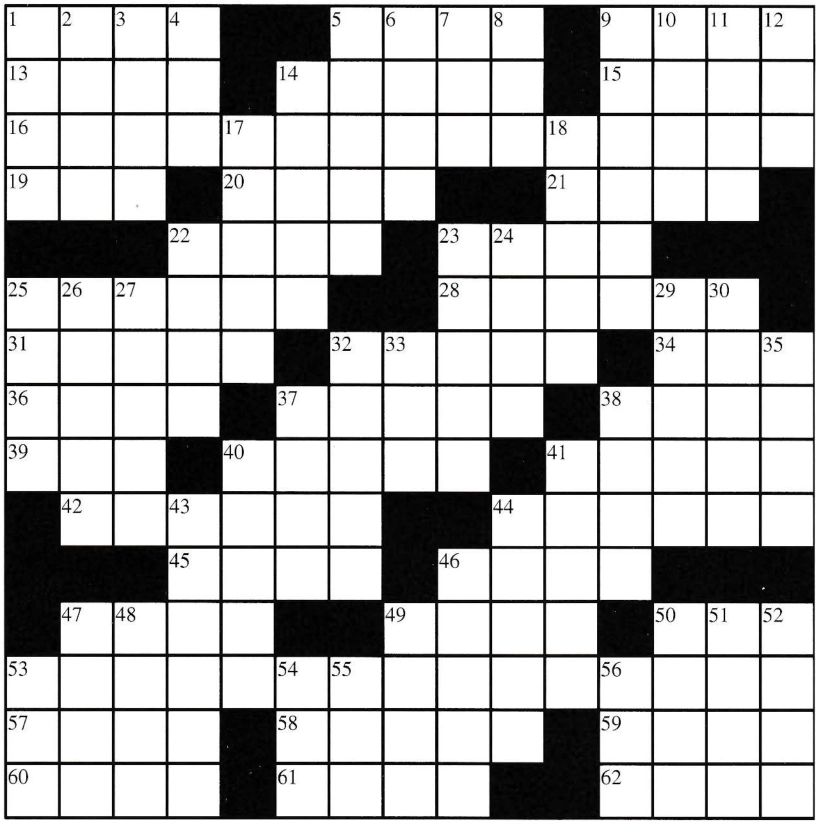


Figure 2

Criss × cross science

by David R. Martin



ACROSS

1 Seasoning
5 Energy units
9 Kiln
13 Fencing sword
14 Orchestra members
15 Ghost: comb. form
16 Austrian physicist (1844–1906)
19 Donkey
20 Orderly
21 Poet Gunther ___ (1907–1972)
22 Act
23 Town along the Sokoto
25 Monochromatic light sources
28 Type of bomb
31 Mature person
32 Ground hominy
34 Energy unit: abbr.
36 1551
37 Type of exams
38 Indonesian island
39 Ten decibels
40 English astronomer John Couch ___ (1819–1892)
41 Number part
42 Angle unit
44 Type of algebra
45 Terrapin
46 Sounds from an audience
47 Wise man
49 Russian weight unit
50 College degree
53 English mathematician (1861–1947)
57 Point of minimum disturbance
58 Chills
59 Sicilian volcano
60 German physicist Ernst ___ (1840–1905)
61 Focusing device
62 Space org.

DOWN

1 Actress ___ Ward
2 Constellation
3 Calculator displays: abbr.
4 Work hard, in London
5 965,549 (in base 16)
6 Square ___
7 Coagulated substance
8 Fast plane
9 Element 76
10 44,204 (in base 16)
11 Hyperbolic function
12 Trig. function
14 S-shaped curves
17 Like a noble gas
18 Goose eggs
22 Sandwich shop
23 Sends
24 Lawyers: abbr.
25 ___ shift (of H spectrum)
26 Totaler
27 Roman general (138–78 B.C.)
29 Real or virtual follower
30 Actress Weston of Alice
32 ___ unified theory
33 Computer memory: abbr.
35 Glass: comb. form
37 Singer Anita ___
38 Storage boxes
40 Actress Anouk
41 Current rectifier
43 Temperature unit
44 Plunders
46 Vaporizes
47 Messy person
48 45,019 (in base 16)
49 Unit of loudness
50 10^{15}: pref.
51 Common German name
52 56,746 (in base 16)
53 Collection of anecdotes
54 650
55 Grief
56 Egg layer

SOLUTION IN THE NEXT ISSUE

SOLUTION TO THE MARCH/APRIL PUZZLE

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

The theorem of Menelaus

by B. Orach

THE SIMPLE AND ELEGANT result known as Menelaus' Theorem is often hidden among more complicated and specialized problems in the problem-solving literature. It is a tiny gem of ancient mathematics.

Menelaus' Theorem concerns a line which intersects all three sides of a triangle. We will call such a line a *secant line* for the triangle. Clearly, there is no case when all three points of intersection lie on the triangle's sides, so at least one point of intersection lies on the extension of a side. This is the case one usually sees in problems.

Menelaus' Theorem states the following:

Let a secant line to triangle ABC intersect the sides at A_1, B_1, C_1 (see figure 1). *Then*

$$\frac{AB_1}{B_1C}\cdot\frac{CA_1}{A_1B}\cdot\frac{BC_1}{C_1A}=1.$$

(To help keep track of the letters in the formula, follow the outline of

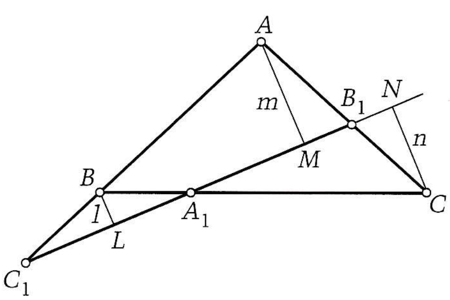
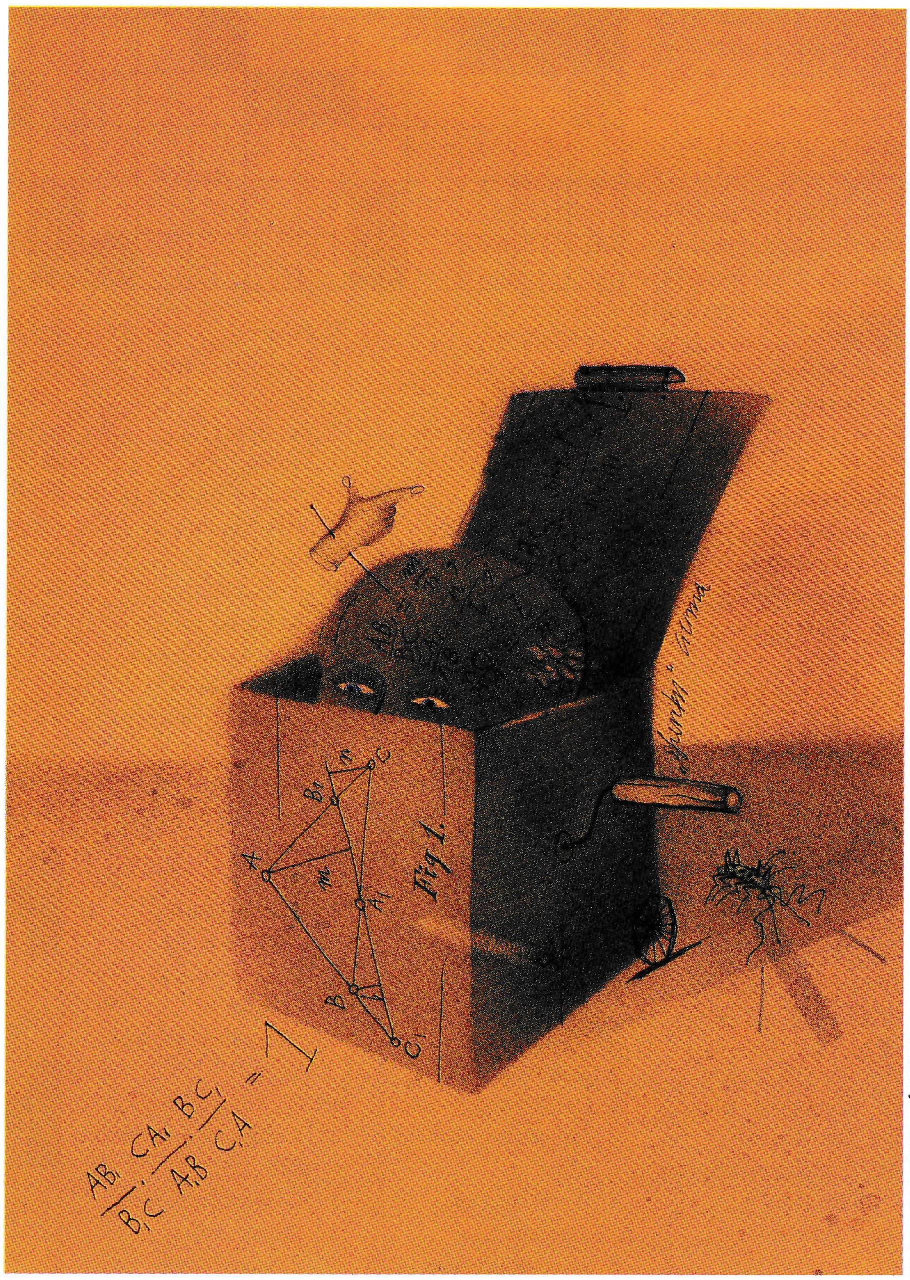


Figure 1



Art by Ekaterina Silina

the triangle from a vertex to the intersection point and then from the intersection point to the next vertex.)

Proof. Draw any set of parallel segments from the three vertices to the secant line. This creates many sets of similar triangles.

From similar triangles AMB_1, CNB_1, we have

$$\frac{AB_1}{B_1C}=\frac{m}{n},$$

From similar triangles CNA_1, BLA_1, we have

$$\frac{CA_1}{A_1B}=\frac{n}{l},$$

From similar triangles BLC_1, AMC_1, we have

$$\frac{BC_1}{C_1A}=\frac{l}{m}.$$

It remains to multiply these equations to obtain

$$\frac{AB_1}{B_1C}\cdot\frac{CA_1}{A_1B}\cdot\frac{BC_1}{C_1A}=\frac{m\cdot n\cdot l}{n\cdot l\cdot m}=1.$$

The theorem is proved.

We know this theorem from an Arabic translation of the book *Spherica* by Menelaus of Alexandria (1st century A.D.).

To demonstrate the effectiveness of this theorem, consider two solutions of a problem—a solution using areas and one based on Menelaus' theorem.

Problem 1. *Let AD be the median of triangle ABC* (figure 2). *A point K is taken on AD such that AK : KD =* 3 : 1. *Find the ratio in which line BK divides the area of triangle ABC.*

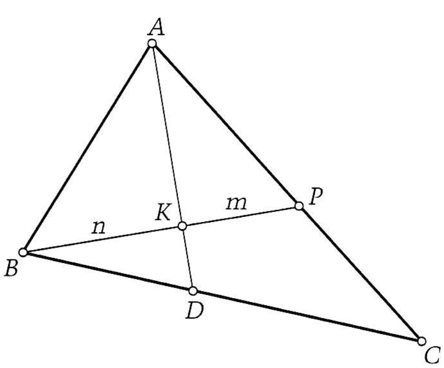


Figure 2

Solution using areas. We will use, over and over, the theorem that the ratio of the areas of triangles with equal altitudes is the ratio of their bases, and absolute value signs to denote area. Draw segment PD, and suppose $|PDK| = S$. Then, since triangles PDK, PAK have the same altitude from P, the $|PAK|$ is $3S$.

Now suppose the ratio $BK : KP = k$. We will solve for k by comparing areas. Triangles ADK, PDK have the same altitude from D, so $|BDK| = kS$. Triangles PBD, PDC have equal bases $BD = BC$, and equal altitudes from P, so $|PDC| = |PBD| = |KBD| + |PKD| = S + kS$. Triangles ABD, ADC also have equal altitudes (from A) and equal bases, and so $|ABD| = |ADC|$.

In terms of the various areas we have represented, this last equation can be written as: $3kS + kS = 3S + S + S + kS$, or $3kS = 5S$, and $k = 5/3$. Similarly, $|ABP| = 3kS + 3S = 8S$, and $|PBC| = kS + S + S + kS = 16S/3$. The ratio $|APB| : |PBC|$ is thus $8 : (16/3) = 3 : 2$.

Solution based on the Menelaus' Theorem. We apply the theorem to triangle ACD and secant line BP to obtain

$$\frac{AP}{PC}\cdot\frac{CB}{BD}\cdot\frac{DK}{AK}=1,$$

$$\frac{AP}{PC}\cdot 2\cdot\frac{1}{3}=1,$$

$$\frac{AP}{PC}=\frac{3}{2}.$$

The advantage of using Menelaus' Theorem is obvious.

The following proposition, which is the converse of Menelaus' Theorem, is often very useful.

Suppose points A_1, B_1, C_1 lie on sides BC, AC, and AB, respectively, of some triangle ABC, or on the extensions of these sides (figure 3). *If*

$$\frac{AB_1}{B_1C}\cdot\frac{CA_1}{A_1B}\cdot\frac{BC_1}{CA}=1,$$

the points A_1, B_1, and C_1 are collinear.

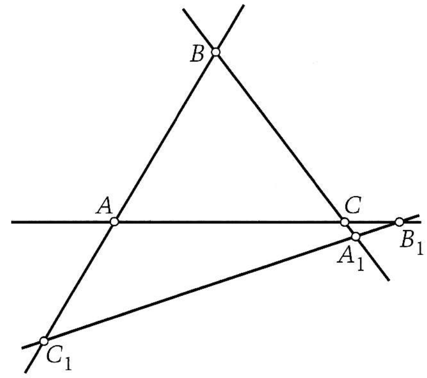


Figure 3

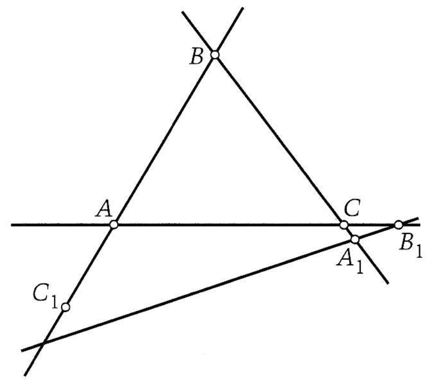


Figure 4

Proof. Suppose point C_1 were not on line A_1B_1 (figure 4). Let C' be the intersection of lines A_1B_1 and AB. Then, by Menelaus' Theorem,

$$\frac{AB_1}{B_1C}\cdot\left|\frac{CA_1}{A_1B}\right|\cdot\left|\frac{BC'}{C'A}\right|=1.$$

However, it's clear that

$$\left|\frac{BC_1}{C_1A}\right|\neq\left|\frac{BC'}{C'A}\right|.$$

Thus the relation in the condition of the theorem cannot be true. The contradiction obtained proves the theorem.

The following elegant problem seems to have been known since ancient times. It has often been used in mathematical olympiads.

Problem 2. *Three circles of different radii are arranged in a plane so that none of them lies inside the other two. For every pair of circles, we construct the intersection point of their two common tangents. Prove that these three intersection points are collinear* (figure 5).

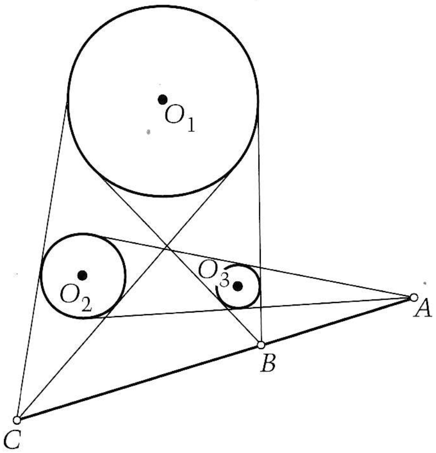


Figure 5

Solution. Let the radii of the given circles centered at points O_1, O_2, and O_3 be r_1, r_2, and r_3, respectively. Then

$$\left|\frac{O_1C}{CO_2}\right| = \frac{r_1}{r_2},$$

since the circles centered at O_1 and O_2 are homothetic with respect to point C with the homothety coefficient $|r_1/r_2|$. Similarly,

$$\frac{O_2A}{AO_3} = \frac{r_2}{r_3},$$

and

$$\frac{O_3B}{BO_1} = \frac{r_3}{r_1}.$$

Thus we have

$$\frac{O_1C}{CO_2}\cdot\frac{O_2A}{AO_3}\cdot\frac{O_3B}{BO_1} = \frac{r_1}{r_2}\cdot\frac{r_2}{r_3}\cdot\frac{r_3}{r_1} = 1.$$

By the inverse of the Menelaus' Theorem, points A, B, and C lie on a line.

Exercises.

1. Points M and N are given on sides AB and AC, respectively, of triangle ABC, such that $AM/MB = CN/NA = 1/2$. Let S be the point of intersection of segments BN and CM. Determine the ratio in which S divides each of these segments.

2. The bisector AD in triangle ABC divides BC in the ratio 2 : 1. Determine the ratio in which the median CE divides this bisector.

3. A point D on side AB and points E and F are taken on side BC of triangle ABC, such that $AD : DB = 3 : 2$, $BE : EC = 1 : 3$, and $BF : FC = 4 : 1$. Determine the ratio in which line AE divides segment DF.

4. The point of intersection of the altitudes of triangle ABC is at the center of the altitude drawn from vertex C of the triangle. Prove that $\cos\angle C = \cos\angle A \cos\angle B$, where $\angle A$, $\angle B$, and $\angle C$ are the angles of the triangle.

5. In equilateral triangle ABC with side a, E and F are midpoints of sides BC and AC, respectively, F is a point on segment DC, and BF intersects DE at M. If the area of triangle BMD is 5/8 of the area of triangle ABC, find, in terms of a, the length of MF (figure 6).

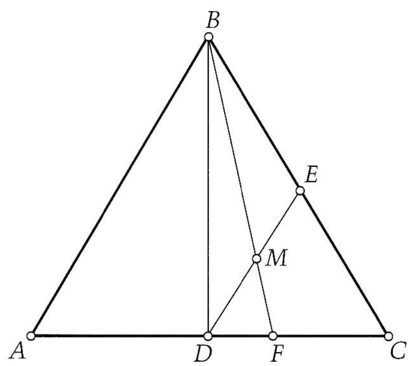


Figure 6

6. A parallelogram $ABCD$ is given. Point M divides side AD in the ratio p, and point N divides DC in the ratio q. Lines BM and AN meet at point S. Determine the ratio $AS : SN$ (figure 7).

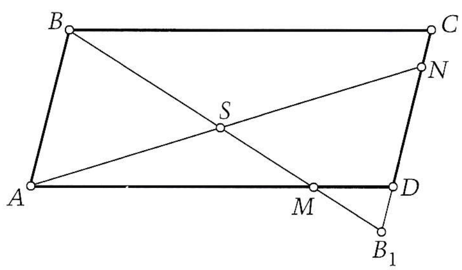


Figure 7

7. The area of parallelogram $ABCD$ is 1. A line is drawn through the midpoint M of side BC, such that it intersects the diagonal BD at a point Q. Determine the area of quadrilateral $QMCD$ (figure 8).

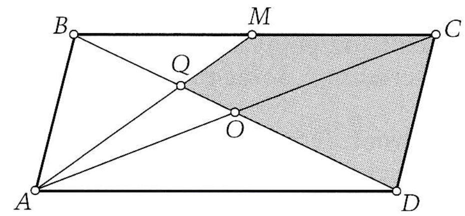


Figure 8

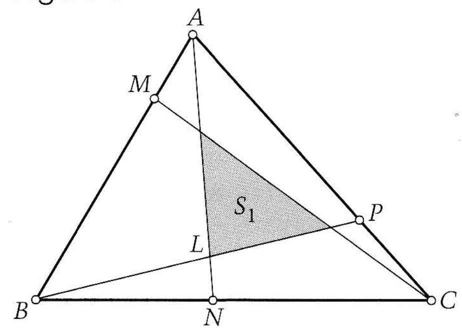


Figure 9

8. The sides of triangle ABC are divided by points M, N, and P, such that $AM : MB = BN : NC = CP : PA = 1 : 4$. Determine the ratio of the area of the triangle bounded by lines AN, BP, and CM to the area of triangle ABC (figure 9). ◘

CONTINUED FROM PAGE 21

inside the sphere. Any stereographic projection preserves angle measures, so the angles α, β, and γ are equal to the original relative angles. In the first case, $\alpha + \beta + \gamma$ is less than the sum of angles of triangle ABC, and in the third case the sum is greater. The proposition is proved.

Assume that the vertices of polyhedron M can be colored black or white so that no two vertices of the same color are neighboring.

Problem 5. *Assume that the same number of faces meet at every vertex of polyhedron M and the vertices are colored black or white such that no two vertices of the same color are neighboring. Prove that the number of black vertices is equal to that of the white ones.*

Problem 6. *Prove that the vertices of a polyhedron can be colored black and white so that no two vertices of the same color are neighboring if and only if every face of the polyhedron has an even number of sides.*

Hint: It is sufficient to take any vertex and color it white, then color the neighboring vertices black, and so on. It remains to prove that no contradictions can occur in the process—that is, any closed polygonal line composed of the edges of the polyhedron has an even number of sides.

Let's return to polyhedron M with its colored vertices. Since three faces meet at every vertex of M, the number of white vertices is equal to the number of black ones. We chose some (or even all) of the black vertices, and cut them off with planes. Every vertex is "cut off" each of the chosen vertices with a plane that (a) intersects only those edges that meet at the chosen vertex and (b) doesn't contain any vertices of M. This forms a new polyhedron, which we will call M'. It differs from M in that it has triangular faces instead of some number of black vertices. We now prove that M' is absolutely uninscribable.

Assume that M' or another polyhedron with the same structure is inscribed in a sphere. Consider the three faces that correspond to the three faces of polyhedron M that meet at a vertex that was cut off. These faces are adjacent in pairs, and the vertex of the trihedral angle that was cut off certainly lies outside the sphere, since M' is inscribed. Thus the sum of the relative angles of this trihedral angle is less than π. To every edge of M, we assign the relative angle of the dihedral angle formed by the faces adjacent to the corresponding edge of M'. For all white and all undesignated black vertices, the sum of all angles thus assigned is π, and the sum of the assigned angles at designated black vertices is strictly less than π. As in the proof of the Steinitz Theorem, write out these equalities and inequalities, multiply the sums corresponding to the white vertices by -1, and add them up. We obtain a strict inequality. However, this inequality reads $0 < 0$, which is impossible. Indeed, the left-hand side of this inequality equals zero, since the magnitude of every angle enters the sum once with a negative sign (at a white vertex) and again with a positive sign (at a black vertex). The right-hand side is also equal to zero, since the number of white vertices is the same as the number of black ones.

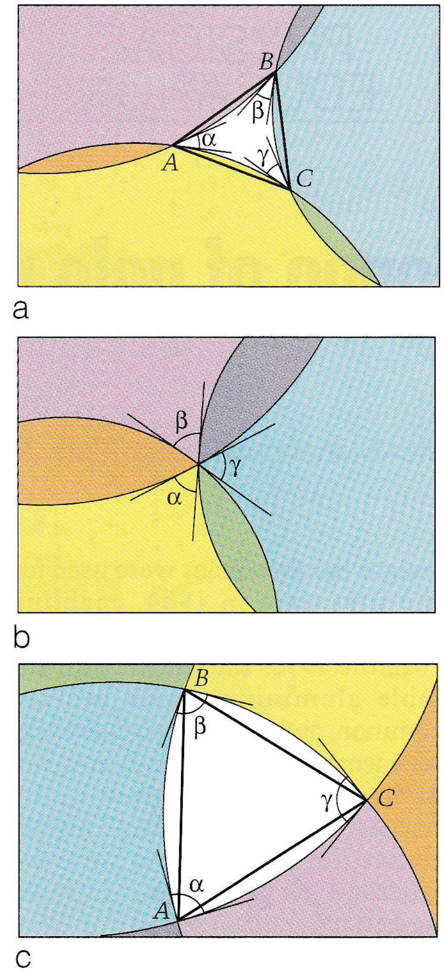


Figure 7

The same reasoning can be used if the designated vertices are cut off by several planes, rather than by a single one. The only condition is that the faces that will appear remain adjacent in pairs and don't intersect at the same vertex.

Similar technique can be used to obtain other sufficient conditions. At present, some necessary conditions for a polyhedron to be absolutely uninscribable are known. However, the problem as a whole remains open.

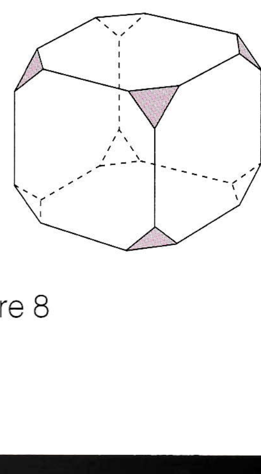


Figure 8

The method of relative angles provides a tool for solving other interesting problems.

Problem 7. *Suppose three faces meet at every vertex of a polyhedron, and every face has an even number of sides. Prove that if all but one vertex lie on a sphere, the polyhedron is inscribed.*

Hint: Use problem 6. Don't neglect the following unpleasant possibility: Edges emerging from a vertex not lying on the sphere might be tangent to it.

Here is a similar but more difficult problem.

Problem 8. *Let all the vertices of a polyhedron be colored black or white as described in problem 5, and let the number of black and white vertices be equal. Prove that if all but one vertex lie on a sphere, the polyhedron is inscribed.*

Thus, if the vertices of a polyhedron are colored black or white as described in problem 5, this polyhedron is either absolutely uninscribable or satisfies problem 8. That is, the fact that all its vertices but one lie on a sphere implies that the polyhedron is inscribed.

The simplest example of an absolutely uninscribable polyhedron is a cube with a vertex cut off (figure 1).

In conclusion, we should point out that Steinitz proved not the theorem named after him, but a related one, which we'll formulate as a problem.

Problem 9. *Assume that all the faces of a polyhedron can be colored black or white so that*

(1) the number of black faces is greater than the number of white ones;

(2) no two black faces are adjacent.

Prove that this polyhedron cannot be circumscribed about a sphere.

Figure 8 shows one of the simplest examples of an absolutely uncircumscribable polyhedron—a cube with all its vertices cut off. ◘

The science of pole vaulting

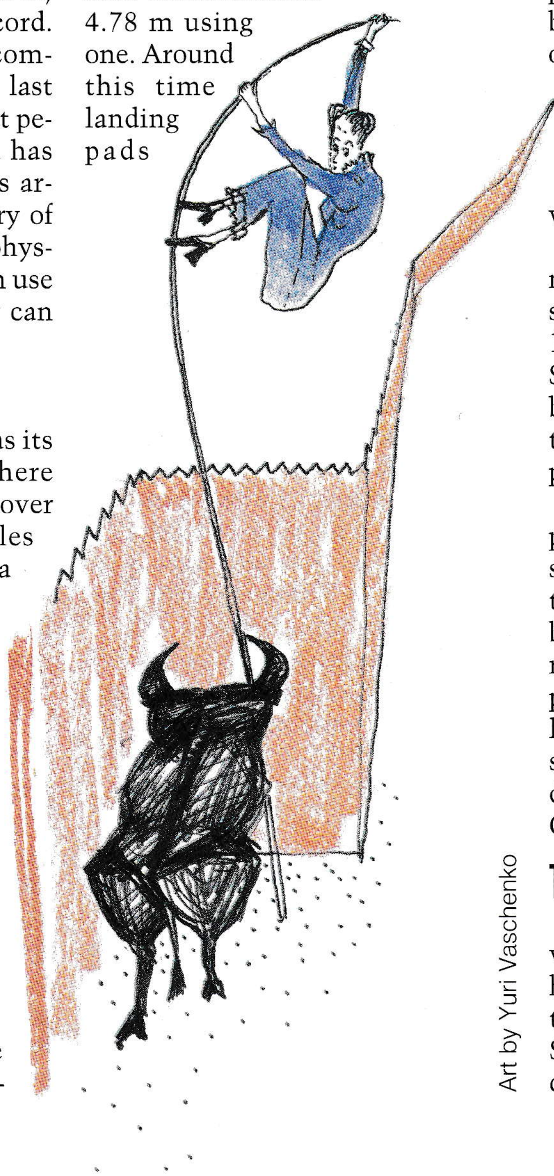
by Peter Blanchonette and Mark Stewart

AT THE OLYMPIC GAMES IN Sydney, Australia, women competed in the sport of pole vaulting for the first time. American Stacey Dragila, the current world record holder, won the gold medal with a height of 4.60 m, just 3 cm below her world record. Women have only seriously competed in pole vaulting over the last five or so years, but in this short period of time the world record has increased dramatically. In this article we will discuss the history of pole vaulting, and some of the physics behind the event that we can use to determine how high Stacey can fly!

History

The sport of pole vaulting has its origins in ancient Greece, where long poles were used to vault over charging bulls. In Europe, poles made of ash were used as a means to cross canals without getting wet. The sport of pole vaulting as it is known today began in the late 1800s when competitors began vaulting for height rather than distance, and they climbed the pole as they vaulted. In 1889 the movement of the hands up the pole was outlawed and the technique of swinging the legs upward, clearing the crossbar with the stomach facing down, was employed, similar to the technique used today. Lightweight bamboo poles were used for the first time in 1904, enabling competitors to jump higher.

In the 1950s the use of more durable aluminium poles became common, and in 1957 Bob Gutowski set a world record of 4.78 m using one. Around this time landing pads were introduced, which improved safety for the competitors. Prior to this the landing material was a combination of sand and wood shavings, making it necessary to land feet first. The fiberglass pole came to prominence at the 1956 Olympics, but a world record was not set using one until 1961. The introduction of fiberglass has been the most significant breakthrough in the sport. This can be seen in the rapid progression of the men's world record in the 1960s (figure 1).



Art by Yuri Vaschenko

American vaulters have dominated the Olympics, their winning streak extending from 1896 until 1968. More recently, Ukranian Sergey Bubka has ruled the event, breaking the world record numerous times and winning six world championships.

Although women's pole vault performances have been recorded since 1911, the International Amateur Athletic Federation has only been ratifying the women's world record since 1995. The women's pole vault has been contested at the last two World Indoor Championships, the most recent World Outdoor Championship, and the Sydney Olympics.

The rise of Stacey Dragila

Stacey Dragila came to pole vaulting from a background in heptathlon (a combination of seven track and field events) at Idaho State University. While a junior in college, Dragila, at the urging of coach Dave Nielsen, tried to vault over 6 feet. While she admits it took her many jumps before she felt comfortable, her improvement has been rapid! With her background in heptathlon, she had a head start on her competitors in terms of the physical requirements of pole vaulting, speed and upper-body strength. Combining this with her competitive instincts, she has won four U.S. outdoor championships. At 170 cm and 64 kg, she is very similar physically to Emma George, who has broken the world record 15 times. Stacey quickly improved and won the first major women's pole vault contest, the world indoor championship in 1997, defeating then world-record holder Emma George. And, of course, she has gone down in history as the first female Olympic champion in pole vault.

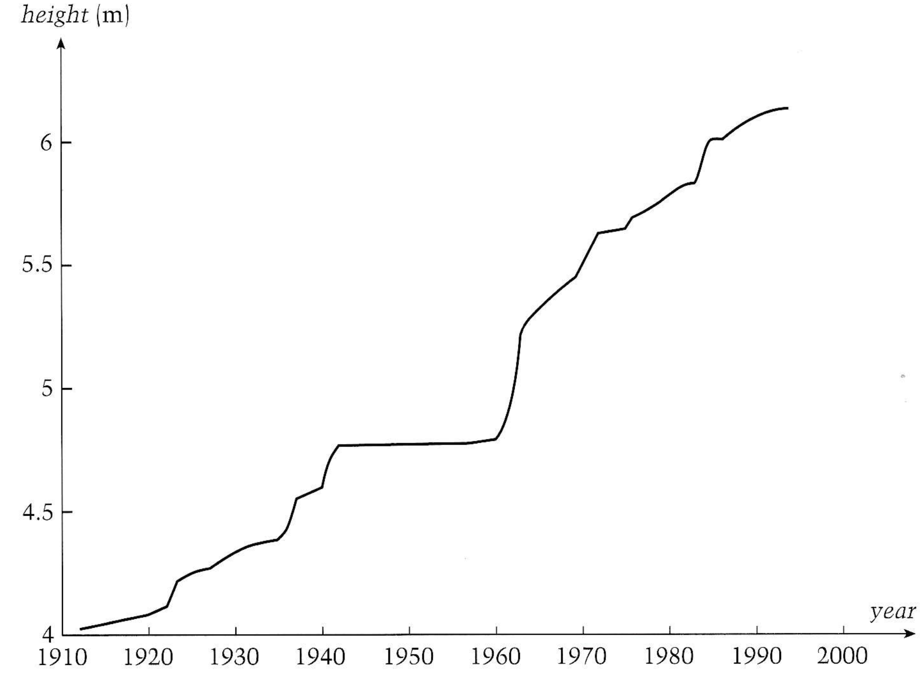


Figure 1

How high can Stacey fly?

Given that women's pole vaulting is a new event, a question to consider is how close the current women's world record is to a "real" world record. To estimate this we can use the current world records in the men's and women's long jump, as well as the men's world record in the pole vault:

Men's long jump world record: 8.95 m (Mike Powell, USA)

Women's long jump world record: 7.52 m (Galina Chistyakova, URS)

The ratio of the women's world record to the men's is 7.52/8.95 = 0.84. Using this ratio and the world record for the men's pole vault, 6.14 m, set by Sergey Bubka in 1994, we can estimate the true world record in the women's event.

$$\text{"real" world record} = 0.84 \times 6.14 \text{ m} = 5.15 \text{ m}$$

Allowing for the fact that the strength of women's upper bodies is proportionally less than their lower bodies compared to men, we can reduce this figure to about 5.00 m. This is about 40 cm above Stacey's current mark. Since Stacey has only been vaulting for a short period, it is possible that she may keep improving. Another question to consider is "Can we estimate the greatest height Stacey could vault?" The answer is yes, using some simple physics. Given that we know the athlete's speed at take-off, we can estimate their maximum vault. Typically, vaulters will use a run-up of between 12 and 16 strides, and as they approach take-off their speed gradually increases, as does their kinetic energy. Kinetic energy is defined as $KE = {}^1/_2 mv^2$, where v is the athlete's speed and m is the mass of the athlete.

As the vaulter reaches the take-off point, she lowers the pole into the box and the pole bends as the kinetic energy is transferred to the pole. The pole then begins to straighten, returning the energy to the vaulter in the form of gravitational potential energy, so that at the peak of her flight she has potential energy of $PE = mgh$, where g is the acceleration due to gravity (9.8 m/s^2) and h is height of the athlete above the ground.

If we assume all the kinetic energy is transformed to potential energy, we can estimate the maximum height of the athlete.

$$h_{max} = v^2/2g$$

Stacey's take-off speed has been measured at 8.3 m/s. Putting this into the equation we see that her center of gravity rises 3.7 m. However, at take-off Stacey's center of gravity is about 1 m from the ground, and due to the techniques she uses her center of gravity actually passes under the bar by around 20 cm. Also, being a good gymnast, Stacey can pull herself up into a handstand position, adding approximately another 70 cm. Taking all these factors into account, we see that Stacey's maximum height is approximately 5.6 m. Obviously there are some mechanical energy losses (transferred to heat, for example), so Stacey couldn't jump this high. However, this result shows she still has room for improvement.

This result shows us that the ability to run fast is vital to a pole vaulter's success, with the height vaulted being proportional to the square of their take-off speed. One thing is certain: The women's world record will continue to increase rapidly for the next several years.

Dr. Peter Blanchonette is an applied mathematician with a keen interest in the physics of sport. Dr. Mark Stewart is a senior lecturer in economics. He coached Emma George to 10 world records from 1994 until the end of 1997.

ANSWERS, HINTS & SOLUTIONS

Physics

P321

If the piece of chalk acquired a speed v_0 during the time interval τ, its displacement on the board is

$$L_1 = \frac{v_0^2}{2\mu g}.$$

When the board is stopped abruptly, the chalk moves the same distance backward and stops at the starting point. So in this case the length of the line drawn is L_1.

If the period τ is too short for the chalk to come to rest on the moving board, the result will be different. The length of the line drawn before the board comes to a full stop is

$$L_2 = v_0\tau - \frac{\mu g \tau^2}{2}.$$

At the instant the board stops, the chalk has gained a speed $v = \mu g \tau$ relative to the laboratory reference system. After that, the chalk will continue on its way with a gradually decreasing speed until it stops completely. The length of this path is

$$L_3 = \frac{1}{2} v \cdot \frac{v}{\mu g} = \frac{v^2}{2\mu g}.$$

Clearly this value is smaller than L_2, so in this case the length of the line drawn is L_2.

P322

Clearly no part of the planet can move with a speed greater than its escape velocity. The escape velocity is determined by the equation

$$\frac{v_{\text{esc}}^2}{R} = \frac{GM}{R^2},$$

where M and R are the mass and radius of this extremely dense planet.

Therefore, in order to hold onto the mass at the planet's equator, the equatorial speed must be less than the escape velocity:

$$v_{\text{eq}} \le v_{\text{esc}} = \left(\frac{GM}{R}\right)^{1/2}.$$

The rotational period is

$$T = \frac{2\pi R}{v_{\text{eq}}},$$

so

$$T \ge \frac{2\pi R}{v_{\text{esc}}} = 2\pi\left(\frac{R^3}{GM}\right)^{1/2}.$$

Taking into consideration the mean density

$$\rho = \frac{M}{\left(\frac{4\pi R^3}{3}\right)},$$

we get

$$\rho \ge \frac{3\pi}{GT^2}$$

and

$$\rho \ge 1.09 \cdot 10^6 \text{ kg/m}^3.$$

P323

Initially the small (inner) sphere was not charged. After the two spheres are connected with a conducting wire, a certain charge q_1 flows to the small sphere, where it is nonuniformly distributed over the surface. At the same time, another charge q_2 appears on the internal surface of the large (outer) sphere, and it is also nonuniformly distributed. The lines of force of the field generated by the point charge q terminate at the small sphere and at the internal surface of the large sphere. Thus the charge q_2 must be

$$q_2 = -(q + q_1).$$

Since these lines of force do not penetrate the large sphere and do not escape from the system, the charge on the external surface of the large sphere is distributed uniformly, so it doesn't generate an electric field anywhere inside the large sphere. Equating the potentials of the two spheres (these potentials are equal since the spheres are connected by a conductor), we can neglect the charge at the external surface of the large sphere. Therefore, the potential of the large sphere can be taken as equal to zero. Thus the potential at the center of the system is also zero (since there is no field inside the small sphere, the potential of the small sphere is equal to the potential at the center of the system). This gives us

$$\frac{q_1}{4\pi\varepsilon_0 a} + \frac{q_2}{4\pi\varepsilon_0 b} + \frac{q}{4\pi\varepsilon_0 c} = 0.$$

Plugging $q_2 = -(q + q_1)$ into this equation, we get

$$q_1 = -q\frac{\frac{1}{c} - \frac{1}{b}}{\frac{1}{a} - \frac{1}{b}}.$$

The value of q_1 doesn't depend on the initial charges of the spheres, since it's determined only by the value of q and the geometrical parameters of the system. All "extra" charge will be located at the external surface of the large sphere, and it will determine the potential of the whole conductor. It's interesting to note that when the charge q moves (more strictly, when the distance c varies), q_1 (and hence q_2 and q_{ext}) can vary. Therefore, in this case, the external field may also vary. In contrast, the field inside such a spherical "screen" doesn't depend on the value and location of the external charges.

P324

While the external magnetic field was fading, an emf was generated in the circuit by two processes: first, by the change in the external magnetic flux moving though the coil, and second, by the change in the flux generated by the coil itself (self-induced emf).

Since the resistors R_1 and R_2 are connected in parallel, the currents flowing through them obey the equation $I_1(t) \cdot R_1 = I_2(t) \cdot R_2$. Therefore, immediately after the external field is switched off, the current flowing through resistor R_2 is

$$I_2 = \frac{IR_1}{R_2}.$$

The electric current I_L flowing in the coil immediately after the magnetic field is cut off equals

$$I_L = I + I_2 = I\frac{R_1 + R_2}{R_2}.$$

Now the system is isolated. At first, all its energy was in the magnetic field of the coil:

$$U = \frac{LI_L^2}{2} = \frac{LI^2}{2}\frac{(R_1+R_2)^2}{R_2^2}.$$

Subsequently, this energy is dissipated by the resistors. Since the voltage drop across the resistors is the same, the amount of heat dissipated by them is inversely proportional to their resistance:

$$Q_1 \sim \frac{1}{R_1} \text{ and } Q_2 \sim \frac{1}{R_2}.$$

Thus we obtain the following system of equations:

$$\begin{cases} Q_1 + Q_2 = \dfrac{LI^2}{2}\dfrac{(R_1+R_2)^2}{R_2^2} \\ Q_1/Q_2 = R_2/R_1, \end{cases}$$

which yields

$$Q_1 = \frac{LI^2}{2}\frac{(R_1+R_2)}{R_2}$$

and

$$Q_2 = \frac{LI^2}{2}\frac{(R_1+R_2)R_1}{R_2^2}.$$

P325

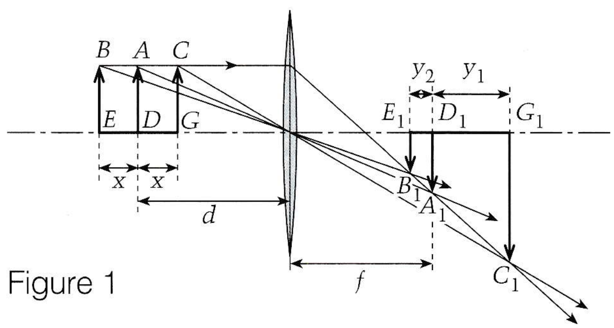


Figure 1

Denote the distance between the center spike of the trident and the lens by d and the distance between its image and the lens by f (figure 1). Let's construct the images of the spikes BE, AD, and CG. Denote the length of segments ED and DG by x, the distance between the images A_1D_1 and C_1G_1 by y_1, and that between the images B_1E_1 and A_1D_1 by y_2. Prove on your own that the magnification of segment AD is

$$\Gamma = \frac{F}{d - F},$$

where F is the focal length of the lens. This yields

$$d = \frac{F(1+\Gamma)}{\Gamma},$$

and

$$f = F(1+\Gamma).$$

Using the lens formula for the spikes CG and BE, we get

$$\frac{1}{d \mp x} + \frac{1}{f \pm y} = \frac{1}{F},$$

where the upper signs in the denominators relate to spike CG, while the lower signs correspond to dent BE. Obvious transformations result in

$$\frac{xy_1}{F} = \frac{y_1}{\Gamma} - x\Gamma$$

and

$$\frac{xy_2}{F} = -\left(\frac{y_2}{\Gamma} - x\Gamma\right).$$

Since the magnifications of segments DG and DE are β_1 and β_2, respectively, we obtain the magnification of segment

$$\Gamma = \sqrt{\frac{2\beta_1\beta_2}{\beta_1 + \beta_2}}.$$

Math

M320

(a) Let $A_1A_2 \ldots A_{12}$ be a regular 12-gon (figure 2). Consider the triangle $A_2A_4A_8$. The lines A_2A_6, A_3A_8, and A_4A_{11} are bisectors of its angles. Similarly, A_3A_8, A_5A_1, and $A_{11}A_4$ are bisectors of the angles of triangle $A_3A_5A_{11}$. This implies that diagonals A_1A_5, A_2A_6, A_3A_8, and A_4A_{11} meet at a point.

(b) Consider the regular 18-gon $A_1A_2 \ldots A_{18}$ with vertices selected from the vertices of the given 54-gon. Its diagonals A_1A_8, A_2A_9, A_4A_{12}, and A_6A_{16} meet at a point. To

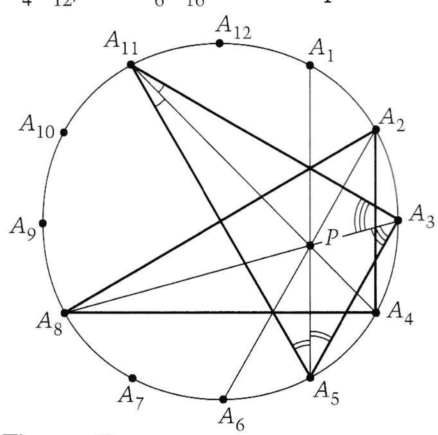


Figure 2

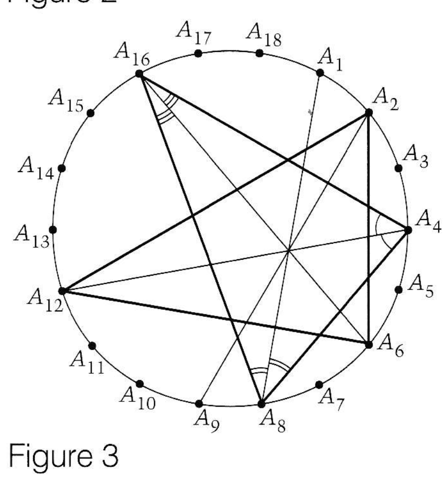


Figure 3

prove this fact, we apply the argument of part (a) to triangles $A_2A_6A_{12}$ and $A_4A_8A_{16}$ with bisectors A_2A_9, A_6A_{16}, $A_{12}A_4$ and A_4A_{12}, A_8A_1, $A_{16}A_6$, respectively (figure 3).

It's interesting to find values of n for which there exist four diagonals of a regular n-gon that meet at a point different from the center of this polygon. Does a regular n-gon exist that has five diagonals that meet at a point different from the center of this polygon?

M321

The answer is $k = 1991$.

If $k \le 1990$, it may happen that the first 10 deputies suggest assigning nothing for the first item and assign $S/199$ for the other items. The next 10 deputies can suggest assigning nothing for the second item and assign $S/199$ for the other items, and so on. As a result, the sum $S/199$ would be approved for every item, and the total budget would amount to

$$\frac{200}{199}S > S.$$

If $k = 1991$, then for every item, there were at most ten deputies who suggested an expenditure less than the approved one. Therefore, for every item, there exists a deputy who suggested an expenditure not less than the approved one. However, the budget proposed by this deputy was not greater than S. Thus the budget approved does not exceed this amount either.

M322

Let $f(x) = x^3 - 3x^2 + 5x = (x-1)^3 + 2(x-1) + 3$, and consider the function $g(y) = y^3 + 2y$. Then the first equation can be written as $g(\alpha - 1) = f(\alpha) - 3 = -2$, and the second as $g(\beta - 1) = f(\beta) - 3 = 2$. A little algebra will show that $g(y)$ is monotone increasing and odd, and it follows that $\alpha - 1 = -(\beta - 1)$. This implies that $\alpha + \beta = 2$.

M323

We use the inequality

$$\frac{1}{a} + \frac{1}{b} \ge \frac{4}{a+b} \quad (a>0,\ b>0)$$

This inequality implies that the sum S of the inverse of the numbers written on the blackboard does not increase after each operation. Initially it was equal to n. Therefore, at the end of the process we will have $S \le n$. This means that the remaining number satisfies the inequality $1/S \ge 1/n$.

To prove the original inequality, note that $(a-b)^2 \ge 0$, so $(a+b)^2 = a^2 + 2ab + b^2 \ge 4ab$. We obtain the required inequality after dividing by $ab(a+b)$, which is positive since a and b are.

M324

The original plan of the investigator can be modified by including additional questions according to the following scheme. First, the investigator asks 13 questions following his original plan. Then he puts the question: "Did you give a false answer to one of the questions in the preceding series?" If the witness answers no, the sets consisting of 12, 11, ..., 2, and 1 questions of the original plan are asked, and the test question is asked after each series. If the answer to one of the test questions is yes, the corresponding series of questions is repeated; then the interrogation follows the original plan without test questions (remember that the witness may give only a single false answer).

Suppose the answer yes was given to the kth test question. Then $k + 14 - k = 14$ additional questions were asked, compared to the original plan. Thus the modified plan guarantees that the investigator can reveal the truth using 105 questions.

Assume that the original plan consists of N questions and one of the answers may be false. Then the modified plan guarantees that the truth will be revealed after $N + q$ questions, where q is the minimum number for which $N \le q(q-1)/2$.

It would be interesting to learn if this number of questions is the minimum possible one and to analyze a similar problem in which k false answers are allowed.

Brainteasers

B321

The story is saying that a 72 cent decrease in the price of one piece results in a 3 cent decrease in the average price. Since $72/3 = 24$, this must be the number of peaches.

More formally, let A be the sum of the original prices (in cents) of all the peaches, and let n be the number of peaches. Then A/n is the original average price of a peach, $(A - 72)/n$ is the new average price, and $A/n = 3 + (A-72)/n$. Simplifying, the A's drop out, and we find that $n = 24$.

B322

In figure 4, right triangles NAD and MBA are congruent (their legs are equal in pairs), so $\angle NAD = \angle MBA$. Similarly, right triangles MCD, NBC are congruent, so $\angle MCD = \angle NBC$. So the sum $\angle ABM + \angle MBN + \angle NBC = \angle ABC = 90°$ is equal to the required sum of the three subtended angles.

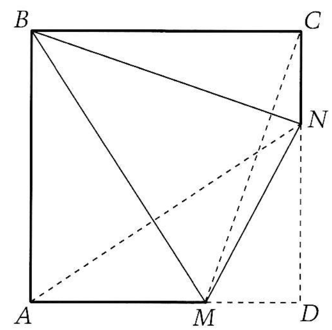


Figure 4

B323

Take any one player, and suppose he won w games playing white and b games playing black. Then the number of all the victories of other players playing black is also w, and so $b + w$ is the total number of victories playing black, including our singled-out player. But this number is constant, independent of the particular values of b and w. That is, this number is independent of the original choice of player to observe, so every participant won this many games.

B324

The numbers assigned to the apartments were 9, 10, 11, and 12. The digits 0, 1, 1, 1, 1, 2, and 6 were ordered. Digit 6 was put upside down on the door to apartment 9. The total sum paid was 1 + 1 + 1 + 1 + 2 + 6 = 12 dollars.

B325

$$3 \times 29 + 13 = 100.$$

Kaleidoscope

1. Earth's magnetic field has a vertical component. In the Northern hemisphere, the north pole will be at the lower end of the bar, and the south pole at the upper end.

2. No, because the effects of external magnetic fields, vibrations, and sharp changes in temperature promote the demagnetization of permanent magnets.

3. A sufficiently strong magnetic field will magnetize a ferromagnet of any shape.

4. As steel is heated and its temperature approaches the ferromagnetic Curie point, its magnetic permeability decreases, so the hot steel is poorly magnetized and weakly attracted by the magnet.

5. Electric current will flow from the central wire to the ring. The ions moving in the solution will be affected by the magnetic field. As a result, the entire liquid will rotate clockwise (if we look down from above).

6. Free electrons in the metal, moving in the magnetic field, are moved by the Lorentz force to one of the faces of the bar. Therefore, an electric field will be generated perpendicular to the direction of the bar's velocity. This is the Hall effect.

7. In contrast to electric lines of force, the lines of force of the magnetic field do not terminate at the surface of the screen. Therefore, a screen can only diminish but not eliminate the magnetic field—and even this effect requires a rather thick screen.

8. The oscillating needle generates an alternating magnetic field, which induces eddy currents in the brass case. These currents consume energy from the needle and damp its oscillations.

9. In order to decrease the induced Foucault (eddy) currents, which decrease the transformer's efficiency.

10. The inductance depends on the magnetic permeability of the core, which in its turn depends on the magnetic field generated by the current in the coil. Therefore, the inductance depends on the current in the coil with an iron core.

11. (a) It will increase many times over; (b) it will increase slightly; (c) it will decrease slightly.

12. The paramagnetic liquids are pulled into the region of the stronger field, while the diamagnetic liquids are expelled from it.

13. The gases formed during combustion (carbon dioxide and carbon monoxide) are diamagnetic substances.

14. (a) Polarization will occur; (b) a short-term induced current will appear; (c) a long-term induced current will appear.

15. The magnetic flux in the superconducting circuit cannot change (otherwise the infinite current must be driven by the induced emf in the infinitely conducting circuit). Since the area of the circuit decreased by a factor of four, the magnetic field increased by the same factor.

Microexperiment

In the flame the iron nail loses its magnetic properties. When cooled, it regains its magnetizing capacity.

Physics Contest

Relativistic conservation laws

In the Contest Problem in the November/December 2000 issue of *Quantum*, we asked our readers to solve a problem using the relativistic forms of the laws of conservation of energy and momentum. A relativistic particle decays into two photons. One of the photons travels along the positive x-axis with frequency f_1, while the second photon travels along the negative x-axis with frequency $f_2 < f_1$.

The relativistic energy and momentum of a particle are given by

$$E = \gamma mc^2$$

and

$$p = \gamma mv,$$

respectively, with

$$\frac{AB_1}{B_1C} \cdot \frac{CA_1}{A_1B} \cdot \frac{BC_1}{C_1A} = 1$$

and

$$\beta = \frac{v}{c},$$

where c is the speed of light in a vacuum. For a photon, we have

$$E_\gamma = hf$$

and

$$p_\gamma = \frac{hf}{c}.$$

A. We now apply energy and momentum conservation to the decay process.

$$\gamma mc^2 = hf_1 + hf_2 \qquad (1)$$

$$\gamma mv = \frac{hf_1}{c} - \frac{hf_2}{c}. \qquad (2)$$

Let's now divide equation (2) by equation (1), cancel common factors, and solve for the velocity:

$$v = c\frac{f_1 - f_2}{f_1 + f_2}. \qquad (3)$$

B. We now solve equation (1) for the mass:

$$m = \frac{h}{\gamma c^2}(f_1 + f_2) = \frac{h}{c^2}(f_1 + f_2)\sqrt{1 - \beta^2}$$

and substitute for β from equation (3) to obtain

$$m = \frac{2h}{c^2}\sqrt{f_1 f_2}. \qquad (4)$$

C. In the rest frame of the particle the two photons must travel in opposite directions with the same size momentum. Therefore, the frequency f will be the same for both photons. Energy conservation requires

$$mc^2 = 2hf$$

or

$$f = \frac{mc^2}{2h}. \qquad (5)$$

D. In this part we are to determine the functions F_1 and F_2 in the expression

$$p_x' = F_1 p_x + F_2 \frac{E\gamma}{c},$$

where the unprimed variable refers to the rest frame of the particle and the primed variable refers to the laboratory frame. Writing this equation for the first photon, we obtain

$$\frac{hf_1}{c} = F_1 \frac{hf}{c} + F_2 \frac{hf}{c} = (F_1 + F_2)\frac{hf}{c}.$$

Using equation (5), this simplifies to

$$F_1 + F_2 = \frac{2hf_1}{mc^2}. \qquad (6)$$

For the second photon we obtain

$$-\frac{hf_2}{c} = -F_1 \frac{hf}{c} + F_2 \frac{hf}{c} = (-F_1 + F_2)\frac{hf}{c},$$

$$F_1 - F_2 = \frac{2hf_2}{mc^2}. \qquad (7)$$

Adding equations (6) and (7) and dividing by 2, we get our expression for F_1:

$$F_1 = \frac{h}{mc^2}(f_1 + f_2).$$

Comparing this to equation (1), we see that

$$F_1 = \gamma.$$

Subtracting equation (7) from equation (6) and dividing by 2 yields

$$F_2 = \frac{h}{mc^2}(f_1 - f_2).$$

Comparing to equation (2) shows that

$$F_2 = \beta\gamma.$$

A good theory

In the January/February 2001 issue of *Quantum*, we asked readers to solve problems emerging from the profound theories of Newton and Bohr. Once again, Art Hovey from Amity Regional High School of Connecticut submitted solutions that were basically correct.

In the first problem, readers were asked to show that if a moon of mass m orbits a planet of mass M closer than a specified distance, loose rocks lying on the surface of the moon will be lifted from the surface.

The gravitational and normal forces acting on a rock of mass μ lying on the surface of the moon provide the centripetal force that keeps it in orbit about the planet. Let's assume that the rock is located on the side of the planet facing the moon, that the radius of the moon is r, and that the distance between the moon and the planet is d. We also assume that the mass of the planet is very much larger than the mass of the moon.

$$\frac{GM\mu}{(r-a)^2} + F_N - \frac{Gm\mu}{a^2} = \mu\omega^2(r-a). \quad (1)$$

The gravitational force of the planet acting on the moon causes the moon to revolve about the planet with the same angular velocity.

$$\frac{GMm}{r^2} = m\omega^2 r. \qquad (2)$$

Solving equation (1) for ω and substituting into equation (2), we get

$$\frac{GM\mu}{(r-a)^2} + F_N - \frac{Gm\mu}{a^2} = \mu(r-a)\left(\frac{GM}{r^3}\right).$$

The loose rock will leave the surface whenever the normal force is negative. The limiting case is found by setting the normal force equal to zero.

$$\frac{GM\mu}{(r-a)^2} - \frac{Gm\mu}{a^2} = \mu(r-a)\left(\frac{GM}{r^3}\right).$$

Recognizing that the distance between the planet and the moon is much greater than the radius of the moon ($r \gg a$), we can ignore the terms of a that are added to much larger values of r.

$$Ma^2r^3 - mr^5 = Ma^2(r^3 - 3r^2a),$$

$$r = a\sqrt[3]{\frac{3M}{m}}.$$

The second problem described an inelastic collision between two hydrogen atoms. For this inelastic collision, the two hydrogen atoms "stick" together, forming a diatomic molecule. Momentum must be conserved.

$$m_1 v_0 = 2m_1 v_f.$$

The loss in kinetic energy can now be calculated.

$$\Delta K = K_f - K_0$$

$$= \frac{1}{2}(2m)\left(\frac{v_0}{2}\right)^2 - \frac{1}{2}(m)v_0^2$$

$$= -\frac{1}{4}mv_0^2 = -\frac{E_0}{2}.$$

Where did the energy go? With billiard balls, the energy of an inelastic collision may be transformed into sound, deformation of the objects, or heat, but these don't make sense at the atomic level. The energy must have raised an electron to a higher energy state. The smallest energy change for a ground state electron in hydrogen can be calculated:

$$\Delta E = E_2 - E_1 = \frac{E_1}{2^2} - E_1 = -\frac{3E_1}{4}$$

$$= -\frac{3(2.18 \cdot 10^{-18}\ \text{J})}{4} = 1.63 \cdot 10^{-18}\ \text{J}.$$

Setting these energy differences equal to one another, we can solve for the initial velocity of the hydrogen atom.

$$v_f = 3.13 \cdot 10^4\ \text{m/s}$$

Since the diatomic molecule is moving with a speed $v_0/2$, the frequency of the emitted photon will be Doppler shifted. For speeds that are small relative to the speed of light, the fractional change in the frequency is approximately equal to the ratio of the speed of the molecule to the speed of light.

$$\frac{\Delta f}{f} \cong \frac{v}{c} = \frac{6.26 \cdot 10^4\ \text{m/s}}{3 \cdot 10^8\ \text{m/s}} = 0.021\%.$$

The frequency is larger if the photon is emitted in the forward direction and smaller if emitted in the backward direction.

—Larry D. Kirkpatrick and Arthur Eisenkraft

Musical chairs

by Don Piele

AT SOME TIME IN YOUR CHILDHOOD YOU have probably played a version of the game called Musical Chairs. In this game, as I played it in elementary school, a group of n kids dance around a set of $n - 1$ chairs, skipping to the music. When the music stops, everyone scrambles to sit down—one person to a chair. Since there are more kids than seats, someone is left standing, and that person is out of the game. Then everyone gets up, one chair is removed from the circle, and the music begins again. The game is repeated until only one chair is left and one person is sitting in it. That person is the winner.

This game is an entertaining way of selecting at random a person from a group of n, since everyone has an equal chance of ending up the winner. Let's change the rules and invent a new game of Musical Chairs. Suppose this time everyone brings their own chair to the game and the chairs are all different. This time we won't take away a chair but instead we'll require that when the music stops everyone must sit down in a chair *different* from their own. Anyone who can't do that takes themselves and the chair they are sitting in, which of course is their own, out of the game. The music begins again and all the remaining players repeat the process with the same rules—you can't sit down in your own chair or you are out!

There are two quick observations about this game. It can never be the case that you have only one person left not sitting in his own chair, since for n players, if $n - 1$ are sitting in their own chairs then the last person must also be sitting in his/her own chair. And on any given round of play, it can happen that everyone finds a different chair and no one is removed. We know that in the original game, where a chair is removed at each play, with n players the game is over in n plays of the music. But what about our new version of Musical Chairs? On average, how long will it take to end this game? The game ends when there are no chairs left.

Simulation

Armed with a computer loaded with *Mathematica*, let's take a look at how to answer this question easily via simulation. Suppose the number of chairs and players is 12. We first assign a number to each of the twelve chairs.

```
n=12;
chairs=Range[1,n]
{1,2,3,4,5,6,7,8,9,10,11,12}
```

We assume that the game is played in such a way that when the music stops, a random permutation of the players {1,2, ..., 12}, is chosen. Any number that does not change position is removed. For example, with the permutation {2, 1, 4, 3, 5, 6, 7, 8, 9, 10, 11, 12}, the eight numbers 5–12 would be removed. So the length of a game with the new rules is a random variable, and we are after its average or expected value.

To generate our random permutations, we will load the *Mathematica* package "Discrete Permutations."

```
<<DiscreteMath`Permutations`
```

Now we can use a simple function, Random Permutation, to generate one play of our new Musical Chairs.

```
stopTheMusic=RandomPermutation[12]
{3,5,8,7,4,6,1,9,12,2,11,10}
```

The question is, which of these numbers are in their own positions? One way to identify them easily is to subtract the position numbers {1, 2, 3, ..., 12} from the corresponding random permutation numbers shown above and see if we have any zeros. The set {1, 2, 3, ..., 12} is generated in *Mathematica* by **Range[12]**.

```
stopTheMusic-Range[12]
{2,3,5,3,-1,0,-6,1,3,-8,0,-2}
```

In this play of the game, two zeros appear at positions 6 and 11, meaning that these players are in their own chairs and must be taken out of the game. One simple way to do this in *Mathematica* is with a replacement rule.

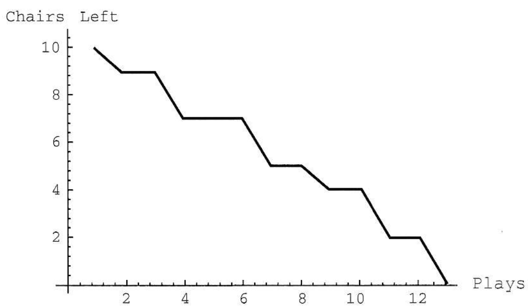
```
stopTheMusic-Range[12] //.{a___,0,b___}
->{a,b}
{2,3,5,3,-1,-6,1,3,-8,-2}
```

Now we have reduced the chair size down to 10 and can start the music again. We repeat this process until all the chairs are gone. Let's simulate one complete game and keep track of how many chairs are left at each step. We begin with 12 chairs.

```
Clear[a,b]

n=12;stage=0; chairsLeft={n};

While [n>0,
 sit= RandomPermutation[n]-Range[n];
 sit=sit //. {a___,0,b___}->{a,b};
 n=Length[sit];
 chairsLeft=Join[chairsLeft,{n}];stage++];

Print["Number of plays",stage]

Print["Chairs left after each play",
chairsLeft]
  Number of plays 13
  Chairs left after each play
{12,10,9,9,7,7,7,5,5,4,4,2,2,0}
```

Now we'll plot the number of chairs left after each play.

```
ListPlot[Rest[chairsLeft],PlotJoined→True,
 AxesLabel→{"Plays","Chairs Left"}]
```



These steps can be collected together into a new *Mathematica* function called musicalChairs.

```
musicalChairs[n_]:=
  Module[{stage=0,sit,m=n},
   While[m>0,sit=
RandomPermutation[m]-Range[m];
   sit=sit //. {a___,0,b___}->{a,b};
   m=Length[sit];stage++]; stage]
```

The function **musicalChairs[n]** returns the number of plays needed to end a simulated game that begins with n chairs. Let's try it for $n = 12$.

```
musicalChairs[12]
11
```

Now we are ready to simulate this game 1000 times and average the results. We'll do the experiment for $n = 12$.

```
experiment=
Table[musicalChairs[12],{1000}];

Apply[Plus,experiment]/1000//N
12.129
```

The answer rounds off to 12, the number needed to end the original game of Musical Chairs, in which one chair is taken away each time. That surprises me.

Probability

A more difficult task is to find the probability density function for this random experiment. Here the arguments can get a bit mind bending and may take a few readings to fully understand them. Our first goal is to develop the formula for computing the number of ways that n people can be arranged in n chairs so that k are sitting in their own chairs and $n - k$ are not. We will call this number $w[n, k]$. Let's see how to construct this number with recursion.

First we notice that

```
w[1,0]:=0

w[1,1]:=1,
```

because with {1} you cannot make an arrangement where 1 is not in his own chair. And with {1} there is one way to arrange a person in his own chair. Also we have

```
w[n_,n_-1]:= 0

w[n_,n_]:=1,
```

because for n people you cannot have $n - 1$ in their own chairs without having all n in their own chairs. And there is only one way all n can be in their own chairs namely $\{1, 2, 3, ..., n\}$.

The next relationship is the key that solves a big part of the problem.

```
w[n_,0]:=w[n,0]=(n-1)w[n-1,0] + w[n-1,1]
```

This says that if you know how many ways you can arrange $n - 1$ people so that none are in their own chairs, $w[n - 1, 0]$, then add one more chair to the game and exchange this chair with any one of the $n - 1$ chairs. You will end up with an arrangement of n chairs with no one in their own chair. Thus you have created $(n - 1)w[n - 1, 0]$ arrangements of n chairs, none of which has a person in his own chair. That's one way to do it. But you can also consider all the arrangements of $n - 1$ chairs where exactly one person is in his own chair, $w[n - 1, 1]$. Now add one more chair and exchange that chair with the person sitting in his own chair. This adds $w[n - 1, 1]$ more new arrangements, which completes the recursion.

Finally we can argue:

```
w[n_,j_]:=w[n,j]= Binomial[n,j]w[n-j,0].
```

This says that to find the number of ways j people are in their own chairs and $n - j$ are not, simply pick j of the n chairs, which can be done **Binomial[n, j]**, and multiply this by the number of ways you can arrange the remaining $n - j$ people with no one in their own chair, $w[n - j, 0]$.

With these recursion relationships established, we are ready to compute $w[i, j]$ for any i and $j \le i$.

```
Table[w[i,j],{i,1,6},{j,0,i}]//TableForm
```

0	1					
1	0	1				
2	3	0	1			
9	8	6	0	1		
44	45	20	10	0	1	
265	264	135	40	15	0	1

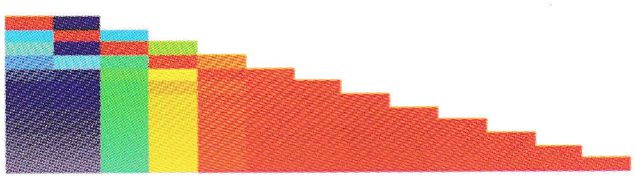
These numbers are transformed into probabilities by simply dividing each row by the total number of ways i people can be arranged in i chairs, which, of course, is $i!$.

```
Table[w[i, j]/(i!), {i, 1, 6}, {j, 0, i}]
// TableForm
```

0	1					
1/2	0	1/2				
1/3	1/2	0	1/6			
3/8	1/3	1/4	0	1/24		
11/30	3/8	1/6	1/12	0	1/120	
53/144	11/30	3/16	1/18	1/48	0	1/720

Each row is the probability that with i chairs, 0, 1, 2, ..., i people are sitting in their own chair when the music stops. Here is a graphical view of this probability distribution.

```
Show[Graphics[Table[{Hue[w[i, j]/(i!)*E
((i/(i + 3)))],
  Rectangle[{j, (-i)}, {1 + j, 1 - i}]}, {i,
1, 12}, {j, 0, i}]],
  AspectRatio→1/4]
```



The red indicates a low probability and the blue and purple the highest probability. Notice that the first two columns are the largest and that they measure the probability that 0 and 1, respectively, are in their own chairs after each play. It can be shown that both of these probabilities converge to $1/e$ as i increases.

How long will the music last?

We are ready to create the probability density function $p[m, k]$, which measures the probability that for a set of m chairs the number of plays of the new Musical Chairs game is k for $k = \{1, 2, 3, ...\}$. Once we know this distribution, we can calculate the expected number of plays and compare it with our simulation.

First we observe that the chances that with m chairs the game is over in one play is $1/m!$ Remember that there is only one way this can happen—$\{1, 2, 3, ..., m\}$.

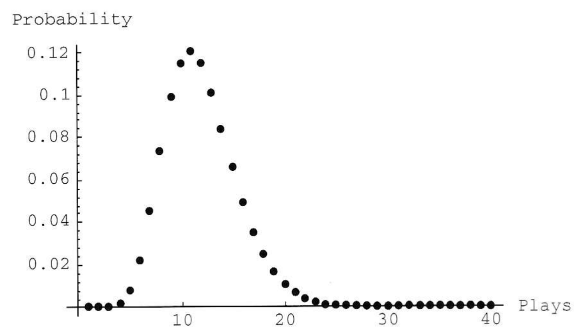
```
Clear[p]

p[m_, 1] := p[m, 1] = 1/(m!)
```

Now for the final observation. If we know the probability $p[j, k-1]$ that it takes $k-1$ plays to end the game with j chairs for $j = 2$ up to m, then we can compute the probability that it takes k plays from m chairs by multiplying $p[j, k-1]$ by the chances of going from m to j chairs, (by removing $m - j$ chairs), in one play which is $w[m, m-j]/m!$. This gives the recursive relationship:

$$p[m\_, k\_] := p[m, k] = \sum_{j=2}^{m} p[j, k-1] \frac{w[m, m-j]}{m!}$$

Here then is a picture of the distribution $p[12, k]$ for $k = \{1, 2, ..., 40\}$.

```
pdf=Table[p[12,k],{k,1,40}];

ListPlot[pdf,AxesLabel
→{"Plays","Probability"},PlotStyle→
PointSize[.02]]
```



The average number of plays to end the game is exactly 12.

$$\sum_{k=1}^{40} k\ p[12, k] //N$$

```
12.
```

Final Thoughts

This column was based on the examination of complete permutations or derangements, as they are also called, where no number in the permutation is left fixed. They are often studied in a first course in probability. However, the game of Musical Chairs that I suggested here is, to my knowledge, original. It took about a day to solve it. I would never have attempted an examination of it without *Mathematica* and the ease with which it allows me to program recursively. This seems to happen all the time. I begin with a simulation, and then move on to the search for the probability distribution. More often than not, the problem yields to force of reason applied with the power of *Mathematica*.