# QUANTUM

NOVEMBER/DECEMBER 2000

US $7.50/CAN $10.50

Springer

NSTA

*Oil on canvas, 35 1/4 × 40 1/8, Gift of Edgar William and Bernice Chrysler Garbisch, © 2000 Board of Trustees, National Gallery of Art, Washington, D.C.*

*Law of the Wild* (1881) by Charles Raleigh

THE ELEMENT OF SURPRISE IS CRUCIAL TO A successful hunt. Animals have many ways of masking their physical presence while stalking prey. The white coat of the polar bear, for example, allows it to blend in with the arctic landscape. In fact, polar bears have been observed covering their black noses with a paw while hunting for a meal, presumably to keep it from standing out like a sore thumb. The bear above seems to have avoided the watchful eyes and ears of the seal by swimming up silently from behind. The hunter may soon become the hunted, however, if that ship on the horizon is looking for bear skins to fill its hold. To explore other conditions that might allow you to stealthily approach your prey, turn to page 40.

# QUANTUM

NOVEMBER/DECEMBER 2000 VOLUME 11, NUMBER 2


*Cover art by Sergey Ivanov*

If someone were to walk a mile in your shoes, it is thought that they would soon share your experiences and outlook on life. If you were to walk along beside them, however, you might also come to share something else—a common velocity. But what velocity would that be and how would it be chosen? To get in step with the phenomena of group velocity, turn to page 47.

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.  Printed in U.S.A.

**Subscription Information:**
*North America*: Student rate: $18; Personal rate (nonstudent): $25. This rate is available to individual subscribers for personal use only from Springer-Verlag New York, Inc., when paid by personal check or charge. Subscriptions are entered with prepayment only. Institutional rate: $56. Single Issue Price: $7.50. Rates include postage and handling. (Canadian customers please add 7% GST to subscription price. Springer-Verlag GST registration number is 123394918.) Subscriptions begin with next published issue (backstarts may be requested). Bulk rates for students are available. Mail order and payment to: Springer-Verlag New York, Inc., Journal Fulfillment Services Department, PO Box 2485, Secaucus, NJ 07094-2485, USA. Telephone: 1(800) SPRINGER; fax: (201) 348-4505; e-mail: custserv@springer-ny.com.

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

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

Printed on acid-free paper.

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

# QUANTUM

THE MAGAZINE OF MATH AND SCIENCE

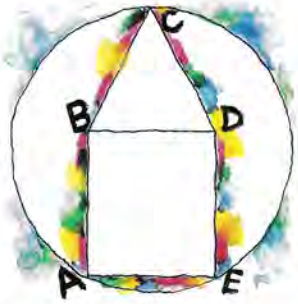
*A publication of the National Science Teachers Association (NSTA)*
*& Quantum Bureau of the Russian Academy of Sciences*
*in conjunction with*
*the American Association of Physics Teachers (AAPT)*
*& the National Council of Teachers of Mathematics (NCTM)*

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

# Just for the fun of it!

## B306

*Quad query.* Side $AD$ of quadrilateral $ABCD$ equals its diagonal $BD$. The three other sides of this quadrilateral are equal to each other. The diagonal $BD$ divides angle $ADC$ into two equal parts. What are the possible angular measures of angle $BAD$?


## B307

*Solve the circle.* Side $AE$ of pentagon $ABCDE$ equals its diagonal $BD$. All the other sides of this pentagon are equal to 1. What is the radius of the circle passing through points $A$, $C$, and $E$?

## B308

*Paycheck envy.* One hundred officials of a government agency were invited to a meeting. Chairs were arranged in a rectangle with 10 rows and 10 chairs in each row. The meeting didn't begin on time, and the officials started talking to their neighbors and exchanging information about their salaries. Those officials who learned that among their neighbors on the left, right, in front, behind, and on the diagonals, there was not more than one person with a higher salary, decided they themselves were highly paid. Which is the maximum possible number of highly paid officials?

## B309

*Meeting on the bridge.* Nick left Nicktown at 10:18 A.M. and arrived at Georgetown at 1:30 P.M., walking at a constant speed. On the same day, George left Georgetown at 9:00 A.M. and arrived at Nicktown at 11:40 A.M., walking at a constant speed along the same road. The road crosses a wide river. Nick and George arrived at the bridge simultaneously, each from his side of the river. Nick left the bridge 1 minute later than George. When did they arrive at the bridge?

## B310

*It's a breeze.* Once, when I returned from a weekend trip to the country, it was very stuffy in the railway car. Therefore, I stepped out onto the platform at the end of the car, but the air was not any better. When the train slowed down for a station, fresh air blew from an air vent. When the train stopped, the air quit flowing. However, the fresh air blew from the vent every time the train slowed down for a station. What caused the air flow? Was I standing on the platform at the front or the rear of the car? I should add that all windows and the door to the platform were open.

Art by Pavel Chernusky

*ANSWERS, HINTS & SOLUTIONS ON PAGE 53*

# The physics of chemical reactions

## *Don't overlook the interactions*

by O. Karpukhin

IN A CHEMICAL REACTION, one set of substances is converted into another set of substances. For example, burning is a chemical reaction that yields water and carbon dioxide gas. Mixing an acid and an alkali produces a salt. Usually when we study chemistry we "close our eyes" while the molecules are interacting and only "open our eyes" again when the products of the chemical reaction have already been formed. And yet it is the processes occurring when the molecules interact that determine the composition of the products of the chemical reaction.

Any chemical reaction consists of two basic stages. First, the reacting particles must meet. Second, a chemical conversion occurs, whereby quantum shells are rearranged and new molecules are formed out of the original ones. Chemical physics is thus subdivided into two major areas: chemical kinetics and elementary event theory.

Chemical kinetics studies how the reacting particles meet, what external forces affect them, and what equations describe the changes in the concentration of the reacting substances in the course of a reaction. In addition, it studies how the rate of a chemical process depends on the concentrations of the reagents (the original substances involved in a chemical reaction), the temperature, and other conditions under which a reaction proceeds.

The elementary event theory of chemical conversion investigates the very process of interaction of the colliding particles as well as changes in the quantum shell configuration and interatomic distances within the molecules.

This article is devoted to chemical kinetics. The major problem of chemical kinetics is the incidence of meetings (that is, collisions) of the reacting particles.

The particles participating in a reaction are not necessarily molecules. Neutral atoms, charged ions, or some other particles can also take part. Before we deal with specific reactants, we'll examine the interaction of arbitrary particles in its most general form.

Let a reaction between particles of type $A$ and $B$ proceed in some volume where the reactants are distributed uniformly. Such a reaction is called bimolecular. We assume that all the particles are spheres of radius $r$. Each cubic centimeter contains $a_0$ particles of type $A$ and $b_0$ particles of type $B$. We know that particles in a gas move chaotically, which means that their speeds and directions keep changing. However, the root mean square speed of all particles $\langle v \rangle$ is virtually constant. This value depends only on the temperature $T$ of the medium and the particle's mass $m$:[1]

$$\langle v \rangle = \sqrt{\frac{3kT}{m}}.$$

During time $t$, particles of type $A$ travel the mean distance

$$s = \sqrt{\frac{3kT}{m}}t.$$

We assume that the $B$-particles are motionless and that the velocities of the $A$-particles don't change after collisions.

As it moves, an $A$-particle collides with all $B$-particles whose centers are located not more than $2r$ from its trajectory—that is, with those $B$-particles inside the cylinder shown in figure 1. During a time interval $t$ an $A$-particle collides with

[1]This is a rearrangement of an equation you probably encountered in your high school physics textbook:

$$E = \frac{m\langle v \rangle^2}{2} = \frac{3}{2}kT.$$

all $B$-particles that are located in the volume

$$V = 4\pi r^2 s.$$

Of course, an $A$-particle can meet not only $B$-particles but other $A$-particles as well. It can also collide with other objects in the medium. However, these collisions don't interest us, because they don't induce a chemical reaction between the substances $A$ and $B$.

As we mentioned, each cubic centimeter of the medium contains $b_0$ particles of type $B$. Therefore, every $A$-particle will collide with $a_0 V$ $B$-particles. Since the number of $A$-particles in one cubic centimeter is $a_0$, the total number of collisions between $A$- and $B$-particles during time $t$ will be

$$n = V a_0 b_0.$$

Correspondingly, $n/t$ collisions between $A$- and $B$-particles occur per unit time:

$$\frac{n}{t} = 4\pi r^2 \sqrt{\frac{3kT}{m}} a_0 b_0.$$

In this calculation we didn't take into account that $B$-particles also move, the velocities of the particles change after every collision, and the distribution of the particles in the volume is homogenous only on average. However, the precise theory, which takes all these factors into account, corrects our results by no more than 10%.

It also should be noted that not every collision between $A$- and $B$-particles results in a chemical conversion. However, the number of elementary chemical events is proportional to the number of these collisions.

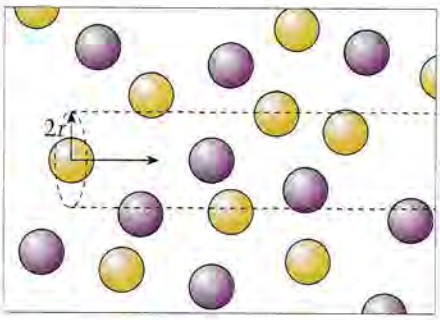



Figure 1

The number of chemical conversions occurring in a unit volume per unit time is called the rate of chemical reaction. The rate of the reaction between $A$- and $B$-particles is determined by the expression

$$W = \alpha 4\pi r^2 \sqrt{\frac{3kT}{m}} a_0 b_0, \qquad (1)$$

where $\alpha$ is a dimensionless proportionality coefficient relating the number of collisions and the number of chemical conversions.

The parameter $\alpha$ essentially signifies the probability of a reaction occuring between two particles if they happen to collide. The value of this probability and the nature of the factors it depends on are considered in the theory of elementary chemical conversion, which is beyond the scope of this article.

Equation (1) describes the law of mass action: at any moment the rate of a chemical reaction is proportional to the product of the concentrations of the reactants at this moment. The proportionality coefficient $K$ that stands before the product of the concentrations is called the rate constant of the chemical reaction at a given temperature.

In a bimolecular reaction between particles of equal mass the rate constant is

$$K = 4\alpha\pi r^2 \sqrt{\frac{3kT}{m}}.$$

Let's calculate the rate constant for a simple chemical reaction.

**Problem 1.** A bimolecular reaction proceeds in an ideal gas under normal conditions (pressure 760 mm Hg, temperature 0°C). The concentrations of both reagents are equal, and there are no other substances in the medium. The mass of molecules $A$ and $B$ is about 30 atomic mass units (1 amu = $1.67 \cdot 10^{-27}$ kg). The radius of the molecules is

$$r = 2.5 \cdot 10^{-8} \text{ cm}.$$

Find the rate of this reaction.

**Solution.** As a first step, let's obtain the mass of the molecules:

$$m = 30 \cdot 1.67 \cdot 10^{-27} \text{ kg} \cong 0.5 \cdot 10^{-25} \text{ kg}.$$

The root mean square speed of the molecular motion is

$$\langle v \rangle = \sqrt{\frac{3kT}{m}}$$

$$= \sqrt{\frac{3 \cdot 1.38 \cdot 10^{-23} \text{ J/K} \cdot 273 \text{ K}}{0.5 \cdot 10^{-25} \text{ kg}}}$$

$$= 480 \text{ m/s}.$$

The constant rate of the reaction is

$$K = 4\alpha\pi r^2 \langle v \rangle$$

$$= 4\alpha\pi (2.5 \cdot 10^{-8} \text{ cm})^2 \cdot (4.8 \cdot 10^4 \text{ cm/s})$$

$$\cong 3.8\alpha \cdot 10^{-10} \text{ cm}^3/\text{s}.$$

We see that the rate constant is expressed in units of $\text{cm}^3$/sec. According to Avogadro's law, one mole of an ideal gas ($6 \cdot 10^{23}$ particles) under the normal conditions occupies 22.4 liters. This means that 1 $\text{cm}^3$ of gas contains $2.7 \cdot 10^{19}$ particles—that is, $1.3 \cdot 10^{19}$ molecules of each type $A$ and $B$ ($a_0 = b_0 = 1.3 \cdot 10^{19} \text{ cm}^{-3}$).

Thus, the rate of the reaction is

$$W = 3.8\alpha (10^{-10} \text{ cm}^3/\text{s}) \cdot$$

$$\cdot (1.3 \cdot 10^{19} \text{ cm}^{-3})^2$$

$$\cong 6.5\alpha \cdot 10^{28} \text{ cm}^3/\text{s}.$$

This means that $6.5 \cdot 10^{28}$ collisions occur per second between $A$ and $B$ molecules in every cubic centimeter of gas.

Let's see what would happen if every collision led to a chemical conversion. In this case, all the molecules would react in a very short time:

$$t' = \frac{a_0}{W} = \frac{1.3 \cdot 10^{19} \text{ cm}^{-3}}{6.5 \cdot 10^{28} \text{ s}^{-1} \text{cm}^{-3}}$$

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

Is this realistic?

A mass of about 30 amu corresponds to molecules of ethane (30 amu) or oxygen (32 amu). Therefore, the reaction described above is a model of gas burning in a conventional gas oven. As you well know, gas doesn't burn all by itself—you need to ignite it with a match or some other device. This means that the value of $\alpha$ in the reaction between ethane and oxygen is very small.

If a chemical process proceeds at a constant temperature, we only need to know the value of the rate constant $K$ to determine how the concentrations of the reactants change during the entire process. Thus the law of mass action can be written as follows:

$$W = K[A][B], \qquad (2)$$

where the symbols $[A]$ and $[B]$ denote the concentrations of the reagents that vary during the reaction.

It's not easy to calculate theoretically the precise value of the rate constant of a chemical reaction, so in most cases it's determined experimentally.

**Problem 2.** Find the rate constant of the chemical reaction between reagents $R$ and $S$ if experiments have shown that 10% of each substance is chemically converted per second. The initial concentrations of each reagent were 1 mole/liter.

**Solution.** In 1 liter of the reacting mixture, 0.1 mole of each reagent is converted per second. Thus $6 \cdot 10^{19}$ elementary events of chemical interaction occur per second in 1 $cm^3$. So the rate of reaction is

$$W = 6 \cdot 10^{19}\ cm^{-3}\ s^{-1}.$$

Plugging the values of the initial concentrations and the reaction rate $W$ into equation (2), we get the value of the rate constant for this reaction:

$$K = \frac{W}{[R]_0[S]_0} = 0.1\frac{\text{liter}}{\text{mol}\cdot\text{s}}$$
$$= 1.7 \cdot 10^{-20}\ cm^3/s.$$

We obtained the value of the rate constant $K$ corresponding to constant concentrations $[R]_0$ and $[S]_0$. If the rate of this process were constant, the reagents $R$ and $S$ would be completely consumed during the first 10 seconds, because 10% of the reagents is processed in one second. However, this doesn't happen in a real chemical experiment—the concentrations of the reagents decrease over the course of the reaction, so the rate at which they're consumed (that is, the reaction rate) decreases as well. Therefore, at different stages of the reaction different amounts of

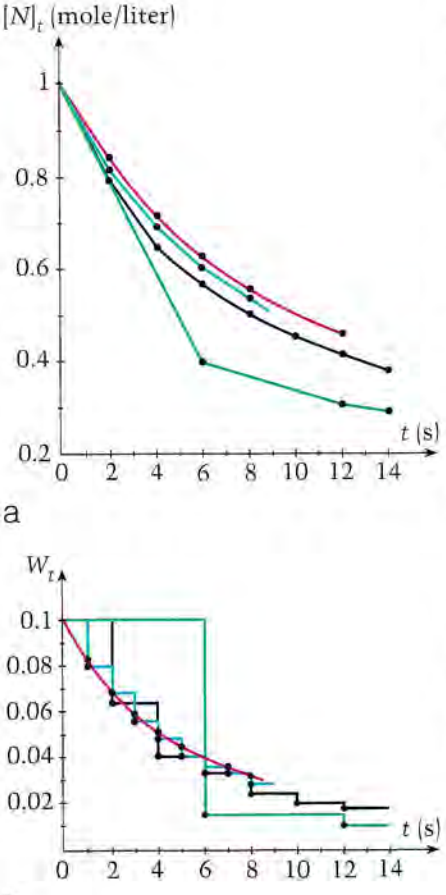



Figure 2

the reagents are consumed during equal time periods.

**Problem 3.** Given the value of the equilibrium constant of a bimolecular reaction between reagents $R$ and $S$ (see problem 2), calculate how the concentrations of the reagents change over time.

**Solution.** To calculate the change in concentration of the reagents over time, let's divide the time into small equal intervals. Since these intervals are small, we assume that the reaction rate doesn't vary during any of them.

First, let's choose time intervals $\tau$ equal to 6 s, 2 s, and 1 s. In figure 2a the green line shows the concentration for $\tau = 6$ s, while the black and blue lines correspond to 2 s and 1 s, respectively. Figure 2b shows the reaction rate over time calculated with these approximations.

In a real reaction, the concentrations of the reagents vary continuously, so our calculations will be more accurate for the smallest values of $\tau$. To obtain precise values for the concentrations at any given moment, we must find the limit as $\tau$ approaches 0, taking into account that the reaction rate varies continuously.

Differential calculus allows us to solve a problem like this precisely. At any given moment, the concentration of a reagent $[N]_t$ is given by the equation

$$[N]_t = \frac{[N]_0}{1 + K[N]_0 t}, \qquad (3)$$

where $[N]_0$ is the initial concentration and $t$ is the time that has elapsed from the beginning of the reaction.[2]

Plugging the values for the equilibrium constant and the initial concentrations into equation (3), we get the formulas describing how the concentrations and the reaction rate change over time:

$$[R]_t = [S]_t = \frac{1}{1+0.1t}\ \text{mol/liter}$$
$$= \frac{6\cdot 10^{20}}{1+0.1t}\ cm^{-3},$$

$$W_t = \frac{0.1}{(1+0.1t)^2}\ \text{mol/(liter}\cdot\text{s)}$$
$$= \frac{6\cdot 10^{19}}{(1+0.1t)^2}\ sm^{-3}\cdot s^{-1}.$$

The red lines in figures 2a and 2b show the plots of these functions. You can see the difference between the curves obtained by this method and those obtained by the approximate method.

Equation (3) makes it possible to calculate the concentrations of the reagents and the reaction rate at any moment and for any initial concentrations of the reagents.

However, chemical processes don't always consist of a single elementary reaction of mutual interaction between the initial substances. Chemical processes are usually far more complex: they incorporate sev-

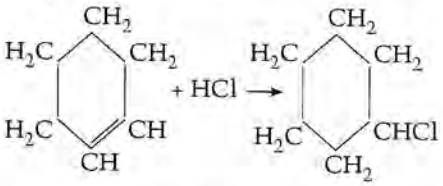
[2]Those who know differential calculus can solve the differential equation:

$$\frac{d[N]}{dt} = K[N]^2.$$

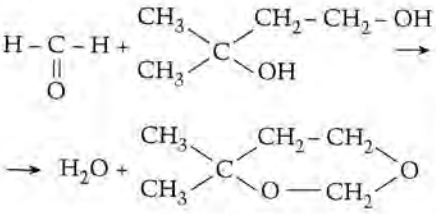
eral elementary reactions occurring simultaneously. For example, the products of a reaction may interact among themselves or react with the initial reagents. Obviously in such a case, equation (3) doesn't describe the kinetics of the process. We need to develop a more complex equation that incorporates the rate constants of all the elementary reactions involved in the process.

But not only do we have to determine the rate constants of all the elementary reactions—we must ensure that the process actually proceeds according to the given set of reactions.

**Problem 4.** In figure 3a the data points represent measurements of the concentration of hydrochloric acid as cyclohexene is chlorinated in the presence of a catalytic agent.




In figure 3b the data points represent experimental data on formaldehyde concentration during the synthesis of dimethyl dioxane.




Which of these reactions can be described by the equation for bimolecular reactions?

**Solution.** If a reaction is bimolecular, the change in the concentration of the reagents over time is described by equation (3).

Let's rewrite equation (3) in another form:

$$\frac{1}{[N]_t} = \frac{1}{[N]_0} + Kt. \qquad (4)$$

This equation describes a straight line in the inverse-concentration–time coordinates. The constant rate of bimolecular reaction equals the slope of this line.

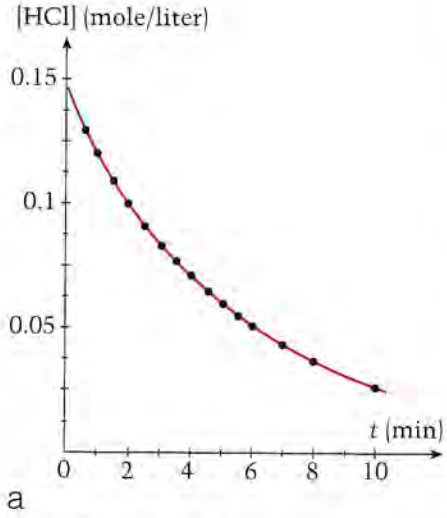



Figure 3

Figure 4

Using the concentration values given in figures 3a and 3b, we plot the dependencies of the inverse concentrations of the initial substances on time (figures 4a and 4b). Indeed, in the case of cyclohexene chlorination we get a straight line. So this reaction is bimolecular. The corresponding equilibrium constant is

$$K \cong 0.02 \text{ liter/mole} \cdot \text{s}.$$

By contrast, the plot illustrating the second reaction is far from being linear. Therefore, this reaction is complex and cannot be characterized by a single rate constant. Special studies revealed that the products of this reaction (dimethyl dioxane and water) interact with each other and yield the original substances. In other words, this process involves two elementary bimolecular reactions occurring simultaneously.

Now we know how the simplest kind of chemical interaction between two substances proceeds. We also know how the reaction rate and the reagent concentrations vary over time. We can calculate the instantaneous values of the concentrations at any moment if we know the rate constant and the initial concentrations of the reagents. Moreover, we can distinguish bimolecular reactions from other processes.

In reality, several elementary reactions proceed simultaneously in most chemical processes, and their products react with one another. Analysis of the entire set of such reactions requires more complex equations to describe the course of a real chemical process. Some of these equations are very difficult even for experienced professionals, who apply the entire arsenal of modern mathematics and computational techniques. It seems that modern chemists must be very skillful in physics and mathematics! ■

# Tackling twisted hoops

## *Untangle these wire pretzels*

by S. Matveyev

LET'S MAKE A CIRCLE OUT OF THIN WIRE, smoothly curve it to give it a more complex shape, and flatten it against a plane (figure 1). What we have is a flat, tangled hoop. *Is it possible to disentangle the wire hoop to obtain the circle again without lifting it from the plane?*

We assume that the wire has zero thickness, so that at the points where one section of the wire passes above another (we'll call them *double points*), the upper section also lies on the plane. The wire is very flexible, but not infinitely flexible, so that the radius of curvature is not zero—otherwise the wire breaks. In particular, the method of straightening loops shown in figure 2 is prohibited.

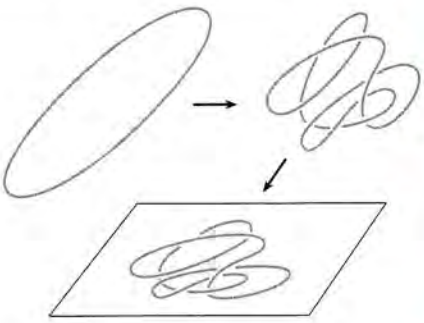


Figure 1

We know that it's possible to disentangle the hoop in space. After all, we obtained the tangled hoop from a wire circle. If we perform the same operations in reverse, we'll get the original circle. Since we started from a circle, we avoided all the difficulties involving knots—if we had started from a knotted hoop, such as the one shown in figure 3, we would be unable to disentangle it into a circle even in space, to say nothing of the plane. (See, for example, O. Viro's article in the May/June 1998 issue of *Quantum*).

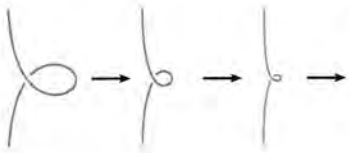


Figure 2

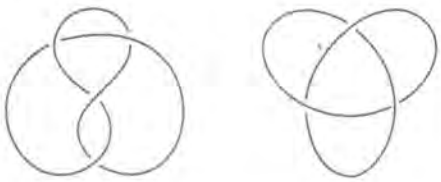
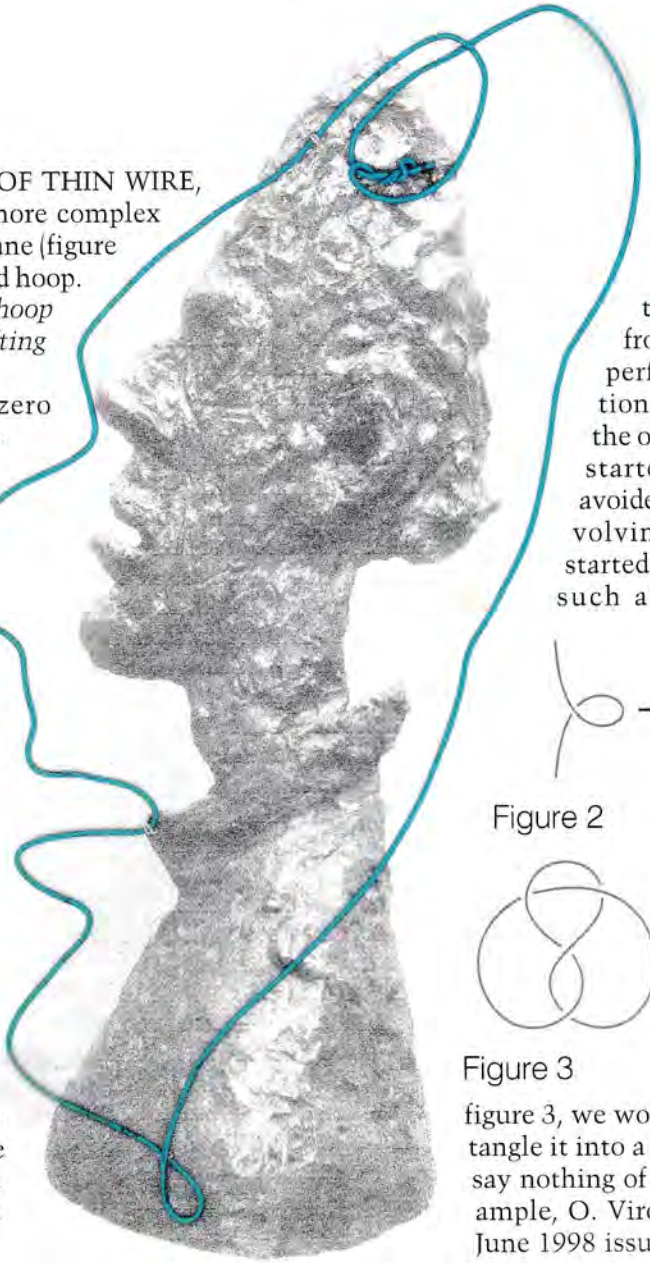


Figure 3



Art by Yuri Vaschenko (Alberto Giacometti, "Large head of Diego" © 2000 ARS, New York/ADAGP, Paris)

## Experiment a bit, think a bit

Let's begin with the wire hoop shown in figure 1. If you play around a bit with a piece of wire (or thread), you'll see that this hoop can be disentangled (see figure 4).

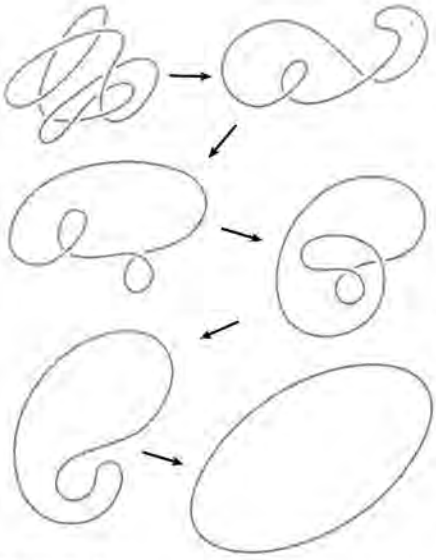


Figure 4

Now I invite the reader to disentangle hoops a–f in figure 5.

I hope you succeeded with hoops b, e, and f. But your inability to disentangle hoops a, c, and d should convince you that hoops exist that cannot be disentangled. How can we prove that a particular hoop cannot be disentangled?

To prove that a certain construction or process is impossible, mathematicians often use the following remarkable method. Every state of the object under consideration is assigned a number that remains the same throughout the process (such a number is called an *invariant*). Then the invariant is determined for the initial and the desired states of the object. If different values are obtained, it means that it is not possible to pass from the initial to the desired state—after all, that's why it's called an invariant: it cannot change during the process!

So let's try to assign a number to every hoop on the plane. The first idea that comes to mind is to count the double points in the hoop. Alas! this is not an invariant, as you can see from figure 4. However, examining this figure, we notice that double points appear and disappear in pairs. This leads to the idea that the parity of the number of double points is an invariant (in other words, the remainder upon division of the number of double points by two is an invariant).

This is indeed the case, as we will see later. This fact implies, for example, that the hoop in figure 5a cannot be disentangled (it has seven double points, while the circle has no double points—this prevents us from transforming the hoop into the circle). The hoop in figure 5d has four double points; thus its invariant is zero, the same as for the circle. Does this mean that it can be disentangled? No, it does not, because we don't know whether the condition of zero invariants is sufficient. Thus the question of whether the hoop in figure 5d can be disentangled remains open.

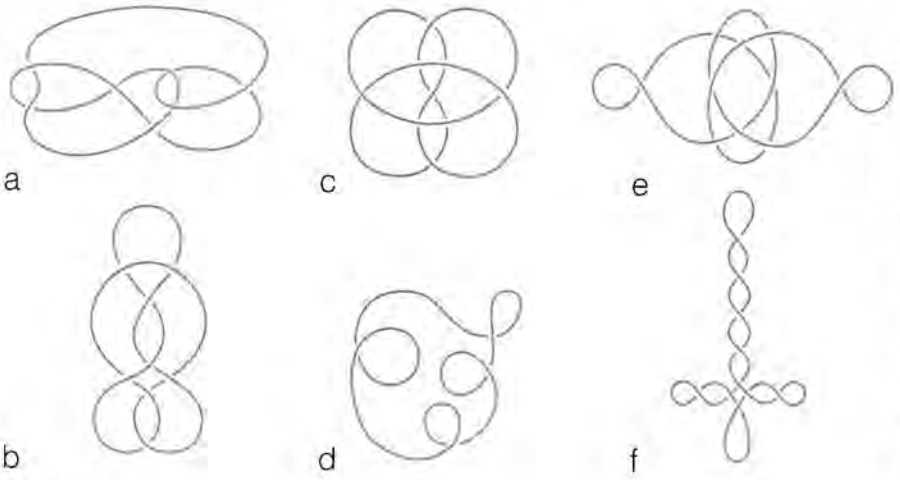
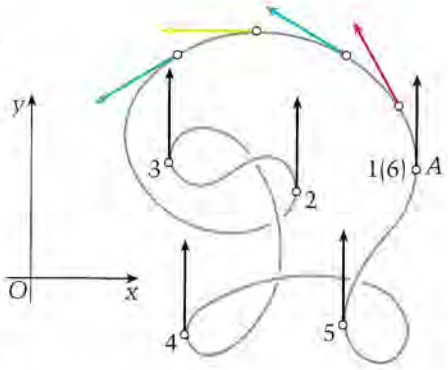



Figure 5



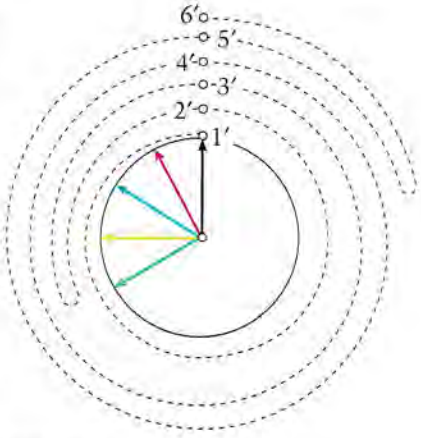




Figure 6

The reasoning above must convince you that it makes sense to search for invariants in this case. Let's do just that.

## The invariant *I*

Let a tangled hoop be given in the plane. Take an arbitrary point $A$ on this hoop and choose one of two possible directions of going around the hoop. We'll move a point along the hoop, starting at point $A$, with unit speed in this direction. The velocity vector will turn about $A'$ and its endpoint will move along a circle centered at $A'$. When we complete the tour around the hoop and return to point $A$, the velocity vector returns to its initial state; therefore, the total number of revolutions of this vector about $A'$ is an integer. We assign revolutions made in the positive direction (counterclockwise) a plus sign and revolutions made in the negative direction (clockwise) a minus sign.

Look at figure 6. In this figure, the endpoint of the velocity vector is shown as a dashed curve and is shifted from the circle to make it

easier to see what's going on. In reality, the red curve is tightly wound on the circle, and points 1′–6′ coincide. All in all, the velocity vector performs –1 revolution: from point 1′ to point 2′, one revolution; from point 2′ to point 3′ and from 5′ to 6′, no revolutions; from point 3′ to point 4′, one revolution in the negative direction; and from 4′ to 5′, one revolution in the negative direction as well.

The invariant we promised (we'll denote it by $V$) equals the absolute value of the total number of revolutions of the velocity vector. It's clearly independent of the choice of the starting point $A$, nor does it depend on the initial direction taken; indeed, changing the direction merely changes the sign of the total number of revolutions. For example, for the hoop in figure 6, the invariant is 1.

We'll show (without giving a rigorous proof) that $V$ is actually an invariant. When the hoop is disentangled, the position of the velocity vector changes smoothly, without making any jumps. Therefore, the number $V$ must also change smoothly. However, $V$ is an integer and it can turn into another integer only by making a jump, which contradicts the criterion of continuity. Therefore, $V$ remains unchanged and is indeed an invariant of the disentangling operation.

Now we can tackle the hoop in figure 5d. For this hoop, $V = 3$ (check this on your own!); therefore, it cannot be disentangled into a circle, for which $V = 1$.

If you actually verified that $V = 3$ for the hoop in figure 5d, you must have noticed that, in practice, it's not so easy to calculate the number of revolutions of the velocity vector. In fact, it's easy to miscalculate. However, there's an easier way to calculate $V$.

For this purpose we choose a direction in the plane—for example, the direction of the axis $Oy$ (see figure 6)—and mark the points of the hoop where the velocity is parallel to $Oy$ and in the same direction. We write the number +1 near a marked point if the small section of the hoop containing this point lies to the left of it; we write the number –1 near a marked point if the section of the hoop containing this point lies to the right of it. (If the section containing the marked point lies on *both* sides of it, we don't write any number. This happens when the vector is traveling along a loop and suddenly starts looping in the other direction at the marked point—it traces a sort of flattened "*S*." and never completes the first loop.) Now we can say *the invariant $V$ equals the absolute value of the sum of all the numbers written.*

For example, in figure 6, we write +1 at points 1 and 2 and –1 at points 3, 4, and 5. Thus $V = 1$ for this hoop. We invite the reader to prove that this method actually gives the value of $V$ for any hoop.

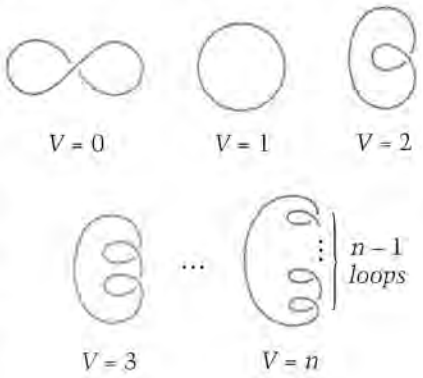



Figure 7

In figure 7, for any nonnegative integer $n$, a hoop whose invariant $V$ equals $n$ is shown. We recall that if a hoop can be disentangled into a circle, its invariant $V$ must be equal to the invariant of the circle—that is, to 1.

## The invariant *R*

The equality $V = 1$ is a necessary condition for a hoop to be disentangled into a circle. But is this condition also sufficient? At first I thought it was, but unsuccessful attempts to disentangle my belt, arranged as shown in figure 8b, convinced me that it wasn't and simultaneously elicited an important observation: when I picked the belt up off the floor (figure 8a), it was twisted completely around twice!

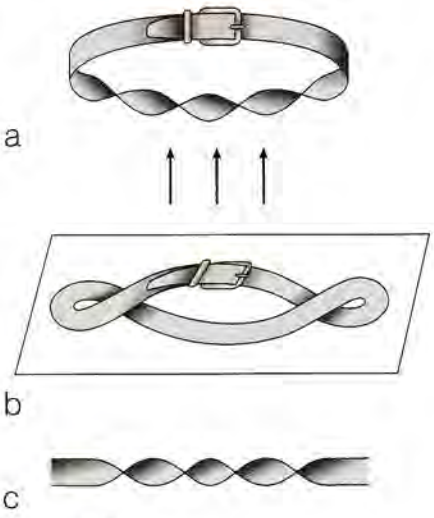



Figure 8

Let's replace the hoop with a band that lies on the plane such that its middle line coincides with the hoop (figure 9a). Disentangling the hoop in space (for example, returning it to the initial state in which it was before placing it on the plane), we obtain a twisted band. We denote the number of complete twists (where the "front" is twisted around and faces the front again) by $R$. This number is our *second invariant*, and

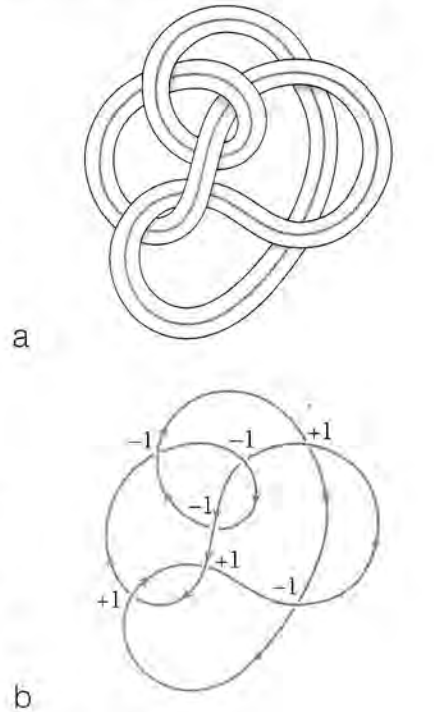



Figure 9

if the hoop is to be disentangled into a circle, this invariant must be zero. To be more specific, the number of complete twists is given a plus sign if the band is twisted as in figure 8a and a minus sign if it's twisted as in figure 8c (recall the difference between a left- and right-threaded screw).

We'll prove that *the number R is indeed an invariant*—that is, it doesn't change when the hoop is disentangled in the plane. It's sufficient to notice that disentangling the hoop determines a method for disentangling the corresponding band. But the number of twists of the band remains unchanged not only when the band is disentangled in the plane, but even for any three-dimensional motion.

It can be proved (we won't do it here) that the invariant $R$ can be calculated as follows. Choose a direction for going around the hoop. Then mark every double point with the number +1 if the lower velocity vector is directed to the left of the upper velocity vector; otherwise, mark this double point with –1. It's easy to see that these numbers are independent of the direction chosen. The invariant $R$ equals the sum of these numbers. For example, the hoop in figure 9b has three positive and four negative double points; thus, its invariant $R$ is –1. Therefore, this hoop cannot be disentangled in the plane.

## Necessary and sufficient conditions

We've already seen that the conditions $V = 1$ and $R = 0$ are necessary for the hoop to be disentangled into a circle. But are these conditions sufficient? In other words, is it sufficient to check that $V = 1$ and $R = 0$ to be sure that the hoop can be disentangled into a circle? The answer is yes.

**Fundamental theorem.** *In order for a hoop to be disentangled into a circle in the plane, it is necessary and sufficient that its invariant V be equal to 1 and its invariant R be zero.*

This theorem gives the complete answer to the question formulated at the beginning of this article. The simple methods described above for evaluating $V$ and $R$ allow us to quickly check the necessary and sufficient conditions in the theorem. We invite the reader to apply this theorem to the hoops depicted in figure 10.

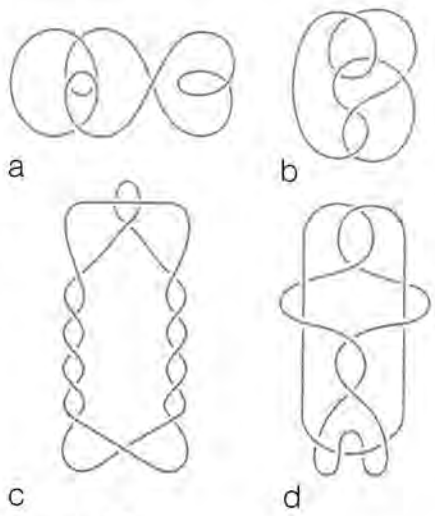


Figure 10

## Proof of the fundamental theorem

We've already proved that $V$ and $R$ are invariants; thus the necessity of the conditions $V = 1$ and $R = 0$ is already proved. To prove sufficiency, we must show that every hoop for which $V = 1$ and $R = 0$ can be disentangled into a circle in the plane.

Consider a hoop of this type. We know that it can be disentangled in three-dimensional space. We denote by $\tilde{K}_t$ the position of the hoop at the time $t$ in the process of disentangling it. The moment $t$ will be called *singular* if the hoop $\tilde{K}_t$ has a vertical tangent at one or more of its points. Assume that there are no singular moments. Then the hoop can be disentangled in the plane. Indeed, assume that the ceiling of the room where we work with the hoop is parallel to the plane to which the hoop belongs. Imagine that the ceiling starts dropping until it reaches the plane with the hoop. In the process, every hoop $\tilde{K}_t$ goes to a certain plane hoop $K_t$. The absence of vertical tangents guarantees that no folds (points with zero radius of curvature) occur in the hoops $K_t$. The family of hoops $K_t$ determines the desired method for disentangling the given hoop into a circle in the plane.

Now consider how the hoop $K_t$ behaves when the moment $t = t_0$ is singular—that is, the hoop passes the state $K_{t_0}$ with a vertical tangent. A typical picture of this passage through the vertical state is shown in figure 11. We see that when the hoop undergoes the transformation $\tilde{K}_{t_1} \to \tilde{K}_{t_0} \to \tilde{K}_{t_2}$ in space, the corresponding plane hoop undergoes the forbidden transformation $K_{t_1} \to K_{t_0} \to K_{t_2}$ during which the break $K_{t_0}$ occurs and a loop appears on the hoop $K_{t_2}$.

It can be proved (but not here) that the process of disentangling any hoop in space can be performed in such a way that only a finite number of singular moments occurs and all of them are typical—that is, *a single loop appears or disappears at each of these moments.*

Assume that a loop disappeared at the moment $t_0$. We cannot destroy it in the plane; thus we contract it into a very small loop and won't change it in future transformations (we consider it "frozen" or glued in place).

Now assume that a loop has *appeared* at a singular moment. We cannot create a loop by transforming the hoop in the plane, but we can create two (mutually annihilating) loops, as shown in figure 12. Thus

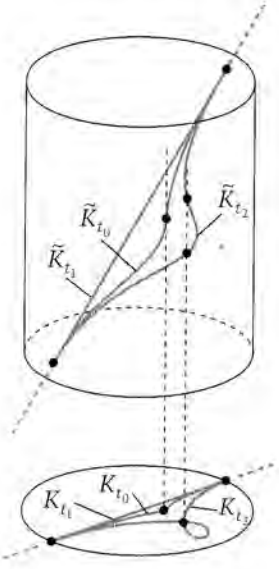


Figure 11

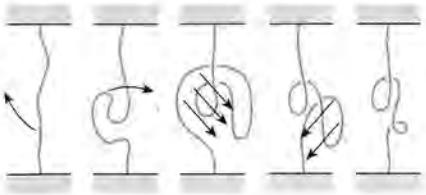


Figure 12

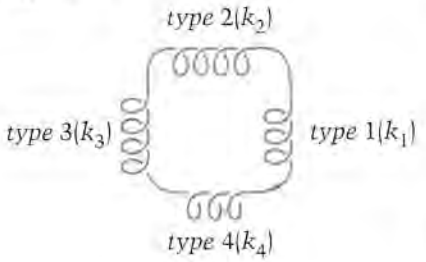



Figure 13

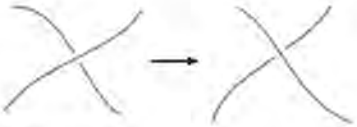


Figure 14

we create two loops, contract the extra one into a very small loop, and "freeze" it.

Continuing the process of disentangling simultaneously in three-dimensional space and in the projection onto the plane, we transform the plane hoop into a circle with a finite number of small ("frozen") loops. These loops can be classified into four types depending on where the loop is situated (inside the circle or outside of it) and in what order its double point passes (first the upper and then the lower thread, or vice versa). Then we can change the order of the loops by pulling one through the other, as shown in figure 13.

If $k_i$ denotes the number of loops of type $i$, then $V = 1 + k_1 + k_2 - k_3 - k_4$ and $R = k_1 - k_2 + k_3 - k_4$. Recalling that $V = 1$ and $R = 0$, we obtain the system of equations

$$\begin{cases} k_1 + k_2 - k_3 - k_4 = 0, \\ k_1 - k_2 + k_3 - k_4 = 0, \end{cases}$$

from which it follows that $k_1 = k_4$ and $k_2 = k_3$. A pair of loops of types 1 and 4 can easily be destroyed, as shown in figure 12; the same is true for pair of loops of types 2 and 3 (figure 4). It remains to transform our circle with loops into a real circle. The theorem is thus proved.

## Disentangling hoops with self-intersections

Now let's change the statement of the problem by saying we're allowed to create self-intersections while we're disentangling the hoop. More precisely, we're allowed to pull the lower part of the loop through the upper part near double points, as shown in figure 14. This problem statement doesn't seem quite natural (indeed, to perform such a transformation, we must cut the hoop and glue it back together, which can wear down even the most patient experimenter). And yet a formal mathematical problem investigated by the American mathematician H. Whitney in the 1930s can be reduced to this very statement. In fact, Whitney's problem served as the starting point for this article.

Since we are now interested in *disentangling hoops in the plane with self-intersections allowed*, the reader is invited to prove the following statements.

1. The number $k_i$ is an invariant of the operation of disentangling with self-intersections. (Hint: recall the method for evaluating $V$ described above.)

2. The number $R$ is not an invariant of the operation of disentangling with self-intersections. (Hint: experiment with a belt, exchanging the upper and lower parts near one of the double points.)

3. The remainder $R'$ upon division of $R$ by 2 is an invariant of the operation of disentangling with self-intersections. (Hint: every operation of self-intersection replaces the number $\pm 1$ marking the double point with $\mp 1$.)

4. The number $R'$ is an invariant of the operation of disentangling (without self-intersections!) the hoop in the plane.

To make further progress, we'll need the notion of a *simple loop*:

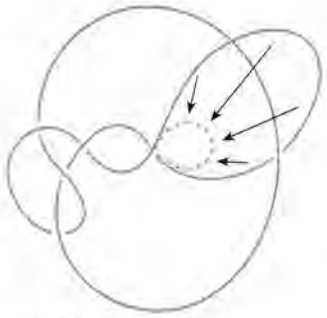


Figure 15

this is a portion of the hoop that begins at a double point, ends at the same double point, and has no self-intersections (though it may intersect other portions of the hoop, as shown in figure 15). Now try to prove the following series of propositions.

5. Every plane hoop has a simple loop.

6. Every simple loop can be contracted (with self-intersections!) into a small loop without affecting other parts of the hoop.

7. Any hoop can be transformed (with self-intersections) into a figure eight, a circle, or a circle with a finite number of small loops inside it.

8. Any hoop can be transformed (with self-intersections) into any other hoop if we first add to one of these hoops several (how many?) loops.

9. (Whitney's theorem) A hoop with invariant $V_1$ can be transformed into another hoop with invariant $V_2$ if and only if $V_1 = V_2$.

In conclusion, we present three more problems related to the initial problem statement (concerning the process of disentangling without self-intersections).

10. For any pair of integers $m$ and $n$ with an odd sum ($m \geq 0$), construct a hoop with invariants $V = m$ and $R = n$. Why don't any hoops exist with invariants $V = 1$ and $R = 1$?

11. Formulate and prove an analogue of Whitney's theorem for disentangling hoops without self-intersections.

12. Prove that any hoop on the sphere can be transformed (without self-intersections) into either a circle or a figure eight. ■

# Challenges

## Physics

### P306

*Model behavior.* A working 1:10 scale model of a helicopter is powered by 30-watt engine. What would be the minimum power needed to lift a real helicopter built from the same materials?

### P307

*Flow factors.* Gas leaks from a balloon through a small hole. How much will this flow change if the temperature of the gas is increased by a factor of four and the pressure by a factor of eight?

### P308

*Spark generator.* The discharge gap of a spark generator (figure 1) is set to a voltage $V$, and a resistor $R$ is chosen to elicit $n$ discharges per second. Find the mean power dissipated by the resistor if the capacitor is completely discharged during a spark. (P. Zubkov)

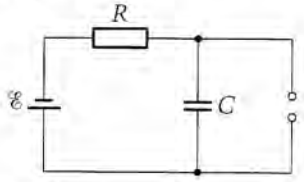



Figure 1

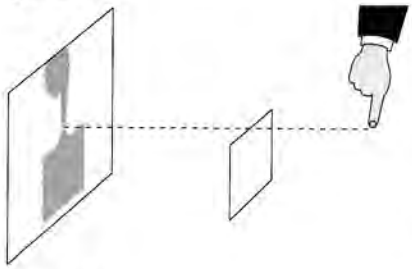


Figure 2

### P309

*Generating heat.* An inductance coil has many windings of a wire with a high resistivity. The ends of the coil are connected. A strong permanent magnet is placed near the coil. The magnet is quickly removed so that it generates an electric current in the circuit. In the first 100 ms, 0.01 J of heat is released; in the next 100 ms, an additional 0.006 J of heat is released. How much heat will be released in the circuit over a long period?

### P310

*Shadow on the wall.* On a bright sunny day, when the Sun is high over the horizon, look at the shadow cast by the clean edge of a piece of cardboard on a smooth screen (say, a white wall). Now place your finger near the cardboard as shown in figure 2. When you bring your finger closer to the cardboard, a second shadow will emerge from the dark region of the screen in addition to that produced by your finger on the bright part of the screen. Explain this result.

(G. Solovyanyuk)

## Math

### M306

*Sit and ponder.* Infinitely many seats are lined up along a racecourse and numbered 1, 2, 3, 4, ... . An incompetent cashier sold tickets for the first $m$ places, but sold more than one ticket for some seats, and no tickets for others. Altogether, she sold $n$ tickets, where $n > m$.

The spectators enter one by one. Each attempts to sit in the seat for which he or she holds a ticket. If no one is in the seat, they occupy it. If that seat is already occupied, the spectator says "Oh!" and moves to the seat with the next higher number. If this seat is unoccupied, the spectator takes it. Otherwise, the person again says "Oh!" and moves to the seat with the next higher number. This continues until the spectator finds an unoccupied seat.

Prove that the number of "Oh!"s uttered is independent of the order in which the spectators enter.

(A. Shen and N. Vasilyev)

### M307

*Prove your point.* Two intersecting circles are given in the plane. $A$ is one of the points where the circles intersect. In each circle a diameter is drawn that is parallel to the tangent to the other circle at point $A$, and these diameters do not intersect. Prove that the four endpoints of these diameters lie on the same circle. (S. Berlov)

### M308

*Reasonable roots?* It is known that $f(x)$, $g(x)$, and $h(x)$ are quadratic trinomials. Can the equation $f(g(h(x))) = 0$ have the roots 1, 2, 3, 4, 5, 6, 7, and 8? (S. Tokarev)

### M309

*Draw a line.* Three points $A$, $B$, and $C$ are given in the plane. Draw a line through point $C$ such that the product of the distances from points $A$ and $B$ to this line is greatest. Does such a line always exist?

(N. Vasilyev)

CONTINUED ON PAGE 24

# Laser pointer

## *Shedding light on this remarkable little device*

by S. Obukhov

IT'S ABOUT THE SIZE OF A ballpoint pen, or even smaller (it's sometimes sold as a key chain gadget), and it's the source of inexplicable delight. I know some very respectable people who bought one the minute they saw it and cannot stop playing with this marvelous little thing—a laser pointer.

Why is it so amazing? Let's back up and look at its venerable ancestor, the flashlight.

Everybody knows that a flashlight can illuminate objects at a distance of 5–20 meters. The resulting luminosity depends mainly on how tightly the flashlight is focused. In an ideally focused flashlight, the radiant tungsten filament in the incandescent bulb must be located at the focus of the parabolic reflector. To adjust the flashlight, we move the reflector in both directions or even remove it and slightly rotate the bulb in the holder before aligning the reflector again, trying to find the best mutual disposition of the bulb and the reflector.

The size of the filament is a few millimeters. Therefore, if one part of the filament is placed at the focus, other parts of the filament will lie outside the focus. This is why a flashlight beam always diverges. The angle of divergence of a flashlight beam (in radians) is approximately equal to the ratio of the filament size to the diameter of the reflector (several centimeters): $\alpha \cong 4$ mm/4 cm = 0.1 rad, or about 5 degrees.

In everyday life, we rarely need to measure angles with the naked eye, except maybe such angles as 90°, 45°, and the like. How can we visualize an angle of 5°? Amateur astronomers know how—they use devices that are always "at hand," so to speak: their own arms. Stretch out your arm in front of you and spread your forefinger and middle finger to form the "V for victory" sign. Now close one eye and look at your fingers with your other eye. The angle formed by your two fingers is about 5°, or 1/10 radian.

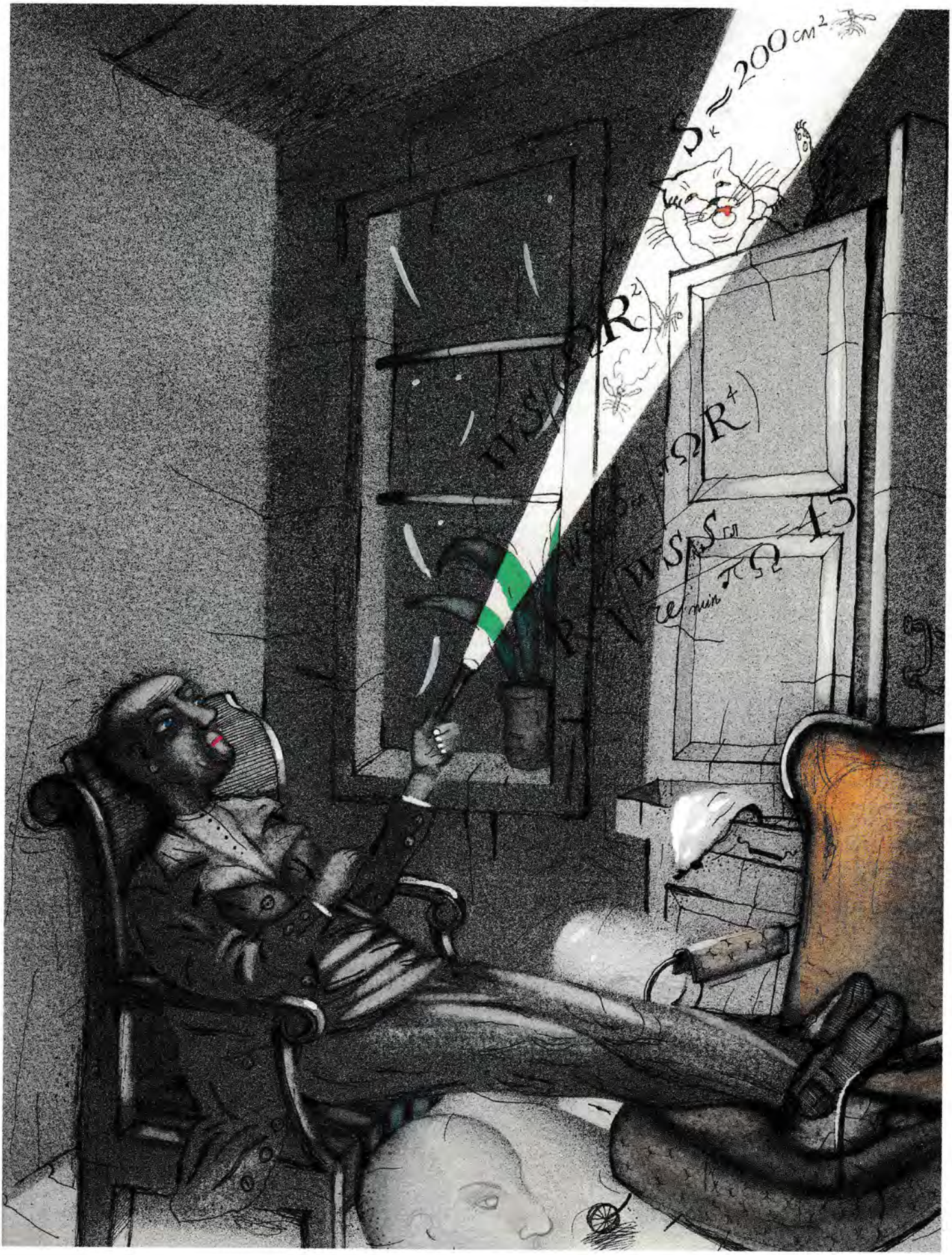
I often use this method to measure the angle between the Sun and the horizon in order to determine the time left before sunset. By the way, the angular size of the Sun is about 0.5°, or 30′, which is the same as that of the Moon. The angle formed by the fingernail of your index finger at the end of your outstretched arm is approximately three times greater (about 1.5°).

Some flashlights have large reflectors (10–12 cm). The divergence angle of the beam for such flashlights is correspondingly smaller (by a factor of 3). Also, very small high-power bulbs are available that focus the beam even better.

At what distance can one see the beam of a flashlight? The power consumption of a conventional flashlight is about 1 W. Only 1/20 of this energy is radiated in the form of visible light; most of the energy is converted to heat and thermal radiation. A flashlight radiates power $P$ within a solid angle $\Pi = \pi\alpha^2$. At a distance $R$ the radiated power incident on a unit surface area is $P/(\Omega R^2)$. In turn, the power $P$ of the light entering the eye from a distance $R$ is $PS_{eye}/(\Omega R^2)$, where $S_{eye}$ is the area of the pupil. In darkness, the pupil of the human eye has a diameter of about 7 mm, so $S_{eye} \cong 0.5$ cm$^2$. If the power of the incident light is greater than a certain threshold value $p_{min}$, we can see the light. The minimal threshold power $p_{min}$ of the human eye can be as low as $10^{-18}$ W, which corresponds to several photons striking the retina per second.

This remarkable sensitivity is possible only after the eye has had time to adjust to the darkness. If we could conduct an optical experiment in complete darkness without an

Art by Ekaterina Silina


≈200 cm²

atmosphere, we might perceive the light of an flashlight from a distance of ten thousand kilometers. In reality, the threshold of the human eye is many orders of magnitude higher, primarily because of the presence of other bright objects in the field of vision—streetlights, houses, stars, the Moon, and so on. In this article, when we analyze optical experiments performed in the open air on a moonless night, we'll use the value $p_{min} = 10^{-13}$ W. Plugging numerical values into the equation

$$R = \sqrt{\frac{PS_{eye}}{p_{min}\Omega}} \qquad (1)$$

yields $R = 27$ km.

At what maximum distance can we illuminate an object with a flashlight? Alas, this distance is only few dozen meters. When we illuminate an object (say, a cat), we want to see the light reflected from the object. The total power of the light incident on an object situated at a distance $R$ is given by the equation $PS/(\Omega R^2)$, where $S$ is the area of the object's surface (for a cat, $S_c \cong 200$ cm$^2$). We'll assume the cat is white (and not gray or black), which means that most of the incident light is reflected (as diffuse light), not absorbed. If the reflected light is scattered in all directions, only a negligible portion of it will enter the observer's eye. This fraction is equal to the ratio of pupil's area to 1/4 the area of a sphere of radius $R$—that is, $S_{eye}/(\pi R^2)$. The factor 1/4 corresponds to the case where the reflecting surface is perpendicular both to the direction of the beam and to the observer. We also take into account that the total light reflected from the surface is proportional to the solid angle at which this surface is viewed by the observer. The power of the light entering the eye is $PS_cS_{eye}/(\pi\Omega R^4)$. Equating this to $p_{min}$, we get

$$R = \sqrt[4]{\frac{PS_cS_{eye}}{p_{min}\pi\Omega}} = 45 \text{ m}. \qquad (2)$$

Therefore, a cat can see the beam of a flashlight from dozens of kilometers away, but we can illuminate the cat from only a distance of 45 m. (A similar situation arises with speed traps, where a police officer uses a radar gun to catch speeders. If the car is supplied with a radar detector tuned to the proper frequency, the driver will know about the trap long before the car becomes "visible" to the radar gun.) Equation (2) shows that if you want to double the detection radius for watching a cat, you must increase the power of your flashlight by a factor of $2^4 = 16$. Correspondingly, the cat will detect your flashlight from four times farther away.

## Laser pointer

An inexpensive laser pointer can project a spot of light on objects situated in the dark hundreds of meters away. The technical description on the package says that the pointer's range may be 200, 500, 800, or even 1,200 m. Keep in mind that the power consumption of a laser pointer is negligible. In the United States one can purchase 5-mW laser pointers, while in Europe the power is restricted to 1 mW. Usually, the specified parameter is the power consumption, while the radiation power is about 60% of this value.

The ability of such a low-power device to put a bright red dot on almost every building on a poorly lit street is truly amazing. Clearly this capacity is related to the extraordinarily small divergence of the laser beam. Theoretically, the divergence angle $\alpha$ is determined only by the diameter of the emerging beam $D$ and the wavelength $\lambda$:

$$\alpha = \lambda/D. \qquad (3)$$

For a laser pointer that throws a red beam (with a wavelength of 600–700 nm), one can use the equation

$$\alpha \text{ (millirad)} = 1/D \text{ (millimeter)}. \qquad (4)$$

It's worth noting that the same equation determines the angular resolution of the human eye, but in that case $D$ is the diameter of the pupil. Since both the diameter of the laser beam and the size of the pupil are almost the same and equal to several millimeters, the beam's divergence angle is approximately equal to the resolution angle of the eye. In actuality the laser beam's divergence is somewhat wider: about 1 cm for every 10 m of its path. Therefore, at a distance of 1 km the beam's diameter will be 1 m. Nevertheless, to the human eye the spot of light looks like a point at any distance, with an angular size equal to that of Jupiter in the night sky.

Now we can explain why equation (2) can't be used for a laser beam. When we deduced it, we assumed that the diameter of the flashlight beam projected at the distant object is far greater than the size of the object. Thus the object is illuminated by only a small fraction of the radiated light. If, however, we use a laser pointer to illuminate a distant object whose angular size is greater than the resolution of the human eye, all the light from the laser beam will strike the object and be scattered. Therefore, the factor $S_c/(\Omega R^2)$ should be replaced by 1; so we get

$$p = \frac{PS_{eye}}{\pi R^2}. \qquad (5)$$

This equation can be illustrated as follows. Imagine a tiny bulb of power $P = 0.003$ W "attached" to the far end of the laser beam. Whatever object we direct the laser pointer's beam at, the bulb "attached" to the end of the beam shows up on the same object and shines in our direction. If we see the light from this bulb, we'll consider this light the reflected beam from our laser pointer. The equation for the largest distance at which the reflected beam will be visible is the same as the equation for the maximum distance at which one can see the bulb "attached" to the end of the beam:

$$R = \sqrt{\frac{PS_{eye}}{\pi p_{min}}}. \qquad (6)$$

Plugging numerical values into this equation, we get $R = 700$ m.

Note that this equation differs from equation (2): to double the

range, it's sufficient to increase the power by a factor of 4 (rather than 16).

## Beam luminosity

As you know, the power of commercially available lasers is strictly limited. So why do some seem brighter than others? Here another factor comes into play: wavelength. The sensitivity of our eyes is strongly dependent on the wavelength, which in commercial laser pointers may be 633, 650, 670, or 680 nm. The beam of a laser pointer with a wavelength of 650 nm seems to be five to ten times brighter than a beam with a wavelength of 670–680 nm, while a 633-nm beam looks twice as bright as a 650-nm beam. The sensitivity of the human eye is greatest for green light with a wavelength of 555 nm, so lasers operating at this wavelength would seem brightest. Indeed, the brightest laser pointer, which has recently come on the market, radiates a green beam with a wavelength of 532 nm. Its luminosity is about eight times that of a 650-nm laser pointer.

The dependence of apparent brightness on wavelength should be taken into account in estimating the distance at which the laser beam is visible. To do so, we assume that the sensitivity $p_{min}$ changes with the wavelength of the radiated light. Advertisements claiming that a certain laser pointer generates a beam visible at a distance of $X$ hundred meters ($X$ = 2, 5, 8, 12, etc.) should be taken with a grain of salt, to say the least, since $p_{min}$ strongly depends on the amount of background light from stars, the Moon, streetlights, and so on.

In addition, the brightness of the laser spot strongly depends on the reflective properties of the illuminated surface. We've assumed that the illuminated surface scatters the reflected light in all directions. What about the special reflective materials used for highway signs, lane markers, and safety vests for road workers? A surface coated with such a material reflects light in the direction almost entirely opposite that of the incident beam (the angle between the incident and the reflected beams doesn't exceed 3°). In this case, the brightness of the reflected light is greater than the brightness of the light reflected from an ordinary light-scattering surface by a factor of about $\pi/(\pi\alpha^2) \cong 400$. If a 3-mW beam is reflected from such a special surface, an observer will perceive the same amount of light as if a well-focused flashlight with angular divergence of 3° were attached to the end of the laser beam and directed back to the observer.

Unfortunately, at large distances the diameter of the beam may exceed the size of the surface coated with the reflective material. If the diameter of the beam is 1 m at a distance of 1 km and the size of the reflective target (say, a street sign) is 0.5 m, only a quarter of the beam's power is reflected back to the observer. To estimate the distance at which the reflected light will be seen, we can use a equation that looks like equation (2) deduced for a similar estimate with a flashlight:

$$R = \sqrt[4]{\frac{PS_cS_{eye}}{p_{min}\pi\Omega\Omega'}}. \qquad (7)$$

In this equation we need to use $\alpha = d/\lambda = 0.0005$ for the solid angle $\Pi = \pi\alpha^2$, introduce an additional factor $\Omega' = 1/400$ in the denominator describing the focus of the reflected beam, and insert the size of the street sign—approximately 0.2 m$^2$—in place of $S_c$. As a result, we get $R \sim 3.5$ km.

## Laser pointers and night vision

Interesting results can be obtained by observing reflected light at night with an infrared viewer. To do this, attach a laser pointer with a rubber band to the body of the viewer such that their optical axes are approximately parallel.

An infrared viewer is a combination of binoculars and a cathode-ray tube, which amplifies the intensity of the incident light.

By themselves, binoculars significantly increase the viewing range in low light, because all the light collected by the objective lenses of the binoculars is transferred to the observer's eyes. For example, if we use binoculars whose objectives have a diameter of 50 mm, the collecting area will be increased by a factor of $(50/7)^2 \cong 50$ (the pupil of our eye has a diameter of 7 mm).

Binoculars with large objectives are most suitable for night viewing, but the magnification should be small. For example, 7x binoculars (50 × 7) are better for this purpose than 12x binoculars (50 × 12). The amount of light collected by both binoculars is equal, but the image will shake more in 12x binoculars than in 7x binoculars.

Note that the first figure in the description of the type of binocular gives the entrance aperture, while the second figure gives the magnification. The value of the exit aperture can be determined as the ratio of the entrance diameter of the binoculars and the magnification. For example, 7x (50 × 7) binoculars have an exit aperture of 50 mm/7 = 7.1 mm. This value is approximately equal to the diameter of the pupil of the human eye adapted to pitch darkness. In 7x (35 × 7) binoculars the diameter of the exit aperture is 35 mm /7 = 5 mm, so these binoculars yield less light. In broad daylight the difference between these binoculars is negligible, because under these conditions the diameter of the pupil is only 2–3 mm, which is smaller than the exit apertures of both binoculars.

If at night we can see an object with the naked eye from a maximum distance $Z$, then with 50 × 7 binoculars we could see objects at a distance of $7Z$. Similarly, if all the stars in the Universe were equally bright and all the visible stars were $Z$ light years away, we could see stars at a distance of $7Z$ through the 50 × 7 binoculars. Thus the volume of the Universe open to our view would be increased by a factor of $7^3 = 343$. The number of visible stars will increase by the same factor! In the real Universe, stars are not distributed homogeneously and they

are not equally bright, but our estimate of the increase in the number of visible stars achieved by using the binoculars remains valid.

The cathode-ray tube of the infrared viewer magnifies light by a factor of several thousand (or even tens of thousands). The maximum sensitivity of this tube occurs in the red and near-infrared part of the spectrum, which is quite suitable for observing the red light of a laser pointer. All laser pointers (from the cheapest to the most expensive) that generate red light of various shades are virtually identical in brightness. An observer equipped with an infrared viewer can detect light emitted by a point source if only a few photons enter the objective of the viewer per second. Note that in this case complete adaptation to darkness isn't necessary. The actual sensitivity threshold could be lowered further still (by a factor of 10 or more), because the collecting surface of the viewer is much larger than the pupil of our eye. Using such a viewer, I was able to see the light from a laser pointer reflected from low-hanging clouds.

Can we aim the beam of a laser pointer directly at a very distant object—say, an orbiting satellite or a distant ship on the open sea at night? These objects are so far away that the light reflected by them cannot reach an observer. Nevertheless, the problem can be solved with the help of an infrared viewer. The sensitivity of this wonderful device is so high, it can trace the trajectory of a laser beam traveling through the air. Maybe it's not so amazing after all. No doubt you've seen the beams from searchlights in the night sky. Sometimes we can see the beam of a flashlight glimmering in the mist. What it means is that some of the beam's energy is dissipated by fluctuations of the air density or by microscopic particles floating in the air.

The length of the segment of a beam's trajectory where it can be observed from the side is necessarily limited—we see the beam only where it is sufficiently concentrated.

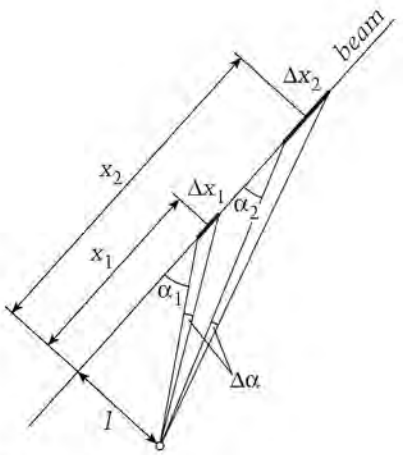



Figure 1

At large distances, the intensity of the beam decreases due to its divergence, so that the beam seems to disappear. The angular divergence of a laser beam is approximately equal to the angular resolution of the human eye, so at large distances a laser beam is seen as a very thin thread of light. And this thread is visible at any distance from the observer! Even if we don't see the light reflected from the distant ship, we'll see the beam from our laser pointer burrowing through the air and striking the ship (or even the satellite).

To understand why the brightness of the laser beam's trajectory doesn't depend on distance, let's examine figure 1. It shows two segments of the beam's path $\Delta x_1$ and $\Delta x_2$, whose lengths are chosen in such a way that they have the same angular size $\Delta\alpha$ as viewed by an observer. The distances from these segments to the observer are approximated by the equations $x_1 = l/\alpha_1$ and $x_2 = l/\alpha_2$, which yield $\Delta x_1 = l\Delta\alpha/\alpha_1^2$ and $\Delta x_2 = l\Delta\alpha/\alpha_2^2$, or $\Delta x_1 = \Delta\alpha x_1^2/l$ and $\Delta x_2 = \Delta\alpha x_2^2/l$. We'll assume that the power of the diffused light per unit length of the beam's path doesn't depend on the distance $x$. In this case, the power of the light scattered in segments 1 and 2 ($\Delta p_1$ and $\Delta p_2$, respectively) are proportional to the lengths of the segments; or, equivalently, to the square of the distances to them: $\Delta p_1 \sim \Delta\alpha x_1^2$ and $\Delta p_2 \sim \Delta\alpha x_2^2$. However, the power of the light traveling to the observer from segments 1 and 2 is inversely proportional to the square of the distances to these segments. Therefore, the power of the light reaching the observer from beam segments 1 and 2, both having the same angular size, does not depend on the distance to these fragments. This means that the beam's trajectory is uniformly bright!

However, at distances of about ten kilometers our approximations are no longer valid. This is due to the fact that at such distances the beam's intensity drops off sharply because of scattering in the atmosphere. Thus the beam becomes dim. In addition, if the beam is directed vertically, the decrease in air density at high altitudes leads to a decrease in the power of the diffused light, which also contributes to a decrease in the brightness of the beam's trajectory. Assuming the length of the visible trajectory to be 10 km, we can estimate that the angular error of the laser pointer is $10^{-4}$ rad for an observer located to one side of the laser pointer at a distance of 1 m. This value is smaller than the angular divergence of the laser beam.

Can astronauts orbiting the Earth see the light from a laser pointer? Apparently it's possible. An estimate obtained by using equation (7) yields a visible range of about 2,000 km. Spacecraft usually orbit the Earth at much lower altitudes (several hundred kilometers). Beam scattering due to fluctuations in air density is significant only in the lower layers of the atmosphere, so this phenomenon won't have much affect on our estimate.

So here's a question for you: if we can see a satellite in the night sky, does that mean the light from our laser pointer can be seen from the satellite? To answer this question, let's compare the intensity of the light received by an observer on Earth and received by an astronaut in an orbiting spacecraft. If we see the spacecraft, it means that the sunlight reflected by the spacecraft and scattered in all directions reaches our eyes in sufficient quantities. Assume the size of the spacecraft to be 3 m. The intensity of sun-

light at the surface of the spacecraft is 1–2 mW/mm$^2$. Our low-power laser pointer generates the same amount of light as reflected by 1 mm$^2$ surface of the satellite—that is, $(1/300)^2 \sim 10^{-5}$ of its radiation. However, this light is concentrated within a solid angle that is smaller by a factor of $10^6$ than the angular divergence of the light reflected by the spacecraft. So, if we can see the spacecraft, the astronauts can see the light from our laser pointer.

## "Do not look directly into the laser beam"

You see this warning in practically any lab where lasers are used. Every scientist knows that laser radiation can cause irreversible damage to the eye. To appreciate the dangers of a laser pointer's beam, let's evaluate its intensity, which is the power incident on 1 mm$^2$ of surface illuminated by the laser. If we assume the diameter of the laser beam is 3 mm and its power is 3 mW, we get an intensity $I_{LP} = 0.3$ mW/mm$^2$. To compare this value with something familiar, let's recall that the intensity of solar radiation is about 1 kW per square meter of the Earth's surface: $I_S = 1$ mW/mm$^2$. So, it's no more dangerous to look at the spot made by a laser pointer than it is to watch a sunbeam playing on the wall. Perhaps this was why the corresponding upper limit was set for the power of lasers used by ordinary consumers.

Remember, though, laser pointers are meant for pointing at things—just as the old wooden pointers were. They're not meant to be pointed at people. A laser pointer should *never* be aimed at a person's eyes. Just as one must be careful not to poke someone in the eye with a wooden pointer, negligence in using a laser pointer can lead to a severe eye injury. This is because the lens in our eye is like a camera lens with a variable focal length; it forms an image on the retina like a camera forms an image on film. If a pencil of light entering the eye is strictly parallel and the lens is focused at "infinity," all of the incident light will be focused and directed by the lens to a single spot on the retina. This spot will be very small (about the wavelength of the incident light, or 1 micron). If, however, the lens is focused at that moment on an object 1 m from the eye, the beam will not be focused sharply on the retina—the size of the blurred spot on the retina will be about 30 microns.

Let's compare the intensities of the light landing on the retina in both cases. In the case of precise focusing, $I_F = 3$ kW/mm$^2$, while for the blurred spot, $I = 3$ W/mm$^2$. It's also instructive to compare these values with the intensity of the light projected on the retina when one looks directly at the Sun. The angular size of the Sun is about 1/100 rad, and the focal length of the lens is about 1 mm. Therefore, the diameter of the Sun's image on the retina will be about 0.1 mm. Assuming that all the sunlight entering the pupil (diameter: 2 mm) is concentrated in a circle with a diameter of 0.1 mm, we get the intensity of the sunlight landing on the retina: $I_S = 0.4$ W/mm$^2$.

These figures convincingly show that one should not look directly into even a low-power laser, because the resulting intensity of the light striking the retina may be $10^4$ times higher than the maximal intensity possible under natural conditions by looking directly at the Sun. On the other hand, if the laser beam "brushes across" eyes focused on some other object (not the laser), only temporary blindness may result without irreparable damage to the eyes. There's no reason to probe the boundary between these two cases. It's much better to heed the warnings and never aim a laser pointer at people.

In the nineties, when laser pointers were very expensive, they were mostly used to indicate targets in shooting galleries. It's not farfetched to think that some people might react to a red spot on their chest by pulling a gun and firing at the person with the laser pointer (the author lives in Florida, where many people have permission to carry a concealed handgun). Many states in the U.S. have enacted legislation that makes it illegal to misuse laser pointers. For example, in California, aiming a laser pointer at people "in a menacing manner" is punishable by 30 days in jail.

## Something to think about

What's brighter: a 5-mW laser pointer, the Sun, or a 1,000-W electric bulb? By definition, brightness is the light radiated in a unit solid angle from a unit surface of the radiating body. Take a sheet of paper and illuminate it alternately with a laser pointer, a sunbeam, and the light from a powerful bulb placed 10 cm away and equipped with a reflector. Calculate the power landing on a unit area of illuminated surface and compare the data obtained. Now imagine a small lens instead of the paper. Estimate the ratio of the brightness at the focal plane of the lens in all three cases and show that you obtained the brightness ratio for the three sources of light. Do you know now why a laser's brightness is tens of thousands of times that of the Sun? ◙

***Quantum*** on lasers and light propagation in the atmosphere:


Y. Nosov, "Lightning in a Crystal," November/December 1990, pp. 13–16.

P. Bliokh, "What Little Stars Do and the Big Old Planets Don't," March/April 1994, pp. 22–27.

V. Surdin, "Optics for a Stargazer," September/October 1994, pp. 18–21.

D. Tarasov and L. Tarasov, "The Play of Light," May/June 1996, pp. 10–13.

A. Buzdin and S. Krotov, "Why is the Sky Blue?," March/April 1998, pp. 47–48.

D. Panenko, "Diffraction in Laser Light," March/April 1999, pp. 33–35.

V. Surdin and M. Kartashev, "Light in a Dark Room," July/August 1999, pp. 40–44.

V. Surdin, "The Eye and the Sky," January/February 2000, pp. 16–20.

# Liberté, égalité, géométrie

## *Gaspard Monge—the father of descriptive geometry*

by V. Lishevsky

MANY SCIENTISTS HAVE had remarkable fates, but few of them lived lives as interesting and full of adventure as Gaspard Monge. He was a talented scientist (mathematician, engineer, chemist, and metallurgist), but he was also a prominent figure in the French Revolution. For example, he signed the death sentence of Louis XVI. Born into a poor family, Monge became a revolutionary and a Jacobin, struggling against the privileges of nobility; yet he became a count and a personal friend of the emperor Napoleon. After the monarchy was restored he was expelled from the French Academy of Sciences and died in exile.

Gaspard Monge was born on May 10, 1746, in the small town of Beaune in eastern France. His father was a semiliterate itinerant peddler, but he tried to give his children the best education that was available at the time for members of the Third Estate (lay citizens who did not own land). Two of Gaspard's brothers became professors, just as he did: the youngest, Jean, became a professor of mathematics, hydrography, and navigation; and the middle brother, Louis, was a professor of mathematics and astronomy. It's interesting that Louis Monge participated in the La Perouse expedition (an early scientific exploration of the Pacific Ocean), and was one of the three men who remained alive.

Gaspard started school at the age of six and soon became a top student. After leaving school in 1762, he entered Holy Trinity College at Lyons, where he taught physics while still a student. Gaspard spent the summer of 1764 at home, as he usually did, and there chance intervened in his life in a big way.

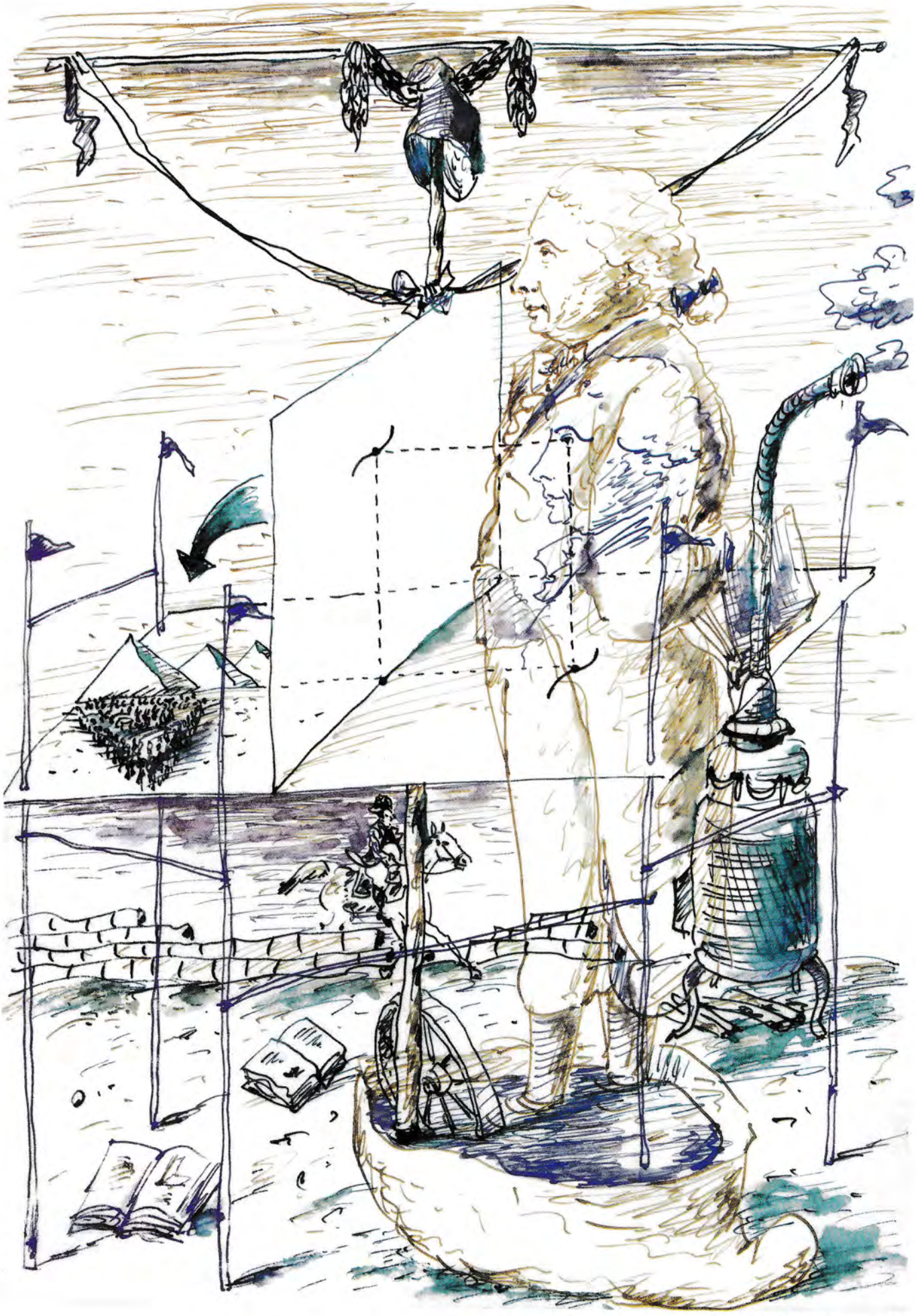
On his days off from school, Gaspard and his friends made a map of his native town. This map caught the eye of a military engineer who headed the military school at Mézières. He invited Gaspard to attend this school, and Monge was admitted to the drafting department. The other department trained military engineers, but only children from noble families could attend.

Monge became interested in a problem that was very important in military engineering: the placement of fortifications so as to make them less vulnerable to guns located at a certain point. Monge solved this problem very quickly, but the professors at first refused to consider his solution, supposing that a student couldn't perform all the complex calculations that were involved. When at last a professor agreed to look at Monge's solution, he was amazed by its simplicity and novel approach. The method was considered so important that it was made a military secret. This is why Monge's method—later called descriptive geometry—remained largely unknown for so long.

### Descriptive geometry

The theory of projection and elements of descriptive geometry were known before Monge. His achievement was to create a new field of science from disjointed facts, individual solutions, and (not always correct) methods of depicting three-dimensional objects. In this sense, Monge can be considered the founder of descriptive geometry, which he defined as *"a method of describing three-dimensional objects on paper having only two dimensions."*

Art by Vadim Ivanyuk



As with many great ideas, Monge's idea was simple. Geometric objects consist of points. So to depict a spatial object, we must find a method for depicting points in space. Consider a point in space, and draw a perpendicular from it onto a horizontal plane. We obtain a projection of the point. However, all points lying on this perpendicular have the same projection. To distinguish among those points, we introduce a vertical plane. Then the two projections (onto the horizontal and the vertical planes) unambiguously determine the position of the point in space (figure 1).

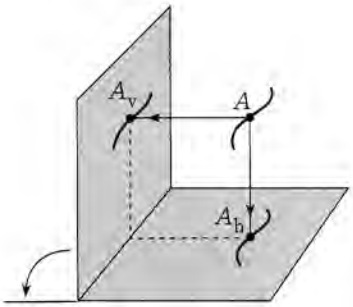



Figure 1

Thus, Monge reasoned, if we depict the orthogonal projections of a point onto two planes, the position is determined unambiguously.

Monge went further. To depict both projections on the same piece of paper, he suggested rotating the vertical plane about the line of its intersection with the horizontal plane. Then all the constructions could be carried out on the same complicated drawing. Using this sort of drawing, it's possible to reconstruct a three-dimensional object, determine the distance between its points, and so on (figure 2).

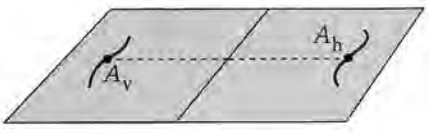



Figure 2

The descriptive geometry invented by Monge underwent certain changes in the course of its development; however, its foundations remained more or less the same. The Russian geometrician B. N. Delone wrote: "*Just as elementary geometry is now presented almost as it was practiced by Euclid and analytic geometry as it was presented by Descartes, descriptive geometry is now seen in a form that is very close to that proposed by Monge.*" His book *Descriptive Geometry* wasn't published until 1799, when the underlying ideas ceased to be a military secret.

## Teacher and scientist

At the age of 23 Monge became a professor at the military school at Mézières. In 1770 he was appointed to a chair in physics, and soon after to a chair in mathematics as well. While teaching, Monge conducted research in various fields; however, he didn't publish anything on descriptive geometry, since it was still a secret.

At this time his mathematical writings appeared in print for the first time—specifically, his work on the theory of the development of surfaces, the calculus of variation, and the integration of certain functions. Monge presented four "memoirs" to the Academy of Sciences, on the calculus of variation, infinitesimal geometry, partial differential equations, and combinatorics. As a result, on April 8, 1772, Monge was elected as a corresponding member of the Paris Academy at the tender age of twenty-five.

Monge devoted much of his time to teaching. He delivered lectures in theoretical and experimental physics, chemistry, mathematics, stone cutting, and the theory of perspective and shadows.

Monge was given to broad gestures while lecturing. In his old age, when it became difficult for him to describe geometric surfaces with his arms, he stopped delivering lectures altogether, remarking that he had "lost his gesture."

Students were fond of the young professor. Monge went on excursions with his students to workshops and factories, strolled in the outskirts of Mézières, and regaled them with many interesting and instructive stories. One of his students later recalled that sometimes Monge waded across a wide creek to get to some factory, not wanting to spend time looking for a bridge, and not interrupting his stream of words. His students followed him (literally and figuratively) without paying any attention to the obstacles strewn in their path—such was the powerful spell Monge cast over these young minds.

In 1777 Monge married a young widow, Marie Catherine Horbou. She was a calm and kind woman; they lived a happy life together and had three daughters.

Madame Horbou inherited a metallurgical workshop from her first husband. Monge became interested in metal processing and particularly in chemistry, as a result of which he organized a chemical laboratory at the Mézières school.

As evidence of his successes in chemistry, consider this: earlier than Lavoisier, Monge proved that water consists of hydrogen and oxygen. He also synthesized water from these gases. (Lavoisier himself acknowledged Monge's priority in this matter.)

However, mathematics remained the primary field of interest for Monge. He developed various practical applications of descriptive geometry, studied partial differential equations, and investigated some aspects of differential geometry.

Monge also participated in the multifaceted activity of the Academy of Sciences. He took part in meetings, worked on various committees, and offered reviews of inventions and scientific papers. At the same time, Monge continued teaching. In 1783, he was appointed examiner of the naval and artillery guards. Reviewing the test results, Monge found that the cadets had a poor grasp of theoretical mechanics. So he wrote a textbook on statics (in 1788).

## Revolutionary

In 1789, revolution burst into flames in France. On July 14, Parisians stormed the Bastille. After Paris, the provinces rose in rebellion. New power structures were organized, as well as new armed forces, called the National Guard. On August 26, the National Constituent Assembly adopted the Declaration of the Rights of Man and Citizen.

The great French scientist Louis Pasteur said that science has no fatherland, but every scientist does. These words certainly apply to Gaspard Monge. He couldn't ignore events—he joined the Patriotic Society, then the People's Society, and finally, the Jacobin Club.

The surrounding countries formed an alliance against revolutionary France. War broke out, and the National Constituent Assembly declared that "the fatherland was in danger." The revolution entered a new stage on August 10, 1792, when the king was dethroned and power passed to a "Provisional Executive Council," consisting of ministers elected by the Legislative Assembly. Gaspard Monge was appointed minister of the Navy and the colonies.

At its first meeting on September 21, a newly elected "Convention" proclaimed the abolition of the monarchy and the establishment of the republic. The king was tried and sentenced to death. Monge, who was the acting chairman of the Council at that time (the ministers occupied this post in rotation), signed the sentence.

The French Republic was in a difficult state. Weapons and food were scarce. Poorly trained, badly armed, hungry soldiers fought against superior enemy forces. On behalf of the revolutionary government, Monge organized the production of gunpowder, guns, and sabers. He found a stock of saltpeter, which was necessary to produce powder. Under his management, iron works began to produce guns (in Paris, as many as 1,000 guns were produced daily). Monge organized foundries to cast gun barrels. He helped train workers and provided food for them, though he was half-starved himself. When his wife offered him a piece of cheese to go with his usual piece of bread, he refused.

As a result of the counterrevolutionary coup on the 9th of Thermidor (July 27), 1794, the leaders of the Jacobin dictatorship—Robespierre, Saint-Just, and others—were executed. They were succeeded by a "Directorate" of five men, who held France together for the next few years. Monge, who was an active Jacobin, had to go into hiding.

The Convention closed the Academy of Sciences and secondary schools. Arms production dropped and many factories and textile mills closed. Monge turned his entire attention to teaching.

## École Polytechnique

Monge played a substantial role in founding the famous École Polytechnique[1] (Polytechnic School) in 1795—for a long time he was its director. The school was his favorite creation; he gave it all his spare time and even money (for scholarships). The school lived up to its expectations. At various times, such prominent scientists as Ampère, Coriolis, Gay-Lussac, Becquerel, Arago, Fresnel, Poinsot, and Poisson studied there, along with many generations of brilliant engineers.

The prominent engineer Brisson, who studied under Monge, recalled that nobody could teach as well as Monge. He used gestures, poses, and changes in his voice to develop and explain ideas. He followed his students' eyes and saw how well they understood the lecture. Monge was a real friend to the students—he used every means to extend their range of interests and talents, and was always glad to be of help. Another student, Dupin, described Monge's appearance: *"He was tall, strong, and muscular. His face, broad and short, resembled a lion's. The eyes were large, lively, and sparkled from beneath his dense black eyebrows. His forehead was high, with deep wrinkles that showed his sharp intellect. His remarkable face was usually calm—the face of a man deep in thought."*

Monge wrote several textbooks on descriptive, analytic, and differential geometry that were used by several generations of students.

## Monge and Napoleon

The situation in France took its course. In February 1796, the Directorate appointed the 26-year-old general Bonaparte to commander-in-chief in Italy. In May of the same year Monge went to Italy on behalf of the Directorate. There he met the future emperor, and this acquaintance played an important role in Monge's life.

Bonaparte and Monge had met earlier (when Monge was minister of the Navy); however, Monge didn't remember his visitor. In Italy, Bonaparte recalled their meeting: *"A young artillery officer visited the minister of the Navy in 1792; he may not have remembered that occasion—he had many visitors. But that obscure officer will remember his kindness forever."*

A trusting relationship quickly developed between the general and the scientist. It was the mutual attraction of two intelligent people; later, it developed into a warm friendship. Although relations between them were not always serene, Bonaparte found in Monge a real friend who remained faithful up to his death.

[1] To learn more about École Polytechnique, read the article "Revolutionary teaching" in the March/April 1998 issue of *Quantum*.

When Bonaparte undertook the Egyptian campaign in 1798–1799, Monge took part in it. This expedition nearly killed Monge—he was taken ill with plague, and the fact that he didn't die was due to the care he received from the renowned chemist Berthollet.

It was in Egypt that Napoleon uttered the immortal phrase: *"Put the donkeys and the scientists in the middle!"* Some people see this as showing disrespect toward scientists. Recognizing the sense of humor of the future emperor, we must note that he was placing in the middle of the square those things that were most valuable: scientists and animals that carried weapons, water, and food.

In Egypt, Monge and other scientists conducted scientific research. Their aim was to contribute to progress and education in Egypt. To this end the Cairo Institute was established; Monge was elected its president and Bonaparte its vice president. The French scientists compiled a "Description of Egypt," studied antiquities and agriculture, and worked on a project to build a canal linking the Mediterranean and Red seas.

However, the French army's situation got progressively worse, and not only in Egypt. The Russian general Suvorov defeated the French in Italy, and the situation on other fronts was just as dire. Napoleon decided to return to Paris. In August 1799, he left the army and sailed for France. Monge, Berthollet, Murat, and others were with him. On October 16, 1799, Bonaparte arrived in Paris. He was welcomed by crowds of enthusiastic people.

On the 18th of Brumaire (November 9), 1799, the Directorate and then the Parliament were abolished. Power passed to three consuls, but in fact all power was concentrated in the hands of the first consul—Napoleon Bonaparte. On December 24, 1799, the first consul appointed Monge senator for life.

Monge left the post of director of the École Polytechnique, but remained a professor there. He continued his studies in applying algebra and calculus to geometry. He also made a major contribution to the theory of machinery.

On August 21, 1803, Monge was appointed vice president of the Senate, and on September 23 the senator from Liège. The Senate administration was mainly carrying out the instructions of the first consul. In particular, Monge was given the task of organizing the production of guns at Liège.

At the end of 1803 Napoleon restored the status of personal awards abolished by the revolution. Monge was the first civilian to receive the Royal Order of the Legion of Honor. Napoleon said: *"I'm envious of you scientists—you must be happy to become famous without besmirching yourself with blood."*

On May 18, 1804, a new Constitution was adopted in France. Napoleon was made emperor for life. Monge carried out various orders that came from the emperor. In particular, he studied the feasibility of constructing a canal from the river Ourcq to Paris; worked on a project for an airborne assault on England using 100 balloons, each 100 meters in diameter, and so on.

On May 20, 1806, Napoleon appointed Monge president of the Senate. Soon he was made a count and received 100,000 francs to purchase an estate. Monge was at the pinnacle of his career, but his health was deteriorating. In the beginning of 1809 he lost the use of an arm. He had to abandon teaching, but continued to advise the emperor on various scientific issues. In 1810 Monge headed a commission studying rockets and gave a report on a study of armor and a monograph on the metallurgy of iron and steel. The emperor asked his advice about the foundries of Tuscany, ores on the island of Elba, the production of cannons, and many other issues.

Eventually Napoleon's empire went into decline. The defeat of his Grand Army in Russia and at the "Battle of the Nations" near Leipzig led to Napoleon's abdication and exile.

When Napoleon returned briefly from his first exile in Elba, Monge came to the Tuileries palace on the first day after his return to the throne. After Napoleon's second abdication, Monge had to leave France and went to Belgium, where he died on July 28, 1818.

His body was moved to Paris and buried in the famous Père-Lachaise cemetery. There was no official ceremony, but many academicians, friends, and students came to pay homage.

Monge is known in the history of science as the inventor of descriptive geometry, as the man who made draftsmanship a working tool of engineers. The well-known Russian scientist V. I. Kurdyumov said: *"If the blueprint serves as the language of engineering, then descriptive geometry is its grammar, since it shows us how to read the ideas of others and to present our own; Monge is the creator of this universal language."* We also must remember his work in other fields of mathematics (calculus and differential geometry), as well as in chemistry, metallurgy, meteorology, optics, hydraulics, and arms and glass production. Monge even suggested a hypothesis for the origin of life on Earth. His life is an impressive example of service to science. ◘

CONTINUED FROM PAGE 13

## M310

*Problem solvers.* Eight students were solving eight problems. It turned out that every problem was solved by exactly five students. Prove that there exist two students such that every problem was solved by one or the other of them. What if every problem was solved by exactly four students?

(N. Vasilyev and S. Tokarev)

ANSWERS, HINTS & SOLUTIONS ON PAGE 51

# Olympic Recap from England

by Mary Mogge

ALL FIVE REPRESENTATIVES of the 2000 United States Physics Team won medals at the XXXI International Physics Olympiad held in Leicester, England, from 8-16 July. Overall, 296 students from 64 countries competed and were awarded a total of 15 gold medals, 11 silver medals, and 42 bronze medals. While the competition is among individuals, unofficial rankings placed the US seventh after China, Russia, India, Hungary, Iran, and Taiwan. China also had the top student, Lu Ying, who scored 43.4 out of 50 points. China's team was the only team with five gold medals. The US won one silver and four bronze medals and was one of only five teams to receive five medals.

Gregory Price of Falls Church, Virginia, was the top US competitor, placing 16th and receiving a silver medal. He is a student at Thomas Jefferson High School for Science and Technology in Alexandria, Virginia, and was nominated by John Dell.

Bronze medalist Anthony Miller graduated from Hopewell Valley Central High School in New Jersey, where he studied physics with Mary Yeomans. This fall he is a student at Princeton. Jason Oh also won a medal at the 1999 Olympiad held in Padua, Italy. Jason graduated from the Gilman School in Baltimore, Maryland, where he studied physics with Edwin Lewis. This fall he is a student at the California Institute of Technology. Michael Vrable of Del Mar, California, graduated from Torrey Pines High School. He was nominated by his physics teacher William Harvie and is currently a student at Harvey Mudd College. Joseph Yu graduated from University High School in Irvine, California, where he studied physics with Glenn Malin. This fall Joseph is a student at the Massachusetts Institute of Technology.

## Selection and training

The selection process for the 2000 U.S. Physics Team began in January, when high school teachers throughout the country nominated over 1300 students. The first round of examinations in late January produced approximately 175 semifinalists who were given a second screening examination in March. Using the results of the second examination, transcripts, and letters of recommendation, the 24 members of the team were selected. More information about the team and its members and additional photos can be found at the US Physics Team website, www.aapt.org/olympiad.

The team members met at the University of Maryland for a nine-day extensive training camp in late May. Their activities at the camp included tutorials, laboratories, problem sets, examinations, and guest lectures on current research topics. At the end of the training camp, five team members were selected to represent the US at the Olympiad.

They and alternate Sean Markan reconvened at the University of Maryland on 2 July for a three-day mini-camp devoted to enhancing their laboratory skills. Then it was on to England accompanied by coaches Mary Mogge and Leaf Turner.

*Medal winners (left to right): Gregory Price, Jason Oh, Anthony Miller, Joseph Yu, Michael Vrable*

## Shifts of parity and time

The travelers arrived in England two days before the start of the competition to adjust to the time and parity shift. Just when our travelers had begun to walk on the left-hand side of the sidewalk and look right first when crossing a street, they would encounter a roundabout where traffic circled clockwise. While they adjusted, there were also sights to see—Big Ben, the Houses of Parliament, and, catty-corner across the Thames,

the 135-meter London Eye, a giant cantilevered Ferris wheel. There were conveyances to ride—a double-decker bus, a tour boat on the Thames, and the Underground. There was cuisine to taste—sandwiches with unusual fillings, strangely flavored potato chips, and "sweets" for dessert.

Then on Saturday, it was on to the University of Leicester. The University is the site of one of the largest space research laboratories in Europe and there are rockets displayed in the lobby of the Physics Building. The town of Leicester was first settled in Roman times and as Britain's first "environment city" features many parks and open spaces. Castle Park is Leicester's "Old Town," containing ancient walls, historic buildings, shops, and restaurants.

## The exams

The five-hour theoretical exam on 10 July consisted of three problems. The first had five major subparts involving a bungee jumper, a Carnot engine, the age of the Earth from radioactive decay, the total electric energy associated with a charged sphere, and a circular ring of thin copper wire rotating in the Earth's magnetic field. The second theory problem modeled two different ways of experimentally determining the charge-to-mass ratio of the electron. In Part A of the third question, students investigated problems associated with detecting gravitational waves using a detector consisting of two perpendicular rods. Part B of the problem concerned the effect of a gravitational field on the propagation of light in space.

The five-hour experimental examination on 12 July consisted of two experiments. In the first experiment, the students determined how the conductance of a light-dependent resistor varied with wavelength across the visible spectrum. They needed to correct for the energy distribution of the emitted light. The second experiment was an investigation of the motion of a magnetic puck as it slid down a *U*-shaped aluminum track. The students were asked to propose a theory and design an experiment to check how the force acting on the puck depended on velocity and track inclination.

Links to recent exams can be found at the International Physics Olympiads website, www.jyu.fi/tdk/kastdk/olympiads.

## Sir Isaac and Alice

Amidst the rolling green hills of England, it took very little effort to imagine an apple falling from a tree—or a tardy rabbit being pursued by a little girl in a pinafore. When not challenged by interesting physics problems, the students toured Cambridge, where Newton studied and taught, or Oxford, where Halley observed and Charles Dodgson wrote as Lewis Carroll. The participants had the opportunity to have lunch at a college of the university. Entering the almost cloisterlike walled college gardens provided welcome refuge from streets bustling with too many tourists.

The students had other opportunities to experience the intersection of Newton's laws and whimsy when they visited Alton Towers, a castle-themed amusement park with gardens and more than its share of plunging thrill rides. Or when they took part in a simulated space mission at the Challenger Learning Center. The center, located in Leicester, promotes hands-on learning of science and technology.

## The 2000 United States Physics Team

The other members of the US Physics Team (with their teachers and high schools) are Badr Albanna (George Lang, Sidwell Friends School, Washington, DC), Dario Amodei (Richard Shapiro, Lowell HS, San Francisco, CA), Owen Baker (Michael Morrill, Columbia HS, Maplewood, NJ), Brian Beck (Robert Shurtz, Hawken School, Gates Mills, OH), Jeffrey Bruidge (LBJ HS, Austin, TX), Kevin Chan (Adam Weiner, The Bishop's School, La Jolla, CA), Susan Dorsher (Gary Anfenson, St. Cloud Technical HS, St. Cloud, MN), David Gaebler (Beth Markham, home schooled, Cedar Rapids, IA), Charvak Karpe (Pratima Karpe, home schooled, Stillwater, OK), Olivia Leitermann (Ronald Francis, Andover HS, Andover, MA), Samuel Lindsay-Levine (Digby Willard, St. Paul Central HS, St. Paul, MN), Sean Markan (Richard Dower, The Roxbury Latin School, West Roxbury, MA), David Marks (Jonathan Bennett, North Carolina School of Science and Mathematics, Durham, NC), Nilah Monnier (Caroline Evans, Brookline HS, Brookline, MA), Vladimir Novakovski (John Dell, Thomas Jefferson HS, Alexandria, VA), Michael Rolish (Hirenda Chatterjee, Cherry Hill HS, West Cherry Hill, NJ), Abigail Shafroth (Manu Patel, T.C. Williams HS, Alexandria, VA), Ryan Timmons (Leonard Klein, Wylie E Groves HS, Beverly Hills, MI), Brian Tsang (Stan Eisenstein, Centennial HS, Ellicott City, MD)

Assisting the author were Leaf Turner—senior coach (Los Alamos National Laboratory, NM), Warren Turner—coach (Brunswick School, Greenwich, CT), Boris Zbarsky—junior coach (MIT undergraduate, member of the 1996 and 1997 teams, and gold medalist in 1997), Jennifer Catelli—senior lab assistant and Ryan McAllister—lab assistant (both University of Maryland graduate students). The support staff is headed by Maria Elena Khoury and Annette Cole-man at the American Association of Physics Teachers. Major financial support is provided by AAPT, the American Institute of Physics, and its member societies.

The XXXII International Physics Olympiad will be held in Antalya, Turkey, from 28 June to 6 July 2001. If you are interested in applying or nominating a student and do not receive an application by early December, please contact Maria Elena Khoury at AAPT [301 209-3344 or mkhoury@aapt.org]. ◙

**Mary Mogge** *(professor of Physics at California State Polytechnic University–Pomona) has been a coach of the US Physics Team since 1995 and is currently academic director.*

# Nonrepeating, Patternless, and

THE NUMBER π = 3.14159 26535897932384626433832 975028841971...,
which equals the ratio of the circumference of a circle to its diameter, has been attracting the attention of mathematicians for thousands of years. For a long time, mathematicians dealt only with integers and fractions representable as the ratio of two integers—these numbers are called rational. All attempts to represent π in this form failed.

The number π occurs in the formula for the area of a circle, $S = \pi R^2$, and many mathematicians, both professional and amateur, tried to solve a famous problem: construct a square that is equal in area to a given circle using only a compass and straightedge. This problem was so well known that any other difficult problem was eventually compared with it, and the term *squaring the circle* has become synonymous for an unsolvable problem.

The notation π comes from the Greek word περιμετρ, which means perimeter.

Mathematicians in ancient Greece knew how to construct a square whose area is twice that of a given square: one merely constructs a square whose side length is equal to the diagonal of the given square (figure 1). However, all attempts to represent the length of the side of this square in terms of the side of the given square using only rational numbers failed. This fact was understood by the followers of Pythagoras, and it undermined the confidence of mathematicians that the number π might be represented as the ratio of two integers. From this time on, a competition began to calculate π with greater and greater accuracy.

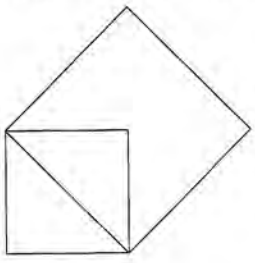


Figure 1

The ancient Egyptians often set π equal to 3; this is equivalent to setting the length of the circumference of a circle equal to the perimeter of the inscribed hexagon. At the same time, Egyptians used the formula

$$S = \left(\frac{8}{9}d\right)^2$$

for calculating the area of a circle. In effect, they were equating π to

$$\left(\frac{16}{9}\right)^2 = 3.16049...$$

Other approximations of π can be found among the records of many ancient civilizations. In the sacred books of the Jains (an Indian religious sect), we encounter an approximation of π as $\sqrt{10}$ = 3.1622777..., and in ancient Chinese texts π was sometimes approximated by the fraction 355/113 = 3.1415929...—an astoundingly high degree of accuracy! But this became apparent only in modern times, since we have been able to compute π to many decimal places. At the time it was unclear which approximation was better: 355/113 or the simpler number 22/7, which was used by the ancient Greeks. Note that 22/7 = 3.1428571... .

In the fifth and fourth centuries B.C., the Greek mathematicians suggested using polygons inscribed in a circle and circumscribed about it

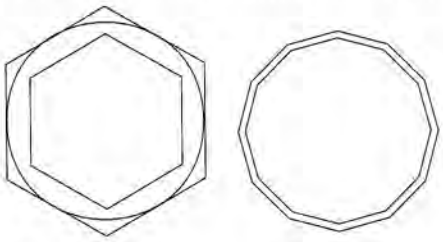


Figure 2

(figure 2) for finding approximate values of π. They noticed that the perimeter of circumscribed circle is greater than, and the perimeter of inscribed circle less than, the circumference of the circle. This idea was exploited by Archimedes, who found the perimeters of the inscribed and circumscribed 6-, 12-, 24-, 48-, and 96-gons using formulas for doubling the number of sides of a polygon. It's surprising that Archimedes could make these precise calculations at that time, repeatedly computing square roots with very high accuracy. As a result, he concluded that π is within the range

$$3\frac{1137}{8069} \text{ to } 3\frac{2669}{18693}$$

—that is, between 3.140995 and 3.142826.

Claudius Ptolemy, who was famous not only as the inventor of the heliocentric planetary system but also as a mathematician, computed the perimeter of the regular inscribed 720-gon and obtained for π the value 377/120 = 3.14166... . He also introduced the notions of angular degree, minute, and second.

# Ter etually a roximated

The next step was made by Francois Viète fifteen hundred years later. He calculated the perimeter of regular inscribed and circumscribed 393,216-gons and obtained the estimate

$$3.1415926535 < \pi < 3.1415926537.$$

This estimate yields 10 valid decimal places for $\pi$. The Dutch mathematician Adrian van Roomen used a $2^{30} = 1{,}073{,}741{,}824$-gon to obtain 17 valid decimal digits. The last mathematician who took this approach was the Dutch mathematician Ludolf van Ceulen. He spent ten years calculating the perimeters of regular polygons by doubling the number of their sides—the same method used by Archimedes. Ludolf reached the 32,512,254,720-gon and obtained 20 valid decimal digits for $\pi$. He concluded his work with the words, "Whoever has the desire, let him go further." In fact, he himself went further and obtained 35 valid decimal digits for $\pi$.

The story of $\pi$ and its approximation continued. Around the turn of the 19th century, the concept of the limit in calculus made it possible, among other things, to consider the sums of an infinite number of summands. In 1671, James Gregory found that the function $\arctan x$ can be represented as an infinite series

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots$$

For $x = 1$, this series (known as the Leibniz series after one of the inventors of the calculus) yields

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots$$

We may group the terms of this series in two ways:

$$\frac{\pi}{4} = \left(1 - \frac{1}{3}\right) + \left(\frac{1}{5} - \frac{1}{7}\right) + \left(\frac{1}{9} - \frac{1}{11}\right) + \dots$$

and

$$\frac{\pi}{4} = 1 - \left(\frac{1}{3} - \frac{1}{5}\right) - \left(\frac{1}{7} - \frac{1}{9}\right) - \left(\frac{1}{11} - \frac{1}{13}\right) - \dots .$$

It's clear that the terms in parentheses are positive. Thus we see from the first equality that, taking an even number of terms, we obtain a number just shy of $\pi/4$; taking an odd number of terms, we obtain a number just a little larger than $\pi/4$.

This series made it a lot easier to calculate $\pi$, although it requires no less than 50 expansion terms to obtain three valid digits; for four valid digits, about 300 terms are required.

Abraham Sharp noticed that with $x = \sqrt{3}/3$, we have

$$\frac{\pi}{6} = \frac{\sqrt{3}}{3}\left(1 - \frac{1}{9} + \frac{1}{45} - \frac{1}{189} + \frac{1}{729} - \frac{1}{2673} + \dots\right),$$

and the first six terms of this series yield $\pi$ with an error of less than 0.0005.

Leonard Euler also took part in calculating $\pi$—he used the relation

$$\frac{\pi}{4} = \arctan\frac{1}{2} + \arctan\frac{1}{3}$$

and found out that Lagny, who had earlier calculated 128 decimal digits of $\pi$, made a mistake at digit 113 (and, therefore, the subsequent digits were also wrong).

The formula

$$\pi = 24\arctan\frac{1}{8} + 8\arctan\frac{1}{57} + 4\arctan\frac{1}{239}$$

turned out to be even more convenient, since the terms of the series decrease more and more quickly as the argument of the arctangent becomes smaller.

The middle of the 19th century was marked by a pursuit of more decimal digits of $\pi$:

1844: 200 digits (Dase);
1847: 248 digits (T. Klausen);
1853: 330 digits (Richter);
1853: 440 digits (Dase);
1853: 519 digits (W. Shanks).

A hundred years later, after computers were invented, the chase continued:

1949: 2,037 digits (von Neumann, ENIAC);
1958: 10,000 digits (F. Jenuit, IBM-704);
1961: 100,000 digits (D. Shanks, IBM-7090);
1973: 1,000,000 digits (J. Guiyu, M. Boiye, CDC-7600);
1986: 29,360,000 digits (D. Bailey, Cray-2);
1987: 134,217,000 digits (J. Kanada, NEC SX-2);
1989: 1,011,196,691 digits (D. and G. Chudnovsky, Cray-2 + IBM-3040).

However, all this has become more of a sport and less of a mathematical activity. It's no wonder the last result was included in *Guinness World Records*. It's interesting that mathematicians have studied the sequence of digits in the decimal representation of $\pi$ and have established that all digits occur in this sequence with the same frequency. ■

# Relativistic conservation laws

by Larry D. Kirkpatrick and Arthur Eisenkraft

CONSERVATION LAWS ARE everywhere! Conservation of energy is one of the most useful laws in all branches of science. Other conservation laws in physics include charge, momentum, angular momentum, and those associated with the more esoteric baryon and lepton attributes. But conservation laws are much more pervasive. They even apply to poker. Certainly, the number of cards is conserved in any one game. The amount of money can also be conserved if the game is carefully constructed.

The World Series of Poker is held in Las Vegas each year. It is not a tournament for the poor or the faint of heart. In the final game each player buys in for $10,000! This year there were 512 players, so the total amount of money in the game was $5,120,000. The game is Texas Hold 'em and the rules require that no money be taken from the game or added to the game—so-called "table stakes."

As the game proceeds, some players accumulate lots of money while others lose. However, at all times the total amount of money in the game is a constant—5.12 million dollars. When players lose all of their chips, they must exit the game. So the number of players is not conserved, but the amount of money is.

After all but one of the players have lost their chips, the winner has accumulated $5,120,000. You can imagine the size of the bets when there are only two players left at the table. No, the winner does not get to keep all of the money; the money is divided among the players according to when they left the game. Of course, those leaving early get smaller amounts and the last player gets the most. This year the winner pocketed $1,500,000 and the second place player walked away with $896,500. Even the players who finished in 37th through 45th place won $15,000.

*In mathematics you don't understand things. You just get used to them.*

*—Johann von Neumann*

Poker may not interest some readers as such, but it is actually one small example from a branch of mathematics called game theory. Games (in mathematics) are situations that involve people or machines with conflicting interests. Simple games can have complete "solutions" but serve as insights into more complex games like checkers and chess and more serious games like politics, warfare, and property law. One type of game is the zero-sum game, where one player's loss is another player's gain. Not all games are zero-sum games. The stock market or the nation's economy can gain value over time.

We are very familiar with the conservation laws of energy and momentum in classical mechanics. Although in special relativity the sizes of time intervals, lengths, energies, momenta, angular momenta, and so on are not the same in different inertial reference systems, the conservation laws are still valid. However, we must modify our classical expressions for energy and momentum to make them work in relativistic situations.

The relativistic momentum $p$ is given by

$$p = \gamma m v,$$

where $m$ is the (rest) mass of the particle, $v$ is its speed, and

$$\gamma = \sqrt{\frac{1}{1-\beta^2}}$$

with $\beta = v/c$, the ratio of the speed of the particle to the speed of light. Notice that the relativistic momentum reduces to the classical value for slow speeds—that is, when $v \to 0$. This must be true because we know that Newton's laws of motion work very well for ordinary speeds.

The photon has no rest mass but it does have momentum. Therefore, we must use a different expression for the momentum of a photon:

$$p = \frac{h}{\lambda} = \frac{hf}{c},$$

Art by Tomas Bunk


TO MY FRIEND VEDRAN JELASKA

where $h = 6.63 \cdot 10^{-34}$ J·s is Planck's constant and $\lambda$ and $f$ are the wavelength and frequency of the photon, respectively.

The relativistic energy of a particle with nonzero mass is given by

$$E = \gamma mc^2.$$

When the particle is at rest $\gamma = 1$, so the rest-mass energy is $E = mc^2$. This is Einstein's famous equation giving the equivalence of mass and energy. The difference between the total energy and the rest-mass energy is equal to the kinetic energy of the particle. We can show that this reduces to the classical formula for the kinetic energy as $v \to 0$:

$$KE = \gamma mc^2 - mc^2 = (\gamma - 1)mc^2.$$

We now use the binomial expansion

$$(1 + x)^n \approx 1 + nx,$$

when $x \ll 1$. This gives us

$$KE = \left(\frac{1}{\sqrt{1-\beta^2}} - 1\right)mc^2$$

$$\approx \left(1 + \frac{1}{2}\beta^2 - 1\right)mc^2 = \frac{1}{2}mv^2.$$

The relativistic energy of a photon is

$$E = hf = pc.$$

With these new expressions, conservation of energy and linear momentum work the same way they do in classical physics. The mathematics can be more difficult because the factor $\gamma$ depends on the speed of the particle.

Our contest problem this month comes to us from the second exam used to select this year's US Physics Team. (See *Happenings* for a report on the success of this year's team—each team member won a medal.) The problem was created by Leaf Turner, who works at Los Alamos National Laboratory and is a senior coach of the team.

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 $f_2 < f_1$.

A. What is the velocity **v** of the particle?

B. What is the rest mass of the particle?

C. What are the frequencies of the photons in the rest frame of the particle?

You are given the formula

$$p'_x = F_1 p_x + F_2 \frac{E_\gamma}{c},$$

where $p'_x$ is the $x$-component of the momentum of either photon in the laboratory reference frame and $E_\gamma$ and $p_x$ are the energy and $x$-component of the momentum, respectively, of the photon in the rest frame of the particle.

D. What are the functions $F_1$ and $F_2$ in terms of $\beta$?

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

## Rolling wheels

Our problem in the May/June 2000 issue of *Quantum* required readers to solve three problems concerning cylinders, hoops, and spheres rolling down inclines. The first two problems, considered standard fare for first-year college physics, were solved correctly by Alex Rifkin, Michelle Chung, and Victoria Butta of Amity Regional High School in Woodbridge, Connecticut. Victoria tried the more difficult and subtle third problem. Their teacher, A. Hovey, correctly solved most of this problem.

A. To show that all uniform, solid spheres arrive at the bottom of the incline with the same speed, independent of their radii and masses, we use conservation of energy. The loss in potential energy is equal to the gain in kinetic energy—both translational and rotational:

$$mgh = \frac{1}{2}mv^2 + \frac{1}{2}I\omega^2$$

or

$$mgh = \frac{1}{2}mv^2 + \frac{1}{2}\left(\frac{2}{5}mR^2\right)\left(\frac{v^2}{R^2}\right).$$

Therefore

$$v = \sqrt{\frac{10}{7}gh}.$$

B. Any object's moment of inertia can be written as $kmR^2$, where $k$ is a constant that depends on the shape of the object. Comparing the relative speeds of a cylinder ($k = 1/2$), a hoop ($k = 1$), and a solid sphere ($k = 2/5$) of the same mass requires us to solve Part A for the general equation:

$$mgh = \frac{1}{2}mv^2 + \frac{1}{2}I\omega^2,$$

which yields

$$v = \sqrt{\frac{2gh}{1+k}}.$$

Notice that the mass of the objects does not appear in the answer and, therefore, does not affect the speed. All objects with the same shape have the same motion. For our three objects, we get

cylinder: $v = \sqrt{\frac{4}{3}gh}$

hoop: $v = \sqrt{gh}$

sphere: $v = \sqrt{\frac{10}{7}gh}$

C. Part C requires us to compare the linear and angular accelerations for three cylinders when they are inclined at an angle $\alpha$. The cylinders all have the same length, outer radius, and mass. The first is solid, the second is a hollow tube with walls of finite thickness, and the third is a hollow tube with walls of the same finite thickness, filled with a liquid of the same density.

This problem requires an analysis of the rolling dynamics of the cylinders. We will need to find the friction required to roll without slipping since the static friction can take on any value less than or equal to $\mu F_N$. (Since it is inconvenient to use $\alpha$ as both the angle of the incline and the angular acceleration, we will refer to the angle of the incline as $\theta$.)

For the linear motion we have

$$\sum F = ma$$

or

$$F_g \sin\theta - F_f = ma.$$

And for the rotational motion,

$$\sum \tau = I\alpha$$

or

$$RF_f = I\alpha = I\frac{a}{R}.$$

Therefore,

$$a = \frac{R^2 F_f}{I}.$$

Solving for the force of friction $F_f$ and the acceleration $a$, we get

$$F_f = mg\sin\theta \frac{I/mR^2}{1+I/mR^2}$$

and

$$a = g\sin\theta \frac{1}{1+I/mR^2}.$$

The limiting angle for rolling without sliding will occur when the frictional force is equal to the normal force $F_N$:

$$\mu mg\cos\theta = mg\sin\theta \frac{I/mR^2}{1+I/mR^2},$$

yielding

$$\tan\theta = \mu(1 + mR^2/I).$$

We must now find the rotational inertia of each cylinder.

1. The first cylinder has a rotational inertia $I = mR^2/2$. Using the equations above, we find that

$$a = \frac{2}{3}g\sin\theta$$

and

$$\tan\theta = 3\mu.$$

2. The second cylinder has a rotational inertia $I = m(R^2 + r^2)$, where $r$ is the inner radius of the cylinder. We can find $r$ by recognizing that the solid cylinder and this hollow tube have the same mass and therefore the densities $\rho$ must be different by a factor $n$, where

$$\rho = \rho_{wall} = n\rho_{solid}.$$

Setting the mass of the solid cylinder equal to that of the tube

$$\rho\pi R^2 l = n\rho\pi l(R^2 - r^2),$$

we get

$$r^2 = R^2\left(\frac{n-1}{n}\right).$$

Therefore

$$I = \frac{1}{2}m(R^2 + r^2) = \frac{1}{2}mR^2\left(\frac{2n-1}{n}\right).$$

The corresponding acceleration and limiting angle are

$$a = \frac{2n}{4n-1}g\sin\theta$$

and

$$\tan\theta = \frac{4n-1}{2n-1}\mu.$$

3. The third cylinder has the same dimensions as the tube but has less mass rotating; that is, the liquid does not rotate due to a lack of friction between it and the walls:

$$r^2 = R^2\left(\frac{n-1}{n}\right),$$

$$m_{tube} = \frac{\pi R^2 l - \pi r^2 l}{\pi R^2 l}m = \left(1 - \frac{r^2}{R^2}\right)m.$$

The rotational inertia is

$$I = \frac{1}{2}m_{tube}(R^2 + r^2) = \frac{1}{2}\left(1 - \frac{r^2}{R^2}\right)m\left(R^2 + r^2\right) = \frac{1}{2}mR^2\left(\frac{2n-1}{n^2}\right).$$

The corresponding acceleration and limiting angle are

$$a = \frac{2n^2}{2n^2 + 2n - 1}g\sin\theta$$

and

$$\tan\theta = \frac{2n^2 + 2n - 1}{2n-1}\mu.$$

Since all of the angular accelerations $\alpha$ are equal to $a/R$, the ratios of the linear and angular accelerations are

$$1 : \frac{3n}{4n-1} : \frac{3n^2}{2n^2 + 2n - 1}.$$

The ratios of the tangents for the limiting angles are

$$1 : \frac{4n-1}{3(2n-1)} : \frac{2n^2 + 2n - 1}{3(2n-1)}.$$

When the cylinders exceed the largest limiting angle and none of them have the necessary friction to roll without sliding, they all have the same linear acceleration:

$$F_g \sin\theta - F_f = ma$$

$$F_f = \mu mg\cos\theta$$

The angular accelerations are given by

$$\alpha = \frac{RF_f}{I} = \frac{R\mu mg\cos\theta}{I}.$$

Since the rotational inertia is different for each cylinder, the corresponding angular accelerations are

$$\alpha_1 = \frac{2\mu\cos\theta}{R},$$

$$\alpha_2 = \frac{2\mu\cos\theta}{R}\frac{n}{2n-1},$$

and

$$\alpha_3 = \frac{2\mu\cos\theta}{R}\frac{n^2}{2n-1}.$$

# Triangular surgery

by O. Izhboldin and L. Kurlyandchik

IN THIS ARTICLE WE'LL DISCUSS several problems in which polygons are sliced up into triangles. For example, a square can be cut into triangles in many different ways (see figures 1–4).

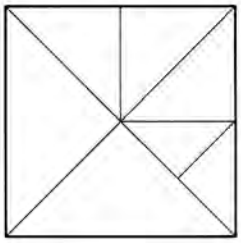


Figure 1

Let's examine these figures carefully. For example, in figures 1 and 4 all the triangles are right triangles, and in figure 2 they are all obtuse triangles. The question naturally arises: can we slice a square into triangles, all of which are acute?

In addition, in all the figures there are at least two triangles with a common side. Is this always the case?

In figures 2 and 3, all the triangles have the same area, and there is an even number of them. Is it possible to cut a square into an odd number of triangles of equal area?

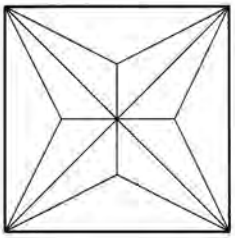


Figure 2

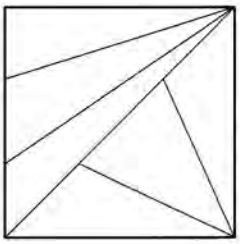


Figure 3

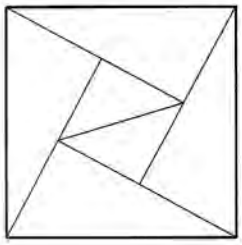


Figure 4

More generally, it will be interesting to find out if a polygon can be broken down into triangles under certain constraints. These may be constraints imposed on the angles of the triangles, their number, their arrangement, and so on.

## Acute triangles

**Problem 1.** *Is it possible to cut a square into acute triangles?*

It's quite natural to begin solving this problem by attempting to cut a square as required. A start might be to cut the square along a diagonal (figure 5) or two diagonals (figure 6). In both cases we reduce the problem to cutting a right triangle into acute triangles. How can we cut a triangle into other triangles? There are three simple possibilities (figures 7–9).

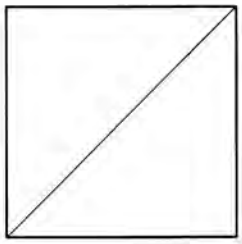


Figure 5

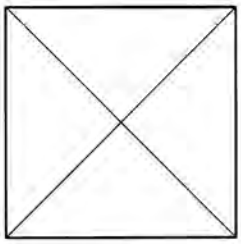


Figure 6

In all three figures, at least one of the resulting triangles is not acute! We invite the reader to try cutting a few triangles—it's very likely you'll end up with the hypothesis that the answer to our question is no. We now have a situation that is familiar to every mathematician: we can continue trying to find the desired

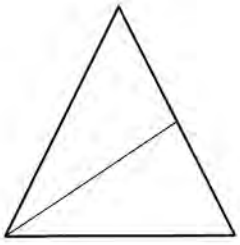


Figure 7

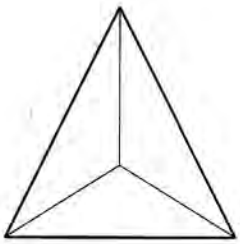


Figure 8

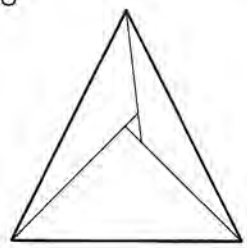


Figure 9

outcome, or we can search for a proof that such an outcome doesn't exist. We invite you to ponder this problem; the answer will be given a little later. Meanwhile, we'll move on to another problem.

## Triangles with clean boundaries

In all the figures shown above, there exists a triangle such that its sides don't contain the vertices of other triangles. We'll call these *triangles with a clean boundary*.

**Problem 2.** *Is it possible to cut a convex n-gon into triangles such that none of them has a clean boundary?*

We'll prove that such a result doesn't exist. Assume the contrary. Let $T$ be the number of triangles created by slicing, and let $V_{\text{int}}$ be the number of "internal" vertices—that is, vertices that lie on the sides of the triangles. It's clear that $V_{\text{int}} \geq T$ since, by our assumption, we can assign to every triangle an internal vertex that lies on its boundary. Notice that different triangles are assigned different vertices, since no vertex can be internal for two triangles simultaneously.

Let's calculate the sum of the angles in all the triangles. On the one hand, it is $180° \cdot T$. On the other hand, the sum of the angles adjacent to internal vertices is $180° \cdot V_{\text{int}}$, and the sum of the angles adjacent to the vertices of the polygon is $180° \cdot (n-2)$. Thus the total sum of the triangles' angles is not less than $180° \cdot V_{\text{int}} + 180° \cdot (n-2)$. Therefore, we have

$$180° \cdot T \geq 180° \cdot V_{\text{int}} + 180° \cdot (n-2) > 180° \cdot V_{\text{int}},$$

which contradicts the inequality $V_{\text{int}} \geq T$ obtained earlier.

Thus we proved the following theorem.

**Theorem 1.** *Whenever a convex n-gon is cut into triangles, at least one triangle has a clean boundary.*

**Exercise 1.** Is this theorem true for nonconvex polygons?

## Without common sides

**Problem 3.** *Is it possible to cut a convex n-gon into triangles such that no two triangles have a common side?*

We begin with the simplest case, where $n = 3$. The desired result is shown in figure 9.

Now try to slice a convex quadrilateral in the desired fashion. It would be quite natural to begin by cutting a square. Look at figures 1–4: each of them contains a pair of triangles with a common side.

Again, we face a dilemma: either try to prove that no such outcome exists or continue the search for the desired outcome.

It turns out that such an outcome is impossible—that is, no matter how a convex $n$-gon ($n \geq 4$) is sliced up, there are at least two triangles with a common side. However, the proof is rather complex.

We'll begin with an important auxiliary proposition.

## The inequality $V \leq T+2$

**Theorem 2.** *Let a convex n-gon be broken down into T triangles and let V be the total number of vertices of those triangles. Then* $V \leq T + 2$.

**Proof.** The sum of all the angles in all the triangles is $180° \cdot T$. We now calculate this sum in a different way. Divide the set of vertices of all the triangles into two parts.

In the first part, we include all the vertices of the given $n$-gon (they are shown in red in figure 10). All the other vertices belong to the second part (they are shown in blue in figure 10).

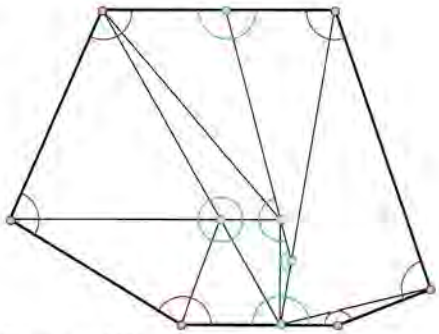


Figure 10

It's clear that the sum of the angles adjacent to the red vertices equals the sum of the angles of the $n$-gon; thus

$$\text{the sum of the "red angles"} = 180° \cdot (n-2).$$

Consider an arbitrary blue vertex. It's clear that the sum of the angles adjacent to it is either 180° or 360° (see figure 10); in any case, it is no less than 180°. Since there are $V - n$ blue vertices, we have

$$\text{the sum of the "blue angles"} \geq 180° \cdot (V-n).$$

Thus

$$\begin{aligned} &180° \cdot T \\ &= \text{the sum of all angles of all triangles} \\ &= \text{the sum of the "red angles"} \\ &+ \text{the sum of the "blue angles"} \\ &\geq 180° \cdot (n-2) + 180° \cdot (V-n) \\ &= 180° \cdot (V-2). \end{aligned}$$

The desired inequality follows:

$$V \leq T + 2.$$

The theorem is thus proved.

## Solution to problem 3

Assume that we have cut a convex $n$-gon into triangles in such a way that no two of them have a common side.

Let's calculate in two different ways the number of segments that are sides of the triangles. We'll call these segments "sides" and denote their number by $S$. It's clear that

$$S = 3T,$$

since every triangle has three sides and no sides of two different triangles coincide.

We divide the set of vertices and the set of sides into two classes:

(i) Boundary vertices and sides—that is, those that lie on the boundary of the given $n$-gon. We denote the number of boundary vertices by $V_b$ and the number of boundary sides by $S_b$.

(ii) Internal vertices and sides—that is, all those that aren't boundaries. These are denoted by $V_{in}$ and $S_{in}$, respectively.

It's clear that

$$V = V_b + V_{in},$$

and

$$S = S_b + S_{in}.$$

We now establish a relationship between the number of boundary sides, $S_b$, and the number of boundary vertices, $V_b$. We do this by "taking a tour" of the boundary of the given $n$-gon. During this tour the "vertices" and "sides" alternate, which means there are just as many of each:

$$V_b = S_b.$$

There exists also a relationship between the number of internal sides $S_{in}$ and the number of internal vertices $V_{in}$; however, it's more complex:

$$3V_{in} \geq S_{in}.$$

To prove this inequality, we select those internal vertices that lie within a side. We'll call such vertices interior and denote their number by $V_{int}$. Clearly $V_{in} \geq V_{int}$. Therefore, to prove the inequality $3V_{in} \geq S_{in}$, it's sufficient to prove that

$$3V_{int} \geq S_{in}.$$

This inequality will be proved if we're able to assign to every interior vertex three internal sides such that every internal side corresponds at least to one vertex. Such a correspondence exists and is illustrated in figure 11.

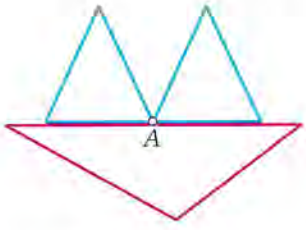



Figure 11. *Vertex* A *corresponds to one red side, for which this vertex is interior, and two blue sides.*

We must make sure that every internal side corresponds at least to one interior vertex. This is obvious for sides that contain a vertex. If an internal side doesn't contain any vertices, it's a part of a side of another triangle. Then at least one of its endpoints is an interior vertex, and it is this vertex that corresponds to the side under consideration.

Thus $3V_{int} \geq S_{in}$ and, therefore, $3V_{in} \geq S_{in}$. Therefore,

$$3T = S = S_{in} + S_b \leq 3V_{in} + V_b = 3(V_{in} + V_b) - 2V_b = 3V - 2V_{in} \leq 3V - 2n.$$

Consequently,

$$V \geq T + {}^2/_3 n.$$

Now, since $n \geq 4$, we have

$$V \geq T + {}^8/_3 > T + 2,$$

which contradicts theorem 1.

Thus we have proved the following theorem.

**Theorem 3**. *Whenever a convex $n$-gon ($n \geq 4$) is cut into triangles, there exist at least two triangles with a common side.*

Examine the above proof carefully and solve the following exercises.

**Exercises**

**2.** Let a triangle be cut into $T$ triangles such that no segment is a common side of two triangles. Let $B$ denote the total number of vertices of the triangles in the decomposition. Prove that

$$V = T + 2.$$

**3.** Let an $n$-gon ($n \geq 4$) be cut into triangles. Prove that there exist at least $n - 3$ segments, each of which is a common side of two of the triangles.

**4.** Theorems 2 and 3 include the condition that the polygon be convex. Are these theorems true for nonconvex polygons?

## Acute triangles revisited

In two of the problems we solved in the preceding sections, the desired division of the polygon into triangles didn't exist. It may appear that if we can't perform the desired division the first time, it doesn't exist at all. However, this isn't the case! Indeed, return to problem 1, where we wanted to cut a square into acute triangles. Although we couldn't do it on our first try, such a result is possible and is shown in figure 12.

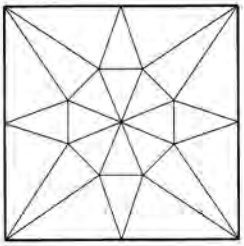


Figure 12

**Exercises**

**5.** In figure 12, there are 24 triangles. Is it possible to achieve the intended result with fewer triangles?

**6.** Prove that any convex polygon can be cut into acute triangles.

## Triangles of equal area

**Problem 4.** *Is it possible to cut a square into an odd number of triangles of equal area?*

The statement of this problem is similar to the problems solved above. However, it is much more difficult. The answer is no. The authors have not been able to prove this fact using elementary methods and would be most appreciative if someone could furnish one. Alert readers may recall a very difficult solution given in the article "2-adic Numbers" in the July/August 1999 issue of *Quantum*.

## An assortment of decompositions

We've merely grazed the surface of the problem of cutting polygons into triangles. In this field, many interesting problems can be formulated, and each of them could be the subject of a research paper. Variations are numerous—we haven't even exhausted the cutting of a square. We invite our readers to solve the following problems.

**Problems**

**4.** In figure 2, the square is cut

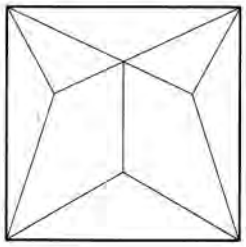


Figure 13

into 12 obtuse triangles; in figure 13, it is cut into 10. Is it possible to cut a square into a smaller number of obtuse triangles? What is the minimum number of triangles that are possible in such a decomposition?

**5.** Is it possible to cut a square into triangles, no two of which are the same, such that all of the triangles are

(i) right triangles,
(ii) isosceles triangles,
(iii) isosceles right triangles,
(iv) similar to each other,
(v) of equal perimeters,
(vi) of equal area?

**6.** Is it possible to cut a square into

(i) "very obtuse" triangles—that is, is it possible to cut a square such that one of the angles of every triangle is greater than 120°? Greater than 179°?

(ii) "almost equilateral" triangles with all angles less than 70°?

(iii) triangles with given angles $\alpha$, $\beta$, and $\gamma$ (for example, with angles of 30°, 60°, and 90°)?

(iv) Find all angles $\alpha$ such that a square can be broken down into triangles whose angles are all less than $\alpha$.

**7.** Is it possible to cut a square into triangles such that every triangle has exactly

(i) two neighbors,
(ii) three neighbors,
(i) $n$ neighbors (where $n$ is a given number)?

Two triangles are called neighbors if they have at least one common point (this is one version of the definition) or a common segment (this is another version). ■

# Our old magnetic-enigmatic friend

by E. Romishevsky

RECENTLY *QUANTUM* PUBlished an article titled "The Enigmatic Magnetic Force" (July/August 2000) that described the properties of the Lorentz force. In this article we'll look at the interplay between its electric and magnetic components and explain the nature of the magnetic force, which affects a current-carrying conductor placed in an external magnetic field.

As a first step, let's investigate the magnetic forces on a conducting bar moving uniformly in a homogenous magnetic field. A magnetic field **B** is perpendicular to the constant velocity vector **v** of a rectangular bar as shown in figure 1. We assume that the bar is thin; its edge length $d$ is much shorter than the edge lengths $a$ and $b$.

The positive ions occupy fixed locations in the bar while the free electrons are evenly distributed throughout the volume of the bar. While the bar moves, the charges experience magnetic forces $\mathbf{F}_M = q\mathbf{v} \times \mathbf{B}$, which act in opposite directions on the positive and negative charges. As a result, the free electrons are shifted downward, so the upper and lower faces of the bar acquire charge densities of $\sigma^+$ and $\sigma^-$, respectively.

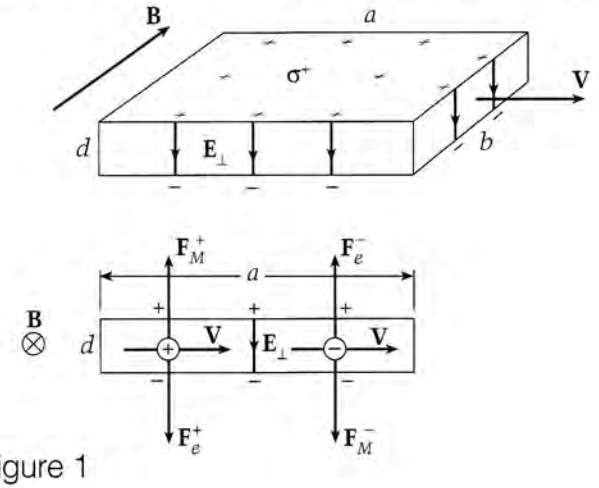



Figure 1

As with a parallel-plate capacitor, these charge densities produce a homogeneous electric field $\mathbf{E}_\perp$ between the upper and lower faces of the bar. The resulting electric forces counterbalance the magnetic forces. In other words, the magnetic force produces an electric field that counteracts its effects:

$$F_M = F_E,$$

or

$$qvB = qE_\perp$$

and

$$E_\perp = \sigma/\varepsilon_0.$$

As a whole, the bar remains electrically neutral because the magnetic field doesn't create any new charges. The field only separates the charges already in the bar. Therefore, $\sigma^+ = |\sigma^-|$ and the total magnetic force acting on a conducting bar moving uniformly in a homogenous magnetic field is zero.

However, if the bar's velocity is increasing, or if the bar enters a region with an increasing magnetic field, a braking force acts on the bar. There will be a corresponding increase in the electric field and the surface charge densities.

A conducting bar moving in a magnetic field is the prototype of the main element in powerful generators used to produce electrical energy. The electromotive force (emf) is produced by the magnetic force $\mathbf{F}_\perp = q\mathbf{v} \times \mathbf{B}$ that we've been examining here. In our case, the emf $= VBd$.

Let's look at another example of the important role played by the magnetic force. Connect a battery with emf $V_b$ to the opposing faces of a fixed metal bar as shown in figure 2. If the resistance between the faces is $R$, the battery generates an electric current $I = V_b/R$ that is homogenous through any cross section parallel to the

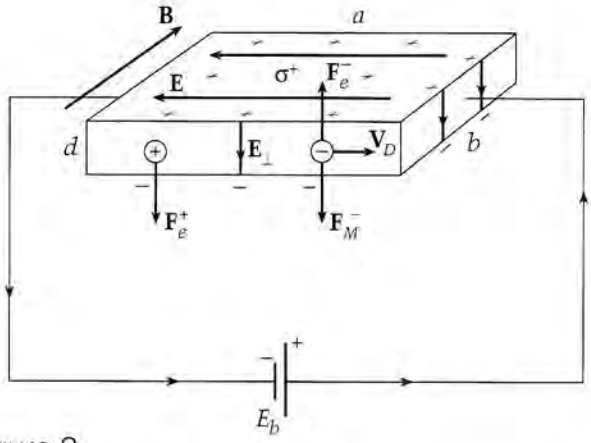



Figure 2

faces. Because the voltage drop between the opposing faces is $V_b$, there is a homogenous electric field **E** within the bar that is perpendicular to the faces connected to the battery.

The free electrons in the bar will be driven by this field to the right with a velocity $\mathbf{v}_d$, creating an electric current

$$I = V_b/R = nev_dbd = nev_dS,$$

where $n$ is the density of the free electrons, $e$ is the charge of an electron, and $bd$ is the cross-sectional area $S$. It should be noted that the velocity of the electrons is opposite the direction of the electric field and the current.

The positively charged ions that form the lattice of the metal are, naturally, at rest during this process. If there is no external magnetic field, the only magnetic field in the system is the internal magnetic field generated by the motion of the free electrons, and this field is negligibly small.

Now let's switch on an external magnetic field with the same strength as in the previous case. Note that the magnetic field must be perpendicular to the direction of the electric current. The electrons, moving with the drift velocity, experience a magnetic force that deflects them downward, thereby producing an extra negative charge on the lower face and an extra positive charge on the upper face.

The charges accumulate until they generate a downward, transverse electric field that counterbalances the magnetic force just as in the case of uniform motion of the bar in a homogenous field. The main difference in this case is that only electrons produce the electric current.

In the steady state (which is established very quickly after the external magnetic field is applied), the average motion of the electrons is again directed horizontally and a transverse electric field $E_\perp = \sigma/\varepsilon_0$ is observed in the reference frame of the bar. This electric field $E_\perp$ counterbalances the magnetic force $F_M = ev_dB$ acting on the moving electrons and creates a force, directed downward, on the motionless positive ions. This is how the magnetic force is transmitted to the metal bar.

The strength of the force acting on a wire of length a placed in a homogeneous magnetic field **B** and carrying a current **I** can be calculated as follows. A positive ion experiences the force

$$F_\perp = eE_\perp = ev_dB.$$

Since the number of positive ions in the bar is $N = nabd$, where $n$ is the density of electrons or positive ions, the total force is

$$F_{tot} = ev_dBnabd = IBa,$$

where $I = nev_dS$ and $S = bd$.

The generation of a voltage drop between the opposite surfaces of a current-carrying wire placed in a magnetic field is called the Hall effect. Edwin Herbert Hall (1855–1938) discovered this phenomenon in 1879, long before J. J. Thomson (1856–1940), discovered the electron. ◙

***Quantum*** on magnetic field:

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

A. Dozorov, "Core Dynamics," March/April 1999, pp. 14–17.

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

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

# Shouting into the wind

by G. Kotkin

WHEN SOMETHING MAKES a noise off in the distance and the wind is blowing away from us and toward the source, why is the sound so faint? At first glance it seems obvious. When you throw a stone into the wind, it doesn't travel as far; a bird heading against the wind doesn't fly as fast. But let's do justice to the title of this magazine and get quantitative.

Let the sound speed relative to the air be $c$ and the wind speed be $u$. Then the speed of the sound signal relative to the Earth is $c - u$, and the time it takes to travel a distance $l$ equals $t = l/(c - u)$. During that time the signal travels (relative to the air) a distance

$$l' = ct = \frac{lc}{c-u}.$$

This value $l'$ determines how much the signal weakens: when we are at a distance $l$ from the source of sound (traveling into the wind), we hear the sound as if we were a distance $l'$ from the source and there was no wind.

Let's assume that $u$ = 15 m/s (a relatively strong wind), $c$ = 330 m/s, and $l$ = 50 m. Then $l' \approx 52$ m. The wind "moves" us nearly 2 m away from the source. We might not even notice it. But actually, the effect of the wind is far stronger. So there must be another explanation.

Note that the wind speed is not the same at every altitude; it increases with height above the Earth. It would seem that this shouldn't affect the propagation of sound when the source and the receiver are at the same altitude. If the receiver (or the source) of a signal is lowered, the signal travels along part of its path through layers of air that move slower. In that stretch, the weakening of the signal will be less pronounced. But in reality, if a person listening (or shouting) sits down, the ability to hear the shout gets worse. Despite this paradoxical result, the key to the puzzle is the dependence of wind speed on altitude.

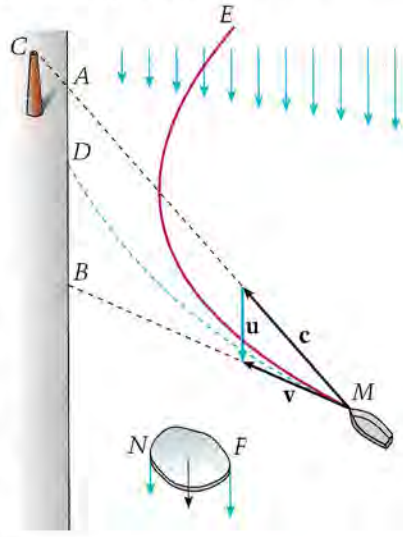



Figure 1

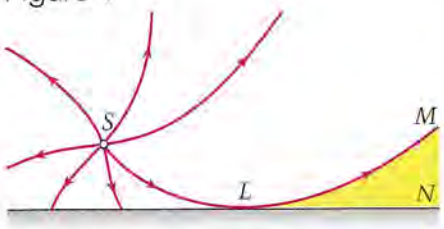



Figure 2

To clarify the peculiarities of sound propagation, let's consider the following example. Imagine a motorboat $M$ moving with velocity $\mathbf{c}$ relative to the water at some angle to the current (figure 1). Let a person sitting in the boat (we'll call her Captain) "aim" at the smokestack $C$ of some factory, which can barely be seen in the distance. If there were no current, the boat would land at point $A$. In fact, the velocity of the boat relative to the riverbank is $\mathbf{v} = \mathbf{c} + \mathbf{u}$. If the velocity of the stream $\mathbf{u}$ were constant the entire distance to the bank, the boat would land at point $B$. But the velocity of the stream decreases as one approaches the shore, so the value and the direction of the boat's velocity $\mathbf{v}$ also vary. As a result, the boat moves along the curve $MD$.

Let Captain fall asleep for a short time while holding the rudder in the same position. While she sleeps, let's look at a drifting ice floe $NF$. While drifting, this ice floe is rotated by the water, which lags behind the floe on one side ($N$) and overtakes it on the other ($F$). The same thing will happen to the boat. It will rotate, so the direction of its velocity will change despite the fact that the rudder is being held firmly in the same position. When Captain wakes up, she will see that the boat is moving along the curve $ME$ and is now heading away from the shore.

This little exercise has made it easier for us to describe the propagation of sound signals. Their trajecto-

Art by Pavel Chernusky

ries (rays) twist like the trajectory of the boat with the sleeping captain (figure 2—the wind is blowing from right to left). The area $MLN$ is a "dead zone"; the sound rays do not reach it. In this area a person can barely hear a sound from the source $S$. (The sound is partly audible in the dead zone because of reflection from the Earth's surface and diffraction.)

Remember, though, that our first explanation was also qualitatively correct, but it was at odds with numerical experimental values. So let's test our second theory with numbers as well. We'll evaluate the distance to the dead zone.

Consider the sound ray $SLM$ (figure 2). Assume that it's almost horizontal. Let's choose a small vertical section $AB$ of the wave front, so small as to be virtually straight (figure 3). After a time $\Delta t$, this section will be shifted to a new position $A'B'$ such that

$$AA' = (c - u_A)\Delta t$$

and

$$BB' = (c - u_B)\Delta t,$$

where $u_A$ and $u_B$ are the wind speeds at the altitudes of points $A$ and $B$, respectively. This displacement will be accompanied by a turn through a small angle $\alpha$:

$$\alpha \approx \tan\alpha \approx \frac{B'B''}{A'B''} = \frac{(c-u_A)-(c-u_B)}{\Delta h} = \frac{\Delta u \Delta t}{\Delta h},$$

where $A'B'' = AB = \Delta h$ and $\Delta u = u_B - u_A$.

Thus the section of the wave front rotates with an angular speed

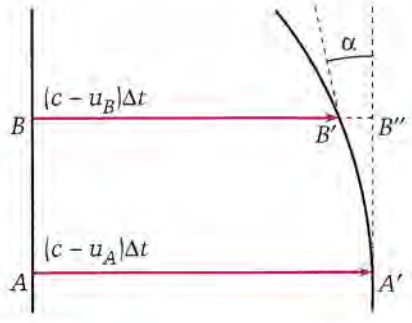



Figure 3

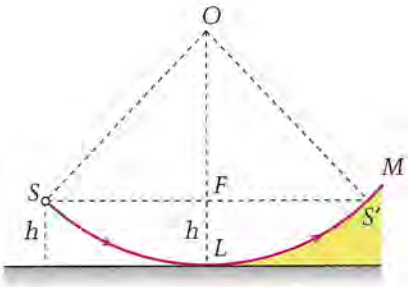



Figure 4

$$\omega = \frac{\alpha}{\Delta t} \approx \frac{\Delta u}{\Delta h}.$$

The vector **c** (sound velocity relative to the air) rotates with the same angular velocity.

Assume the angular speed $\omega$ to be constant along the entire trajectory of the sound ray, so the section $SLM$ (figure 2) can be considered an arc of a circle. The speed of the sound signal relative to the Earth $v = c - u \approx c$ is related to the radius $r$ of this circle as follows: $v = \omega r$. Thus (figure 4),

$$SO = OL = r = \frac{v}{\omega} \approx \frac{c\Delta h}{\Delta u}.$$

Triangle $OSF$ yields

$$SF = \sqrt{OS^2 - OF^2} = \sqrt{r^2 - (r-h)^2} \approx \sqrt{2rh} \approx \sqrt{2ch\frac{\Delta h}{\Delta u}}.$$

Thus an observer located at the same altitude $h$ as the sound source will enter the dead zone at the distance

$$SS' = 2SF \approx 2\sqrt{2ch\frac{\Delta h}{\Delta u}}.$$

In numerical estimates we usually replace the ratio of increments with the ratio of the values themselves. Thus we get $\Delta u/\Delta h \approx u/h$. Finally,

$$SS' \approx 2h\sqrt{\frac{2c}{u}} \approx 3h\sqrt{\frac{c}{u}}.$$

Plugging the values $h = 1.5$ m, $c = 330$ m/s, and $u = 15$ m/s into this formula, we get

$$SS' \approx 20 \text{ m},$$

which is quite reasonable. An exact numerical calculation of the shape of the sound ray yields the same result.

Of course, there are many other factors that reduce our ability to hear a word shouted into the wind, but our calculation gives some assurance that we've found the main culprit.

**Exercises**

**1.** One summer day, a beetle decided to fly to the Sun and took off with a speed $c = 2$ m/s. However, it didn't take into account that there was a light breeze from the south at $u = 1$ m/s. To what angle will the beetle deviate from its target in the reference system of a sparrow sitting on a branch? This heroic flight occurred near Novosibirsk, where the height of the Sun over the horizon at midday is $\alpha = 60°$.

**2.** Find the trajectory of the beetle flying to the Sun (see the previous problem) when the wind speed $u$ increases with altitude $h$ according to a linear law: $u = bh$ (where $b$ = const).

**3.** A plane $AB$ separates a region of still air from air moving with velocity **u** (figure 5). A sound wave arrives at this boundary plane from the area of still air at an angle $\alpha$ to

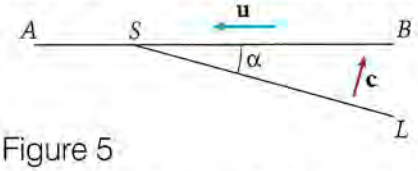



Figure 5

$AB$. The velocity of the incident wave is **c**. At what angle to $AB$ will the wave front of the refracted wave move after entering the region of moving air?

***Quantum*** on sound and sound refraction:


A. Varlamov and A. Malyarovsky, "The Oceanic Phone Booth," May/June 1993, pp. 37–39.

Kaleidoscope: "Songs that Shatter and Winds that Howl," January/February 1994, pp. 32–33.

A. Eisenkraft and L. D. Kirkpatrick, "Sea Sounds," March/April 1996, pp. 34–46; September/October 1996, pp. 36–37.

L. Brekhovskikh and V. Kurtepov, "Waves beneath the Waves," January/February 1998, pp. 16–19.

NSTApress™
NATIONAL SCIENCE TEACHERS ASSOCIATION

# Building the Best Learning Environment...

***with the Science by Design series***


A hands-on approach to the physics of product design. Students apply concepts such as heat transfer, buoyancy, elasticity, or insulation to the development of everyday items. All materials necessary are inexpensive and easy to find.

**PK152X**

**Members: $63.85** Non-Members: $71.82

Buy any 3 in the series and get the 4th volume FREE!

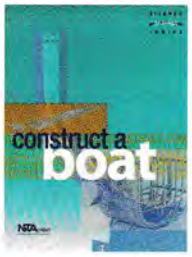

## Construct-a-Glove

(grades 9–12, 2000, 96 pp.)
Physics and technology go hand-in-hand in this practical demonstration of basic concepts in thermodynamics. By testing a simple prototype of an insulated glove, students learn about the process of homeothermic regulation and the variables that influence heat transfer. The challenge to improve upon their initial model introduces them to the design process and the relations between form and function.

**#PB152X1**

**Members: $17.96** Non-Members: $19.95

## Construct-a-Boat




(grades 9–12, 2000, 104 pp.)
By working with a simple model powered by a battery-driven fan, students get a feel for the forces involved in moving a boat through water. Students use their understanding of mass, buoyancy, friction, and acceleration to analyze parameters, improve performance, and design a faster boat. This immersion in learning-by-doing translates abstract concepts into tangible objectives and teaches students seaworthy lessons in modeling and design.

**#PB152X2**

**Members: $17.96** Non-Members: $19.95

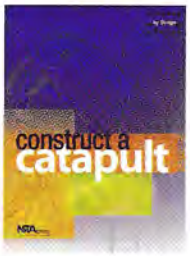

## Construct-a-Greenhouse

(grades 9–12, 2000, 104 pp.)
Students design structures that will convert light to heat during germination and then reconfigure their greenhouses to promote photosynthesis. By comparing the effects of design characteristics on plant growth, students observe the connections between plant biology and basic principles in thermodynamics and energy transfer.

**#PB152X3**

**Members: $17.96** Non-Members: $19.95

## Construct-a-Catapult




(grades 9–12, 2000, 104 pp.)
Launch your students into a hands-on application of concepts such as torsion and elasticity as they learn the physics behind overcoming gravity and hurling objects through the air—SAFELY. By relating the effects of different force settings to distance projections, you can also introduce your students to the analysis of frequency distributions.

**#PB152X4**

**Members: $17.96** Non-Members: $19.95

**To order, call: 1-800-722-NSTA or visit www.nsta.org**

# How many bubbles are in your bubbly?

by A. Stasenko

SOMETIMES A LIQUID "BOILS over" for no obvious reason. This happens when a liquid is pumped up to the surface from deep wells, when a pipe carrying coolant at a power plant bursts, or when a bottle of champagne, beer, or soda pop is opened. Who hasn't looked with pleasure on a carbonated beverage, sparkling with dancing bubbles, on a hot summer day?

When such liquids are transported in pipelines, it's important to know what volume of the gas dissolved in them has separated from the liquid as bubbles. Of course, one could simply take a sample, but in the time it takes to analyze it, what relationship will it have to the mixture at the time the sample was taken? It would be better to use an electromagnetic field—information about how it is changing travels with a speed of the order of the speed of light, so that the actual technological processes will appear "frozen" or *quasi-static*, to use the scientific term.

Let's consider how an ordinary parallel-plate capacitor can test the properties of a liquid flowing through it almost instantaneously. Assume that the liquid has a dielectric constant $\varepsilon$ and contains gas bubbles (of various sizes) in which $\varepsilon_1 = 1$ (figure 1). We'll consider a bubble "large" if its size is comparable to the length $l$ and width $d$ of the capacitor; correspondingly, "small" bubbles are those whose dimensions are much less than $d$.

Let's say we connect the plates of a capacitor (each of area $S$) to a battery with constant emf $V$. We intuitively feel that "something" will differ in the cases when the capacitor is filled with liquid or gas. What is this enigmatic "something" and how should we measure it?

If we neglect the resistance of the connecting wires and the internal resistance of the battery, the voltage drop across the capacitor will always be a constant value $V$. More precisely, the electrical conductance of the gas–fluid mixture is assumed to be negligibly small compared to the conductance of the wires or internal conductance of the battery. Therefore, in the extreme cases under consideration (liquid or gas inside the capacitor), the strength of the electric field in the capacitor will be the same: $E = V/d$. In contrast, the charge on the capacitor will be different in the two cases. Indeed, the capacitance of an empty parallel-plate capacitor is $C_1 = \varepsilon_0 S/d$, while the capacitance of a capacitor filled with a dielectric is $\varepsilon$ times greater: $C_\varepsilon = \varepsilon C_1$. The charge on the filled capacitor is $q_{1\varepsilon} = C_{1\varepsilon} V$. In other words, in the cases considered, the charge and its surface density on the plates will differ by a factor of $\varepsilon$:

$$q_\varepsilon = \varepsilon q_1,\ \sigma_\varepsilon = \varepsilon \sigma_1,$$

where

$$\sigma_1 = \frac{q_1}{S} = \varepsilon_0 \frac{V}{d}.$$

By the way, the strength of the electric field between the plates will be the same at every point, even if the dielectric substance is only partially "inserted" into the capacitor (figure 1). Were this not so, the work performed in moving a test charge along the path *abcfa* would not equal zero, and this is strictly forbidden in electrostatics.

We can see that if at a given moment the dielectric occupies part of the capacitor's length $l'/l$, the total charge on the capacitor will be

$$q = q_1\left(1 - \frac{l'}{l}\right) + q_\varepsilon \frac{l'}{l}$$

$$= \frac{\varepsilon_0 SV}{d}\left(1 + \frac{l'}{l}(\varepsilon - 1)\right). \qquad (1)$$

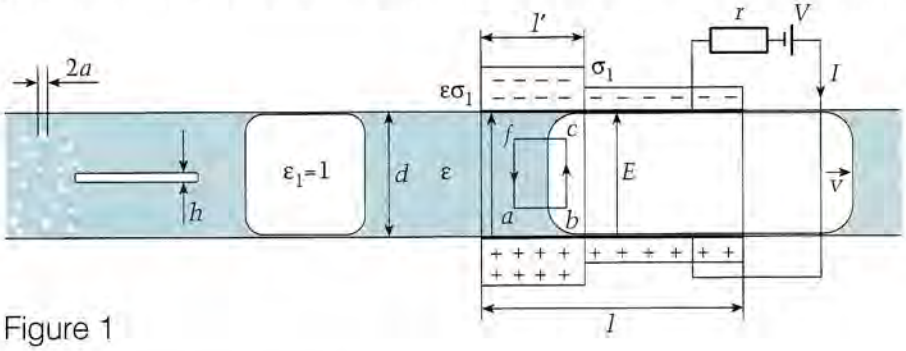



Figure 1

If the dielectric moves into the capacitor with a constant speed $v$ ($l' = vt$), a direct current will flow in the circuit:

$$I = \frac{dq}{dt} = \frac{\varepsilon_0 SV}{d}\frac{v}{l}(\varepsilon - 1) \quad \text{at } 0 < t < \frac{l}{v}. \tag{2}$$

When liquid fills the entire length of the capacitor, the charge reaches its largest value, $q_\varepsilon = \varepsilon q_1$, and remains at this value.

In contrast, when a bubble enters the capacitor, its charge will decrease at the same rate, and the electric current will flow in the opposite direction (figure 2). Therefore, even if our flat pipe is opaque, we can "see" the motion of the gas and liquid parts of the fluid mixture through changes in the electromagnetic field.

From the technological perspective, this type of flow—a gas–fluid mixture in which large bubbles fill the entire cross section of the pipe—is undesirable. For example, in the production of carbonated water, both phases are separated in such a flow, while the intent is that they be mixed. Let's look at a more useful type of flow.

This time, let the gas "bubble" take the shape of a flat slit of width

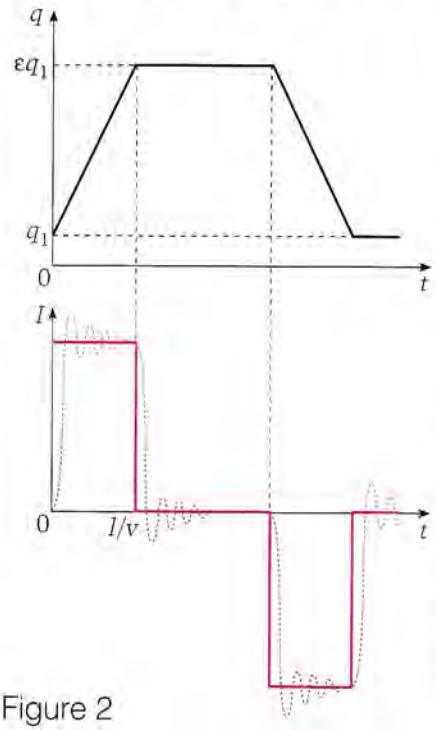



Figure 2

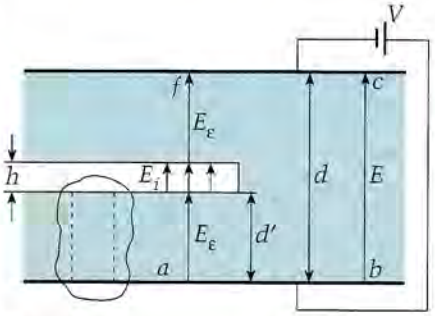



Figure 3

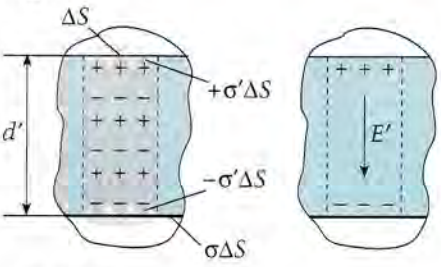



Figure 4

$h$, parallel to the plates of the capacitor (see figures 1 and 3). As before, we must perform zero work in moving the test charge along the path $abcfa$ (figure 3). In other words, the potential difference between points $a$ and $f$ is equal to that between points $b$ and $c$:

$$E_\varepsilon(d - h) + E_i h = V, \tag{3}$$

where $E_i$ is the field strength in the slit and $E_\varepsilon$ is the corresponding value in the dielectric (that is, in the liquid) on both sides of the slit. In addition,

$$E_i = \varepsilon E_\varepsilon. \tag{4}$$

By the way, this is what your physics textbook says: *the relative dielectric constant of a medium is a physical magnitude that shows by what factor the electric field strength ($E_\varepsilon$) inside a uniform dielectric substance is less than the field strength in a vacuum.* And this stipulation is tacked on: the definition is correct only in particular cases—say, for plates in a uniform field (and it's invalid for a hollow sphere). So let's think more carefully about the physical meaning of ε. Previously we considered it to be the factor by which the capacitance of a parallel-plate capacitor filled with a dielectric is greater than that of an "empty" capacitor.

Imagine that we cut a prism with a cross-sectional area $\Delta S$ out of our device (figure 4). The capacitor plate carries an electric charge $+\sigma\Delta S$, and in a vacuum the electric field above this plate is $E_i = \sigma/\varepsilon_0$ (below the plate—that is, outside the capacitor—it's zero). Our dielectric prism is located in an external electric field $E_i$, so it's polarized by that field. This property is illustrated by the dipoles, which are oriented vertically.

Inside the prism, the unlike charges of adjacent dipoles cancel each other out, while their noncompensated "tails," with charges of $\pm\sigma'\Delta S$, protrude outside the prism. Therefore, the dipole moment of the prism is $\Delta p = \sigma'\Delta S d'$ and is directed upward—that is, from the negative charge to the positive charge. The strength of the electric field generated by these polarized charges equals $E' = -\sigma'/\varepsilon_0$ and is directed counter to the dipole moment and the external field. Thus the strength of the total electric field is

$$E_\varepsilon = E' + E_i = -\frac{\sigma'}{\varepsilon_0} + \frac{\sigma}{\varepsilon_0}.$$

One step remains. Let's introduce the concept of the volume density of the dipole moment:

$$P = \frac{\Delta p}{\Delta S d'} = \sigma' = -\varepsilon_0 E'.$$

This physical value is related to the total field in the dielectric according to following equation:

$$P = \varepsilon_0(\varepsilon - 1)E_\varepsilon,$$

which may be considered a general definition of the local relative dielectric constant valid for any point in both a uniform (homogeneous) and nonuniform (heterogeneous) dielectric.

Equations (3) and (4) yield the electric charge on the plates when the gas slit is longer than the capacitor of length $l$ (that is, the bubble projects beyond its edges):

$$q = \varepsilon_0 \varepsilon S E_\varepsilon = \varepsilon_0 S \frac{V}{d}\langle\varepsilon\rangle.$$

Here we used the notation of the volume dielectric constant

$$\langle \varepsilon \rangle = \frac{\varepsilon}{1+(\varepsilon-1)h/d},$$

which takes into account the portion of the volume ($h/d$) occupied by the flat slit.

If we use equation (1) analogously to examine the process by which the "bubble" gradually moves into the capacitor with a constant speed $v$, we can use equation (2) analogously to find the current flowing in the circuit:

$$I = \frac{-\varepsilon_0 SV}{d}\frac{v}{l}\frac{\varepsilon(\varepsilon-1)h/d}{1+(\varepsilon-1)h/d}.$$

This formula is quite different from equation (2), although it coincides with it when $h/d \to 1$—that is, when the gas bubble is moving into the capacitor (see the falling branch of $I(t)$ in figure 2).

However, the time has come to say a few words about small bubbles (the left side of figure 1). Although these bubbles are small, their total relative volume can vary within broad limits—from zero (when the gas phase is absent) to one (when all the small bubbles are fused into a single large gas bubble). The difficulty in describing such a heterogeneous medium is aggravated even more by the fact that the bubbles may have different radii, and the distances between them may vary randomly. Moreover, they may collide and fuse into larger bubbles, or a large bubble may disintegrate into smaller bubbles. To make the picture even more confused, there is an electric field, which polarizes the bubbles and therefore converts them to interacting dipoles.

Speaking of which, do you know in what field a bubble-dipole is situated? Answer: in the total electric field generated by all sources—the free charges in the plates and the bound (polarized) charges. So what is the precise meaning of the phrase "a bubble is situated in the field"? Perhaps it refers to the field that would be located in the place occupied by the bubble if the bubble itself were removed from this place (a charge must not affect itself)—then in the "emptiness" left by the bubble there would remain a field generated by all the remaining electric charges. Many outstanding scientists have racked their brains over this problem: Irving Langmuir, Rudolf Clausius, Ottaviano Mossotti, Hendrik Lorentz ...

Now you see that our problem is not trivial. In such cases physicists often say: let's break the problem down into smaller parts. First, we'll consider a single spherical bubble in an infinite volume of liquid, where a uniform electric field $E_\varepsilon$ is generated sufficiently far ("at infinity") from the bubble. Then we'll assume that there are many such bubbles in the liquid ($N$ bubbles in one cubic meter), but all these bubbles are identical and located (on average) at the same distance from one another (this distance is about $1/\sqrt[3]{N}$). As a result, we'll obtain some effective dielectric constant, averaged over the entire volume, that characterizes this bubble–liquid mixture.

However, even this moderate plan of theoretical work cannot be realized easily. But then, we don't have to go the whole way: the two examples given above showed that the result depends on the total volume of the bubbles in the space between the plates. Therefore, the dependence of the electric current in the real circuit on time will differ from that in these examples.

Have we taken all the important factors into account? Not by a long shot. For example, the dielectric substance will be drawn into the capacitor due to the interaction between the charged plates and the induced charges in the substance. This means that, in the first case (shell-like large bubbles), a bubble in the capacitor will be compressed from the left and right by two "pistons" of fluid. Similar compression takes place in the gas–liquid mixture if the total volume of the bubbles varies within the space, so that the motion of the fluid is not uniform.

Also, in reality the resistance of the connecting wires and the internal resistance of the voltage source are not negligibly small. If their sum is $r$, the voltage difference across the plates of the capacitor will be

$$\frac{q}{C(t)} = V - rI(t),$$

which in contrast to the previous approximation is not a constant value. In a precise theory we would also have to take into account the inductance $L$ of the circuit and the corresponding self-induced emf $-LdI/dt$, so that Kirchhoff's law assumes the form of an awful differential equation for the electric charge:

$$L\frac{d^2q}{dt^2} + r\frac{dq}{dt} + \frac{q}{C(t)} = V,$$

which describes the damped oscillations. This equation is a tough nut to crack, because the capacitance $C$ varies with time (this variation is the cornerstone of our method of testing the gas-liquid mixture). But we would expect that the simple dependences of charge and current (the solid lines in figure 2) on time will show an oscillating pattern (the dotted curves in the same figure).

We might propose other methods of measurement. For example, we may charge the capacitor to some voltage and switch the battery off. Since the conductance of the dielectric liquid is negligible (and the conductance of the gas is even smaller), the charge on the capacitor plates will be constant. When a liquid with a different bubble content flows through such a capacitor, the difference in potential across the plates will change. Such devices are known as capacitor-type transducers and are widely used in science and technology.

We must admit that such measurements yield only the total relative volume of the gas phase, not the number of bubbles per unit volume. It would also be nice to know their average size. For this, we'd need to exploit other physical phenomena and use other instruments (for example, optical devices). So, before you open that bottle of soda, think about the number of bubbles in it and the laws of nature. Bon appétit! ◙

AT THE BLACKBOARD III

# Group velocity

by Helio Waldman

IT'S NICE TO SEE GROUPS OF PEOPLE WALKING together on a path, groups of birds flying together in the sky, groups of ducks swimming together on a lake. It seems so natural that they do so. And yet, we know that to do so, each member of the group is exercising some kind of control over its position and velocity: the velocity of the group is unlike the velocity that each of its members would have if travelling by itself. Are they following a leader, or is the group leading itself? What kind of rules are the individuals following, and how do these individual rules relate to their collective, group behavior? These are nice questions to ask, but they may be tough—or fun—to answer.

## Waves

The concept of group velocity also arises in the context of waves. In this domain, it relates to the speed with which perturbations in a wave move in space, and with respect to its basic structure. The basic wave is generally thought to be something like:

$$s(x, t) = A\cos(\omega t - kx), \tag{1}$$

where $s$ is the wavelike variable quantity (for instance, an electric or magnetic field, pressure, or deformation), $t$ is time, $x$ is the wave direction in space, $\omega = 2\pi/T$, $k = 2\pi/\lambda$, $T$ is the time period, and $\lambda$ is the wave period in the $x$-direction of space, also called the wavelength.

The basic waves have their own speed, called the phase velocity. Consider, for example, a crest of the wave described by equation (1), that is, a point where $s = A$. In equation (1), crests occur whenever the argument of the cosine function is a multiple of $2\pi$. The equation of motion for the crest is then

$$\omega t - kx = m2\pi.$$

By differentiating this equation, we get the phase velocity, which is the speed with which crests, or any other "constant phase" points of the wave, move:

$$v_p = \frac{dx}{dt} = \frac{\omega}{k}. \tag{2}$$

In practice, however, one would never meet with a basic wave in its "pure" form as described by equation (1). Let's see what happens when two waves (a "group" as elementary as possible) with the same amplitude, but slightly different frequencies and wavelengths, superimpose, forming a spatial-temporal pattern given by

$$s_{12}(x, t) = A[\cos(\omega_1 t - k_1 x) + \cos(\omega_2 t - k_2 x)]. \tag{3}$$

From the rules of trigonometry:

$$s_{12}(t) = 2A\cos\left(\frac{\omega_1 + \omega_2}{2}t - \frac{k_1 + k_2}{2}x\right)\cdot$$
$$\cdot\cos\left(\frac{\omega_1 - \omega_2}{2}t - \frac{k_1 - k_2}{2}x\right), \tag{4}$$

which is the *product* of two basic waves! We can see that the sum (superposition) of two basic waves is equivalent to the product (or "modulation") of two other basic waves. These two other basic waves may be identified by direct inspection of the sinusoidal factors in equation (4). The first factor, which may be called a *carrier* in the language of communications, has its temporal and spatial frequencies given by the averages of the temporal and spatial frequencies of the basic component waves. The second factor, which may be called the *modulation* of the carrier, has its frequencies (temporal and spatial) given by one half of the corresponding differences between the frequencies of the basic waves.

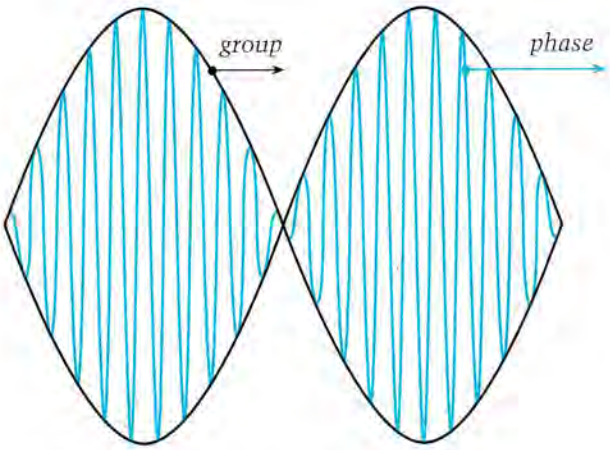



Figure 1. *A "group" of two waves.*

If $\omega$ and $k$ are the averages, and $\Delta\omega$ and $\Delta k$ the half-differences, we have

$$s_{12}(t) = 2A\cos(\omega t - kx)\cos(\Delta\omega \cdot t - \Delta k \cdot x). \quad (5)$$

Figure 1 shows a snapshot of this wave for some instant of time. One sees that the carrier is multiplied by another wave represented by a sinusoidal envelope.

Because the figure is stationary, it cannot show that the waves will, in general, have two different velocities. By applying equation (2) to each wave, we see that the basic wave crests move with velocity $\omega/k$ while the envelope wave crests move with velocity $\Delta\omega/\Delta k$. The former is called the *phase velocity* of the composite wave; the latter is called the *group velocity*, which in the limit (for very small $\Delta\omega$) is given by

$$v_g = \frac{d\omega}{dk}. \quad (6)$$

This example is still artificial, since the modulation itself is also an infinite wave. A more natural instance of the same phenomenon is given by an electromagnetic pulse with finite duration, formed by a finite number of wave crests. In this case, it is possible to show that the pulse may also be decomposed into basic waves given by equation (1), only there are not just two anymore, but rather a continuum of frequencies that occupies a spectral range about as wide as the inverse of the pulse duration. In spite of the increased complexity of the situation, the group velocity given by equation (6) will keep its validity insofar as the pulse is not distorted beyond recognition by the effect of the higher-order derivatives of $\omega$ with respect to $k$.

We started with the discussion of moving objects (birds, people, and so on), and went on to discuss waves. One might ask whether we are not mixing things up. Let us remind ourselves that waves are a succession of identical objects, which may be thought of as cycles, crests, or individuals with a standardized behavior. The next sections show that this individual behavior determines the velocity of any group of individuals that for some reason has detached itself from the regular pattern of a periodic queue. This "group" velocity is in general different from the velocity of the surrounding crowd (the "phase" velocity).

## Interference patterns (beats)

Let's see how the group velocity given by equation (6) emerges naturally from the behavior of the members of the "group." For this purpose, let's establish a relationship between the phase and group velocities given by equations (2) and (6), respectively:

$$v_g = \frac{d\omega}{dk} = \frac{d(kv_p)}{dk} = v_p + k\frac{dv_p}{dk}. \quad (7)$$

Since $\lambda k = 2\pi$ = constant, one has

$$v_g = v_p - \lambda\frac{dv_p}{d\lambda}. \quad (8)$$

In order to visualize how this relationship between phase and group velocities emerges from the behavior of the succeeding objects in each wave, the reader may perform a very simple experiment at home. It is enough to take two combs of similar, but not equal, spacings, and just superimpose them: observing them against the light, one can see an interference pattern (called "beats") between the two periodic patterns. If we slowly move one comb against the other, we can see that this interference pattern moves much faster than the moving comb, and not necessarily in the same direction. If the comb with the largest spacing moves to the right with the other comb fixed, the interference pattern will move to the left.

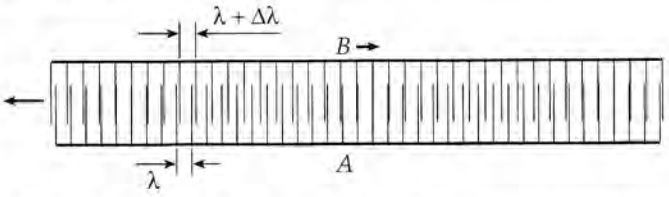



Figure 2. *When B is shifted $\Delta\lambda$ to the right, the pattern moves $\lambda$ to the left.*

Let's see what is happening. Let $\lambda$ be the spacing between the spokes of comb $A$, and $(\lambda + \Delta\lambda)$ of comb $B$ (figure 2). The motion of the beat pattern may be inferred by following the point of coincidence between spokes of both combs as one of them moves with respect to the other. If $\Delta\lambda > 0$, every time $B$ moves $\Delta\lambda$ to the right with respect to $A$, the point of coincidence moves to the $A$ spoke immediately to the left; that is, it moves $\lambda$ to the left. Thus, the pattern moves against the direction of $B$ with respect to $A$, but $\lambda/\Delta\lambda$ times faster. Now suppose $A$ moves with velocity $v_p$ and $B$ with velocity $v_p + \Delta v_p$ in the same direction in space. The pattern will then move with velocity $v_p - (\lambda/\Delta\lambda)\Delta v_p$, which approaches equation (8) when all increments tend to zero.

The examples above show that a pattern formed by the superposition of two periodical structures moves with respect to them with a speed that depends on the

derivative of the basic wave velocity with respect to its spatial period. Hence the distinction between phase and group velocities.

This distinction disappears, of course, if all basic periodic patterns move with the same exact velocity. This happens with electromagnetic waves in vacuum. Since they all move with the same velocity $c = 300{,}000$ km/s, any interference pattern between them will also move with the same velocity $c$. In matter, however, propagation velocities depend on the wavelength, because the "frequency response" of the atoms to alternating excitations is not flat. This results in a group velocity different from the phase velocity.

## Traffic

Let's now see how we can uncover this same concept from the individual behavior of objects in a moving queue. Consider, for example, a sequence of cars in a traffic-saturated road, with the vehicles traveling with speed $v$ and spacing $\lambda$ between them. Along the traffic flow there is a "congested" zone, where the spacing is reduced to $\lambda - \Delta\lambda$ and the speed to $v - \Delta v$ (figure 3). In this case, if we can show that the boundaries of the congested zone move with the same speed and direction, this zone is characterized as a "group," and its boundaries move with a group velocity $v_g$.

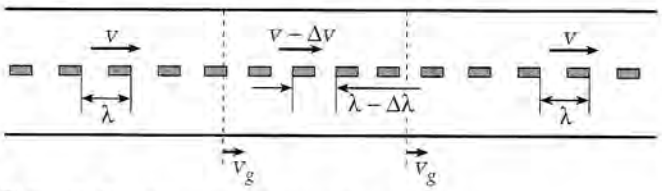



Figure 3. *"Congested" zone in a road.*

Consider the front of the congested zone moving ahead with speed $v_g < v - \Delta v$. A car in the congested zone approaches the front with speed $v - \Delta v - v_g$, which is also the speed with which the car sees the end of the congestion approaching it. When it arrives at the end of the congestion, it increases its speed to $v$, thus moving away from the car behind it with relative speed $\Delta v$. After time $\Delta\lambda/\Delta v$, the spacing between them will have increased from $\lambda - \Delta\lambda$ to $\lambda$, meaning that the congestion front reached the next car behind; that is, it moved $\lambda - \Delta\lambda$ backward as seen from any car in the congested zone. Therefore,

$$v - \Delta v - v_g = \frac{\lambda - \Delta\lambda}{\left(\frac{\Delta\lambda}{\Delta v}\right)}.$$

In the limit for small increments:

$$v_g = v - \lambda \frac{dv}{d\lambda}. \qquad (9)$$

This equation is a repetition of equation (8), with undisturbed traffic speed $v$ taking the role of phase velocity, and congestion speed representing group velocity $v_g$. Using similar arguments, the same conclusion may be reached about the tail of the congested zone.

In this example, one can see explicitly how the group velocity, now represented by the congestion velocity, depends on the behavior of the drivers, that is, on the spacing $\lambda$ between cars as a function of the speed $v$ of the traffic flow. The conventional behavior (recommended by the transit authorities, and followed by the typical driver) implies spacing proportional to speed, thus making $v$ proportional to $\lambda$. Setting $v = A\lambda$ for any constant $A$, one gets $v_g = 0$ in equation (9), meaning that the group velocity is zero in this case. Therefore, conventional driving behavior produces static traffic jams that move neither forward nor backward! In practice, one may observe that such congested zones last for hours, even after the generating cause has disappeared.

Some people wonder why traffic jams do not move forward with the traffic flow. In equation (9), one can see that for this to happen ($v_g = v$), the moving cars would have to maintain their speed independent of the spacing $\lambda$, that is, drivers would have to be extremely imprudent. This vision, however, suggests some intriguing possibilities. In a futuristic situation, one might consider cars being driven by networked automata. Their driving behavior might then be safely reprogrammed so that traffic jams move forward or backward, thus restoring traffic fluidity after some time! For example, if the spacing is proportional to the square of the speed (doubling when the speed increases only 41.4%), we can see from equation (9) that traffic jams would move along with the traffic at one half of the traffic's speed ($v_g = v/2$). On the other hand, a more aggressive behavior in which the spacing is proportional to the square root of the speed (thus doubling only when the speed is quadrupled), would make traffic jams migrate against the flow of traffic, with the same speed as the traffic ($v_g = -v$).

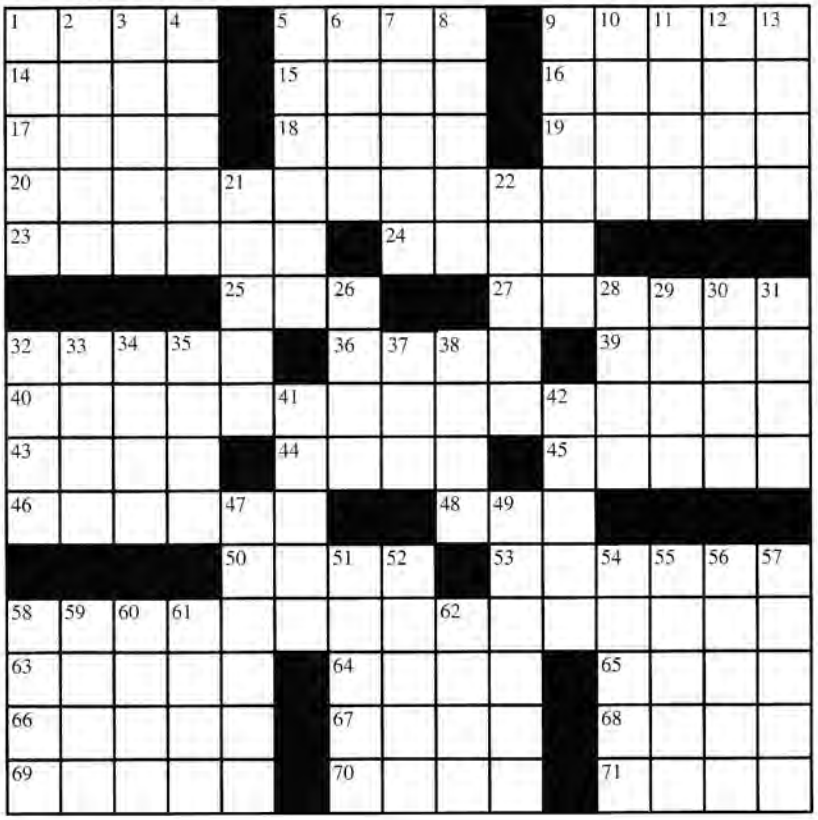
## Conclusions

Group velocity is usually seen as a concept that emerges in the analysis of the behavior of groups of waves. We have tried to argue that it is actually more widely applicable, as it characterizes the collective motion of groups of objects that follow a standard individual behavior given by a functional relation between velocity and spacing. The double interpretation reflects an ambiguity in the way we may decompose such sequences of moving objects. Decomposing them into individuals seems more natural in the analysis of "social" situations such as traffic. Decomposing them into superposed periodic structures (waves) is more convenient, for example, in the analysis of the propagation of electromagnetic pulses in linear media. In modern physics, this ambiguity reemerges in some exotic contexts, such as the wave-particle duality of the behavior of elementary particles. ◙

*Helio Waldman is a professor in the School of Electrical and Computer Engineering at the State University of Campinas, São Paulo, Brazil.*

# Criss × cross science

by David R. Martin



## Across

1 Soviet physiologist ___ Stern (1878–1968)
5 Number specialists: abbr.
9 Ship detector
14 Winglike
15 Pediatrician Luther ___ (1855–1924)
16 Psi follower
17 Fall month: abbr.
18 Botanist Katherine ___
19 Mended a shoe
20 Great similarities?
23 Breakfast dish
24 Alphabet run
25 Donkey
27 Brazilian river
32 Molecular biologist Werner ___
36 Decorative case
39 Miner's pick
40 Pluto or Mercury, e.g.
43 Gen. Robert ___
44 60 coulombs: abbr.
45 970,442 (in base 16)
46 Representative
48 Southern constellation
50 Space org.
53 Herbaceous plants
58 Auto repairman?
63 Below
64 Ireland
65 Million: pref.
66 Gland: comb. form
67 Swedish botanist ___ Afzelius (1750–1837)
68 Jannings or Ludwig
69 Writers Carl and Mark Van ___
70 Common flower
71 Miami's county

## Down

1 Cowboy's rope
2 Intestine part
3 Conic section part
4 Worker's cooperative (in Russia)
5 Storage boxes
6 Small bouquet
7 Texas landmark
8 Cut-off tree trunk
9 Pulpits
10 Certain asteroid
11 Sandwich shop
12 Author James ___
13 Units of radiation
21 Raises
22 Supply
26 Clothing joint
28 Shady mountainside
29 Almost: pref.
30 60,908 (in base 16)
31 Sporogonium stalk
32 Birds
33 Depend
34 British gun
35 Stared at
37 Thallium iodide
38 Arm bone
41 Portuguese territory
42 Dog's restraint
47 Nontranslated gene sequence
49 Type of inflorescence
51 Smudge
52 Indicating $NH_2$
54 Domesticated
55 Clyster
56 Stiff
57 Balance
58 College courtyard
59 Delete
60 ___ wax (ozocerite)
61 Hawaiian goose
62 Paleozoic and Mesozoic

*SOLUTION IN THE NEXT ISSUE*

*SOLUTION TO THE SEPTEMBER/OCTOBER PUZZLE*

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| G | A | F | F | ■ | C | L | A | I | M | ■ | S | L | U | E | D | ■ | M | I | C | A |
| A | L | L | O | ■ | H | E | N | R | I | ■ | W | E | N | D | Y | ■ | O | R | A | L |
| B | E | E | R | ■ | I | S | A | A | C | N | E | W | T | O | N | ■ | D | O | S | E |
| A | C | E | T | O | N | E | ■ | ■ | R | O | D | E | O | ■ | A | G | E | N | T | S |
| ■ | ■ | ■ | R | N | A | ■ | A | B | O | D | E | S | ■ | O | M | A | R | ■ | ■ | ■ |
| A | L | G | A | E | ■ | P | L | E | B | E | S | ■ | T | R | I | A | N | G | L | E |
| M | O | R | N | ■ | S | H | I | N | A | S | ■ | B | R | A | T | S | ■ | L | A | G |
| E | D | E | ■ | S | C | O | T | E | R | ■ | T | R | I | T | E | ■ | L | O | V | E |
| B | A | A | ■ | K | A | T | E | S | ■ | I | R | A | T | E | ■ | W | A | V | E | S |
| A | R | T | T | U | R | I | ■ | ■ | F | R | I | G | O | ■ | S | A | N | E | S | T |
| ■ | ■ | ■ | E | L | E | C | T | R | O | M | A | G | N | E | T | I | C | ■ | ■ | ■ |
| C | A | S | A | L | S | ■ | R | U | R | A | L | ■ | ■ | L | A | T | E | R | A | L |
| A | L | U | M | S | ■ | J | O | L | T | S | ■ | A | M | A | T | E | ■ | A | G | I |
| R | U | G | S | ■ | T | A | P | E | S | ■ | E | M | O | T | E | D | ■ | D | O | T |
| A | L | A | ■ | T | U | N | I | S | ■ | S | I | M | O | E | S | ■ | H | I | R | E |
| T | A | R | T | A | R | I | C | ■ | A | C | C | O | R | D | ■ | L | U | N | A | R |
| ■ | ■ | ■ | O | R | B | S | ■ | A | R | R | O | N | E | ■ | C | A | Y | ■ | ■ | ■ |
| H | A | L | L | O | O | ■ | A | R | I | E | S | ■ | ■ | T | A | N | G | E | N | T |
| O | P | A | L | ■ | J | A | M | E | S | W | A | T | S | O | N | ■ | E | D | E | R |
| L | I | M | E | ■ | E | B | E | C | E | ■ | N | A | T | T | A | ■ | N | I | N | A |
| T | S | P | N | ■ | T | O | R | A | N | ■ | E | X | P | E | L | ■ | S | T | E | M |

# ANSWERS, HINTS & SOLUTIONS

# Physics

## P306

Both the helicopter and the scale model are kept in the air by the reactive force arising when the rotors force the air downward. According to Newton's third law, the reactive force acting on the rotors and supporting the helicopter has the same magnitude as the force exerted on the air stream by the rotors.

Let's denote air density by $\rho$, the cross-sectional area of the airstream by $S$, and the air speed by $v$. During a short time $\Delta t$ the rotors push air with a volume $Sv\Delta t$ and mass $m = \rho Sv\Delta t$. Therefore, the momentum of the propelled air changes by

$$\Delta p = mv = \rho Sv^2\Delta t.$$

According to Newton's second law, the air experiences a force **F** equal to

$$F = \frac{\Delta p}{\Delta t} = \rho Sv^2.$$

A force of the same magnitude acts on the helicopter. To hold the helicopter in the air, this force must be equal to its weight:

$$\rho Sv^2 = Mg. \qquad (1)$$

The power $P$ of the engine is equal to the energy imparted to the airstream in one second:

$$P = \frac{mv^2}{2\Delta t} = \frac{\rho Sv^3}{2}.$$

Plugging in

$$v = \sqrt{\frac{Mg}{\rho S}}$$

from equation (1), we get

$$P = \frac{1}{2}Mg\sqrt{\frac{Mg}{\rho S}}. \qquad (2)$$

Since the mass of the helicopter is proportional to its volume (that is, to the third power of its linear size)—$M \sim L^3$, while $S \sim L^3$—equation (2) yields

$$P \sim L^{7/2}.$$

Thus the ratio of the power of the helicopter's motor to the power of the model's motor must be equal to the ratio of their linear sizes raised to the 7/2 power—that is,

$$\frac{P}{P_{\text{model}}} = \left(\frac{L}{L_{\text{model}}}\right)^{7/2},$$

from which we get

$$P = P_{\text{model}} \cdot 10^{7/2} \approx 95 \text{ kW}.$$

## P307

The number of molecules that escape from the balloon during time $t$ is

$$Z = \frac{1}{2}nS\langle v_x\rangle t,$$

where $\langle v_x\rangle$ is the mean value of the projection of the molecular velocity onto the $x$-axis (which is perpendicular to the wall where the hole is), $S$ is the area of the hole, and $n$ is the concentration of gas molecules in the balloon (the number of molecules per unit volume).

$\langle v_x\rangle$ is proportional to the speed $v$ of the thermal motion of the molecules. Since

$$v = \sqrt{\frac{3RT}{\mu}}$$

($\mu$ is the molar mass of the gas), $\langle v_x\rangle \sim \sqrt{T}$. The ideal gas equation $P = nkT$ yields $n = P/kT$.

Therefore,

$$Z = \frac{P}{T}\sqrt{T} = \frac{P}{\sqrt{T}}.$$

This equation says that a fourfold increase in temperature, combined with an eightfold increase in pressure, increases the leakage rate by a factor of four.

## P308

To ensure stable operation of the spark generator, the discharge in the spark gap must not affect the charging of the capacitor. This is possible when the time it takes for the capacitor to discharge across the spark gap is far less than the time required for a battery to charge it to the voltage $V$.

In this case, there is no current in the spark gap when the voltage across the capacitor is zero. Under these conditions, the spark gap restores its dielectric properties. The next discharge will occur when the voltage across the gap (and across the capacitor) reaches $V$.

While the capacitor is being charged to the voltage $V$, the battery performs work $W = q\mathcal{E}$, where $q = CV$ is the charge on the capacitor. According to the energy conservation law, $W = Q + CV^2/2$, where $Q$ is the energy dissipated by the resistor while the capacitor is being discharged. Therefore,

$$CV\mathcal{E} = Q + CV^2/2,$$

from which we get

$$Q = CV\mathcal{E}(1 - V/2\mathcal{E}).$$

Since the duration of the discharge across the spark gap is small, we neglect the energy dissipated by the resistor during this time.

If the capacitor is charged $n$ times per second, the mean power dissipated by the resistor is

$$P = nCV\mathcal{E}\left(1 - \frac{1}{2}\frac{V}{\mathcal{E}}\right).$$

## P309

The rate of decrease of the current in the circuit drops as the current decreases, because the induction emf is equal to the product of the current in the circuit and the resistance of the wire in the coil:

$$-L\frac{\Delta I}{\Delta t} = RI$$

or

$$-L\frac{\Delta I}{I} = R\Delta t.$$

This equation says that the current decreases by the same factor in equal time intervals. This is also true for the dissipated power. Therefore, in the next 100 ms, $0.01 \cdot 0.6^2$ J of heat will be dissipated by the coil. The total amount of heat is given by the following sum:

$$Q_{\text{total}} = 0.01\left(1 + 0.6 + 0.6^2 + 0.6^3 + \ldots\right) \text{J}$$

$$= \frac{0.01}{1-0.6}\ \text{J} = 0.025\ \text{J}.$$

## P310

We can't use the concept of diffraction to explain the phenomenon of "two shadows" because the size of the diffraction pattern at such a distance is very small.

The problem can be solved if we take note that the Sun is not a point source of light. It has a certain finite angular size. With this in mind, and referring to figure 1, we can explain the strange phenomenon of a "countershadow."

The angular size of the Sun is about a half-degree. Although this isn't much, it's enough to produce indistinct shadows on a sunny day. The degree of blurring of the shadow's edge depends on the distance from the object to the screen. Although the shadow's edge in our experiment appears sharp, that's not really the case. In addition to an area of "complete shadow," there is also a gray region called the penumbra. When your finger approaching the object cuts off some of the rays falling into the penumbra, a countershadow appears on the screen. Further motion of the finger downward

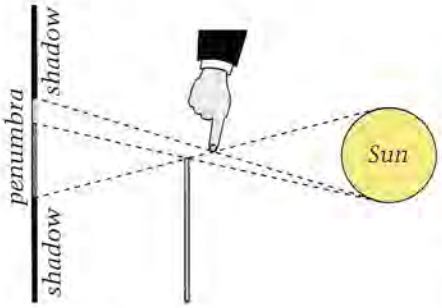



Figure 1

produces a fusion of the complete shadow with the penumbra at the boundary between the illuminated part of the screen and the penumbra.

By measuring the distances between object (cardboard) and the screen and between your finger and the object, as well as the size of the countershadow just before both shadows fuse, we can determine the angular size of the Sun using some very simple geometry.

# Math

## M306

Suppose $k > 0$ is the largest number on a seat for which a ticket has been sold. Let $K > k$ be the largest number with the following property: for any $i$ from $k$ through $K$, the number $f(i)$ of tickets sold for seats numbered from $k$ through $i$ is at least a great as the number $g(i) = i - k + 1$ of these seats.

Now it is clear that after everyone has been seated, then for any $ii = k, k + 1, \ldots, K$, the $i$th seat will be occupied, and $f(ii) - g(ii)$ "Oh!"s will have been uttered in going from the $i$th place to the $(i + 1)$st place. After all, this is exactly the number of seats lacking for the spectators with numbers $k$ through $i$.

If there are more spectators than those who occupied seats $k$ through $K$, we consider the least number $k' > K$, and repeat the same reasoning for the spectators occupying seats $k'$ through a $K'$, (where $K' > k'$), and so on. (Note that this description holds even if $n < m$.)

Thus, not only is the total number of "Oh!"s independent of the order in which the spectators arrive, so is the number of "Oh!"s uttered as spectators pass from place $i$ to place $(i + 1)$, for each value of $i$.

## M307

Denote the given diameters by $BC$ and $DE$ and the centers of the corresponding circles by $O_1$ and $O_2$. From $O_1$ and $O_2$ we draw perpendiculars to the corresponding diameters and call the point of their intersection $F$ (see figure 2). We'll prove that $F$ is the center of the desired circle.

Notice that $O_1F \parallel AO_2$, since $O_1F$ and $AO_2$ are perpendicular to $BC$. Similarly, $FO_2 \parallel AO_1$. Therefore, $AO_1FO_2$ is a parallelogram, and so $FO_2 = AO_1 = BO_1$ and $FO_1 = AO_2 = DO_2$. Hence the triangles $BO_1F$ and $FO_2D$ are congruent, and we have $FB = FD$. Moreover, from the construction, point $F$ lies on the perpendicular bisectors of segments $BC$ and $DE$; therefore, $FC = FD = FE$, as was to be proved.

## M308

The situation is *not* possible.

Suppose that the numbers 1, 2, 3, 4, 5, 7, and 8 are the roots of the equation $f(g(h(x))) = 0$. If the line $x = a$ is the axis of the parabola defined by the equation $y = h(x)$, then $h(x_1) = h(x_2)$ if and only if $x_1 + x_2 = 2a$. The polynomial $f(g(x))$ has no more than four roots. However, the numbers $h(1), h(2), \ldots, h(8)$ are a list of its roots, with some repeats. It follows that $a = 4.5$ and $h(4) = h(5)$, $h(3) = h(6)$, $h(2) = h(7)$, and $h(1) = h(8)$. In addition, we incidentally proved that the numbers $h(1)$, $h(2)$, $h(3)$, and $h(4)$ form a monotonic sequence.

Similarly, considering the trinomial $f(x)$ and its roots $g(h(1))$, $g(h(2))$,

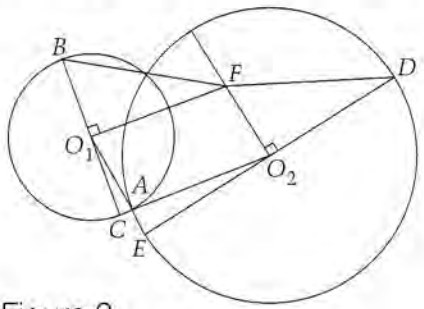



Figure 2

$g(h(3))$, and $g(h(4))$, we find that $h(1) + h(4) = 2b$ and $h(2) + h(3) = 2b$, where the line $x = b$ is the axis of the parabola defined by the equation $y = g(x)$. However, a little algebra shows that if $h(x) = Ax^2 + Bx + C$ and $h(1) + h(4) = h(2) + h(3)$, then $A = 0$. Thus we have arrived at a contradiction.

## M309

First, consider a line $l$ through $C$ that intersects segment $AB$—that is, those passing inside the angle $ACB$. Let $\angle ACB = 2\gamma$. The product $P$ of the distances from points $A$ and $B$ to line $l$ equals (see figure 3)

$$P = ab \sin\phi \sin\psi$$
$$= ab(\cos(\phi - \psi) - \cos(\phi + \psi))/2,$$

where $\phi$ and $\psi$ are the angles formed by $l$ with segments $AC$ and $BC$, respectively. In this expression, the quantities $a$, $b$, and $\phi + \psi = 2\gamma$ are constant, so $P$ is largest when $\cos(\phi - \psi)$ is largest, which is when $\phi = \psi$. This maximum is given by

$$P' = ab(1 - \cos 2\gamma)/2 = ab \sin^2\gamma.$$

If $l$ passes outside angle $ACB$, the formula for $P$ is the same (see figure 4). In this case, the sum $\phi + \psi = \pi - 2\gamma = 2\delta$ is also constant and equals the exterior angle at the vertex $C$ of triangle $ABC$. The maximum value of $P$ is attained for $\phi = \psi$ and equals

$$P_2 = ab \sin^2\delta.$$

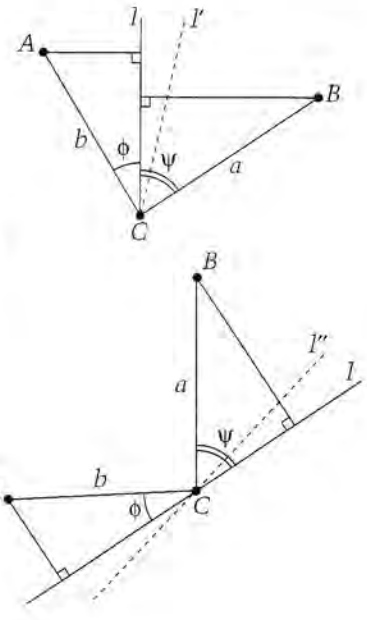


Figure 3

Thus if the angle $2\gamma = \angle ACB > \pi/2$, then the desired line is the bisector $l'$ of angle $ACB$ (in this case, $\sin^2\gamma > \sin^2\delta$ and $P_1 > P_2$). If $2\gamma < \pi/2$, then the desired line is the bisector $l''$ of the exterior angle at vertex $C$ of triangle $ABC$ ($P_1 < P_2$). If $2\gamma = \pi/2$, then $P_1 = P_2$ and there exist two lines with the same maximum product $P$: $l'$ and $l''$.

## M310

Each problem was not solved by three of the eight students. Suppose that no two students exist who (together) solved all the problems. This means that, for every pair of students $\{X, Y\}$, there exists a problem $P_{\{X, Y\}}$ that they didn't solve. There exists $8 \cdot 7/2 = 28$ pairs $\{X, Y\}$. However, every one of the eight problems can play the role of $P_{\{X, Y\}}$ only for three pairs $\{X, Y\}$, and $8 \cdot 3 = 24 < 28$. Thus we arrive at a contradiction.

The main reasoning we used here is making the transition to complementary sets and "negations." Similar reasoning makes it possible to show that if there are $p$ problems and $n$ students, every problem was solved by no less than $n - m$ students, and $n(n-1) > pm(m-1)$, then there exist two students who solved (together) all problems. If $n(n-1) \ldots (n-k+1) > pm(m-1) \ldots (m-k+1)$, then, for a certain $k > 1$, there exist $k$ students who solved (together) all the problems.

# Brainteasers

## B306

Let $\angle BDA = \angle CDB = \alpha$. Since $BC = CD$, we have $\angle CBD = \alpha$. Therefore, $BC \parallel AD$. Two cases are possible (figure 4).

(1) $AB$ is parallel to $CD$. Then $ABCD$ is a rhombus, $ABD$ is an equilateral triangle, and $\angle BAD = 60°$.

(2) $AB$ is not parallel to $CD$. Then $ABCD$ is a trapezoid. By the statement of the problem, it is isosceles and $\angle BAD = \angle ABD = 2\alpha$. Thus we have $\alpha + 2\alpha + 2\alpha = 180°$. Therefore, $\alpha = 36°$, and $\angle BAD = 72°$.

So the answer is 60° or 72°.

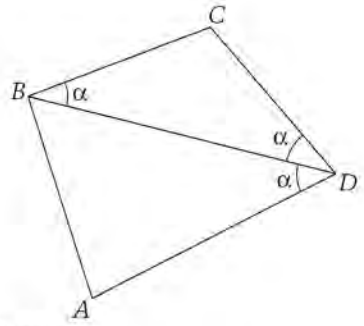


Figure 4

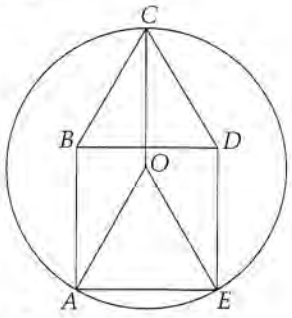


Figure 5

## B307

Let's construct a triangle $AOE$ congruent to $BCD$ as shown in figure 5. It follows from the condition of the problem that $ABCO$ and $EDCO$ are rhombuses. Indeed, in quadrilateral $ABDE$, $BD = AE$, and $AB = DE$. Thus it is a parallelogram, so $AB \parallel DE$ and $AE \parallel BD$. Now $CD$ and $OE$ make equal angles with the parallel lines $BD$ and $AE$, and so $OE \parallel CD$, so that $OEDC$ is a rhombus. It's clear that $O$ is the center of the desired circle, and its radius equals 1.

## B308

We can break the $10 \times 10$ square down into twenty-five $2 \times 2$ squares. It's clear that there can't be more

| | | | | |
|---|---|---|---|---|
| 2 1<br>3 1 | 2 1<br>3 1 | 2 1<br>3 1 | 2 1<br>3 1 | 2 1<br>3 1 |
| 4 1<br>5 1 | 4 1<br>5 1 | 4 1<br>5 1 | 4 1<br>5 1 | 4 1<br>5 1 |
| 6 1<br>7 1 | 6 1<br>7 1 | 6 1<br>7 1 | 6 1<br>7 1 | 6 1<br>7 1 |
| 8 1<br>9 1 | 8 1<br>9 1 | 8 1<br>9 1 | 8 1<br>9 1 | 8 1<br>9 1 |
| 10 1<br>11 1 | 10 1<br>11 1 | 10 1<br>11 1 | 10 1<br>11 1 | 10 1<br>11 1 |

Figure 6

than two highly paid officials in each square; thus the number of such officials cannot exceed 50. Figure 6 shows an arrangement and salaries of the officials such that 50 of them can consider themselves highly paid.

## B309

It took Nick 3 hours 12 minutes—that is, 16/5 hours—to reach Georgetown, and it took George 2 hours 40 minutes—that is, 8/3 hours—to reach Nicktown. Denoting the distance between the towns by $L$ miles, we find that Nick was walking at a speed of $5L/16$ mph and George's speed was $3L/8$ mph. We can determine the length of the bridge $l$, since we know that George crossed it one minute faster than Nick: $16l/5L - 8l/3L = 1/60$. This yields $l = L/32$. Let $t$ be the moment the boys reached the bridge. At this moment, the total distance walked by both boys was $L - L/32 = 31L/32$. On the other hand, this equals the sum of the distances walked by each of them—that is,

$$\frac{5L}{16}\left(t-\left(10+\frac{3}{10}\right)\right)+\frac{3L}{8}(t-9)$$
$$=\frac{L}{16}\left(11t-\frac{211}{2}\right).$$

Setting these expressions equal to each other, we obtain

$$\frac{L}{16}\left(11t-\frac{211}{2}\right)=\frac{31L}{32},$$

which gives us $t = 11$ o'clock.

## B310

The key to the problem is the inertia of the air in the car. When the car slowed down, the air continued to move forward. This increased the pressure at the front platform and decreased the pressure at the rear platform. Hence, the storyteller was standing on the rear platform as the decreased pressure would cause air to flow from the vent.

HAPPENINGS

# Cyberteaser winners

THIS MONTH'S CYBERTEASER was a bridge too far for some contestants, however, the majority of you were able to bridge the mental gap and correctly calculate the time that Nick and George arrived at the span in question. The following were the first ten correct solutions to the problems received at *Quantum* headquarters:

**Jerold Lewandowski** (Troy, New York)
**Nick Fonarev** (Staten Island, New York)
**Marco Devigili** (Verona, Italy)
**Theo Koupelis** (Wausau, Wisconsin)
**Anastasia Nikitina** (Pasadena, California)
**Michael H. Brill** (Morrisville, Pennsylvania)
**Bruno Konder** (Rio de Janeiro, Brazil)
**Jacopo De Simoi** (Treviso, Italy)
**John Beam** (Bellaire, Texas)
**Manny Dekermenjian** (Menlo Park, California)

Our congratulations to this month's winners, who will receive a copy of this issue of *Quantum* and the coveted *Quantum* button. Everyone who submitted a correct answer (up to the time the answer is posted on the web) is entered into a drawing for a copy of *Quantum Quandaries*, a stimulating collection of 100 *Quantum* brainteasers. Our thanks to everyone who submitted an answer—right or wrong. The new CyberTeaser can be found at http://www.nsta.org/quantum. ■




STATEMENT OF OWNERSHIP, MANAGEMENT, AND CIRCULATION (Required by 39 U.S.C. 3685). (1) Publication title: Quantum. (2) Publication No. 008-544. (3) Filing Date: 10/00. (4) Issue Frequency: Bi-Monthly. (5) No. of Issues Published Annually: 6. (6) Annual Subscription Price: $56.00. (7) Complete Mailing Address of Known Office of Publication: 175 Fifth Avenue, New York, NY 10010-7858. Contact Person: Joe Kozak, Telephone: 212-460-1500, ext. 303. (8) Complete Mailing Address of Headquarters or General Business Office of Publisher: 175 Fifth Avenue, New York, NY 10010-7858. (9) Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor: Publisher: Gerald F. Wheeler, National Science Teachers Association, 1840 Wilson Boulevard, Arlington, VA 22201-3000 in cooperation with Springer-Verlag New York Inc., 175 Fifth Avenue, New York, NY 10010-7858. Editors: Larry D. Kirkpatrick, Mark E. Saul, Albert L. Stasenko, Igor F. Sharygin, National Science Teachers Association, 1840 Wilson Boulevard, Arlington, VA 22201-3000. Managing Editor: Sergey Ivanov, National Science Teachers Association, 1840 Wilson Boulevard, Arlington, VA 22201-3000. (10) Owner: National Science Teachers Association, 1840 Wilson Boulevard, Arlington, VA 22201-3000. (11) Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities: None. (12)Tax Status. The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes: Has Not Changed During Preceding 12 Months. (13) Publication Title: Quantum. (14) Issue Date for Circulation Data Below: September/October 2000. (15) Extent and Nature of Circulation: (a.) Total No. Copies (Net Press Run): Average No. Copies Each Issue During Preceding 12 Months, 6675; No. Copies of Single Issue Published Nearest to Filing Date, 5500. (b.) Paid and/or Requested Circulation: (1) Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541: Average No. Copies Each Issue During Preceding 12 Months, 2352; No. Copies of Single Issue Published Nearest to Filing Date, 2117. (2) Paid In-County Subscriptions: Average No. Copies Each Issue During Preceding 12 Months, 0; No. Copies of Single Issue Published Nearest to Filing Date, 0. (3) Sales Through Dealers and Carriers, Street Vendors, and Counter Sales, and Other Non-USPS Paid Distribution: Average No. Copies Each Issue During Preceding 12 Months, 583; No. Copies of Single Issue Published Nearest to Filing Date, 33. (4) Other Classes Mailed Through the USPS: Average No. Copies Each Issue During Preceding 12 Months, 1210; No. Copies of Single Issue Published Nearest to Filing Date, 1346. (c.) Total Paid and/or Requested Circulation: Average No. Copies Each Issue During Preceding 12 Months, 4145; No. Copies of Single Issue Published Nearest to Filing Date, 3496. (d.) Free Distribution by Mail: (1) Outside-County as Stated on Form 3541: Average No. Copies Each Issue During Preceding 12 Months, 85; No. Copies of Single Issue Published Nearest to Filing Date, 77. (2) In-County as Stated on Form 3541: Average No. Copies Each Issue During Preceding 12 Months, 0; No. Copies of Single Issue Published Nearest to Filing Date, 0. (3) Other Classes Mailed Through the USPS: Average No. Copies Each Issue During Preceding 12 Months, 0; No. Copies of Single Issue Published Nearest to Filing Date, 0. (e.) Free Distribution Outside the Mail: Average No. Copies Each Issue During Preceding 12 Months, 76; No. Copies of Single Issue Published Nearest to Filing Date, 76. (f.) Total Free Distribution: Average No. Copies Each Issue During Preceding 12 Months, 161; No. Copies of Single Issue Published Nearest to Filing Date, 153. (g.) Total Distribution: Average No. Copies Each Issue During Preceding 12 Months, 4306; No. Copies of Single Issue Published Nearest to Filing Date, 3649. (h.) Copies Not Distributed: Average No. Copies Each Issue During Preceding 12 Months, 2369; No. Copies of Single Issue Published Nearest to Filing Date, 1851. (i.) (Sum of 15g. and h.): Average No. Copies Each Issue During Preceding 12 Months, 6675; No. Copies of Single Issue Published Nearest to Filing Date, 5500. (j.) Percent Paid and/or Requested Circulation: Average No. Copies Each Issue During Preceding 12 Months, 96.26%; No. Copies of Single Issue Published Nearest to Filing Date, 95.80%. (16) Publication of Statement of Ownership will be printed in the Nov/Dec 2000 issue of this publication. I certify that all information furnished on this form is true and complete. (17) Signature and Title of Editor, Publisher, Business Manager, or Owner:

Dennis Looney
Senior Vice President and Chief Financial Officer
Date: 10/04/00

# Perfect shuffle

by Don Piele

HOW MANY TIMES DO YOU NEED TO SHUFFLE a deck of cards to feel reasonably certain that the cards are in random order? If you shuffle more times, do you feel the randomness improves? If you are very good at shuffling and are able to get the maximum amount of intermingling with each shuffle, will the randomness improve? Let's investigate these questions assuming we perform a "perfect shuffle." You may be surprised at the outcome.

What is a perfect shuffle? Take a deck of 52 cards and divide it in the middle (called a cut) so you have two piles of 26 cards each. Call the top pile *A* and the bottom pile *B*. Now shuffle the cards together by taking one card from the bottom of pile *A*, followed by one card from the bottom of pile *B*. Repeat this process, alternating cards from pile *A* and pile *B* until all cards are interlaced into one pile. This is a perfect shuffle.

Actually, it is called a "perfect in-shuffle" because all the cards, including the top and bottom card, are moved inside the deck to new positions. Had we decided to take our first card from the bottom of pile *B*, then we would have performed a "perfect out-shuffle." In this case the top and bottom cards always stay on top and bottom.

It is easy to see that in a perfect in-shuffle all the cards move to new positions. If you number the cards consecutively from 1 to 52 (from bottom to top), then the bottom card from pile *B* (originally in the 1 position) moves to position 2, and all other cards in pile *B* move to even higher positions. For pile *A*, all cards move into lower positions. Thus, all cards change their position after one perfect in-shuffle. All of the perfect shuffles done here will be in-shuffles and I will leave as an exercise the examination of out-shuffles.

What happens if we continue to do more perfect shuffles?

Let's take a small deck of 4 cards {1, 2, 3, 4} and see what happens.

Cut the cards $A = \{3, 4\}$, $B = \{1, 2\}$. Do a perfect shuffle to get {3, 1, 4, 2}.

Cut the cards $A = \{4, 2\}$, $B = \{3, 1\}$. Do a perfect shuffle to get {4, 3, 2, 1}.

Cut the cards $A = \{2, 1\}$, $B = \{4, 3\}$. Do a perfect shuffle to get {2, 4, 1, 3}.

Cut the cards $A = \{1, 3\}$, $B = \{2, 4\}$. Do a perfect shuffle to get {1, 2, 3, 4}.

We get the same order we started with after 4 perfect shuffles.

The matrix PS stores the order of the cards through all perfect shuffles.

```
PS = {{1, 2, 3, 4}, {3, 1, 4, 2}, {4, 3, 2, 1},
{2, 4, 1, 3}, {1, 2, 3, 4}}
```

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \\ 4 & 3 & 2 & 1 \\ 2 & 4 & 1 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

If we represent each card by a rectangle and color it by its original position number, we can easily get a visual representation of the PS matrix. Notice that the top and bottom rows are identical and the middle row is the reverse of the first row. Also note how short the expression to create this visualization is in *Mathematica*. **Hue** assigns the color for each position in the PS matrix and **Rectangle** draws the corresponding rectangle. **AspectRatio** makes the rectangles square.

```
Show[Graphics[Table[{Hue[(PS[[j,i]])/4],
Rectangle[{i, -j}, {1 + i, 1 - j}, {i, 1, 4},
{j, 1, 5}]], AspectRatio → 5/4]
```

I wonder how many perfect shuffles it takes to get back to the original order for a deck of 52 cards? What about a deck of *n* cards? Before we can answer that ques-

tion we need to develop an algorithm to automate the perfect shuffle. Let's implement the algorithm in *Mathematica* with a small deck of six cards.

## The algorithm

First make a deck of six cards with **Range**.

```
cards=Range[6]
```

{1, 2, 3, 4, 5, 6}

Divide them in two with **Partition**.

```
Partition[cards, 3]
```

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

Now take one card from the bottom of each pile and interlace them in a perfect in-shuffle. This can be done in *Mathematica* as follows:

```
RotateRight[%]
```

$$\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$$

Note: To do an out-shuffle you would not apply **RotateRight** to interchange the rows.

```
Transpose[%]
```

$$\begin{pmatrix} 4 & 1 \\ 5 & 2 \\ 6 & 3 \end{pmatrix}$$

```
Flatten[%]
```

{4, 1, 5, 2, 6, 3}

And there you have the results of the first perfect shuffle. Now compose these commands to create a shuffle function that will be applied over and over again.

```
shuffle[cards_]:= Flatten[Transpose
[RotateRight[Partition[cards,Length[cards]/
2]]]]
```

The built-in *Mathematica* command **NestList** applies the shuffle functions to the deck of cards 3 times.

```
NestList[shuffle,cards,3]
```

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 5 & 2 & 6 & 3 \\ 2 & 4 & 6 & 1 & 3 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

For a deck of six cards, it returns to its original order in 3 perfect shuffles. There goes any conjecture that a deck of $n$ cards requires $n$ perfect shuffles to get back to its original order.

If we use the **NestWhileList** command, we can make our program even smarter. Now it knows when to stop iterating the shuffle function. It applies this function as long as the shuffles are all unequal.

```
PS = NestWhileList[shuffle, cards, Un-
equal, All]
```

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 5 & 2 & 6 & 3 \\ 2 & 4 & 6 & 1 & 3 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

Showing the graph of our PS matrix, we see the perfect shuffles in color.

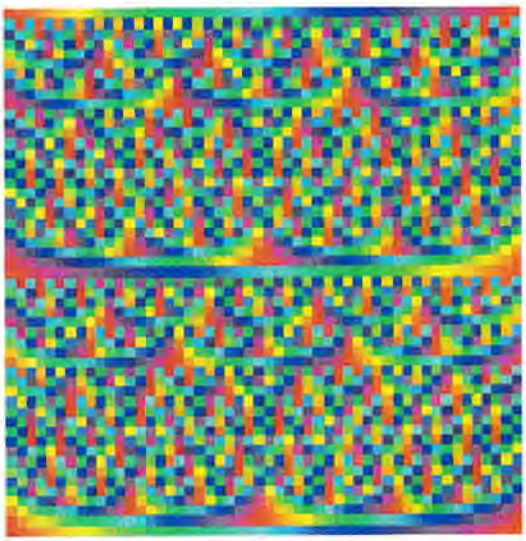
```
Show[Graphics[Table[{Hue[(PS[[j, i]])/
6], Rectangle[{i, -j}, {1 + i, 1 - j}]},
{i, 1, 6}, {j, 1, 4}]], AspectRatio → 4/6]
```

Let's see what happens with 52 cards.

```
cards = Range[52];

PS = NestWhileList[shuffle, cards, Unequal,
All];

Show[Graphics[Table[{Hue[(PS[[j, i]])/52],
Rectangle[{i, -j}, {1 + i, 1 - j}]}, {i, 1, 52},
{j, 1, 53}]], AspectRatio → 1]
```



The card order is reversed after 26 shuffles and returns to the original order in 52 shuffles.

## Other decks

We're not done yet. We need to investigate other deck sizes. First we define a **PerfectShuffle** function that will investigate a deck of size $n$ ($n$ is even) and return the minimum number of shuffles necessary to get back to its original order. We call this number the Perfect Shuffle Number. We compose the commands into a

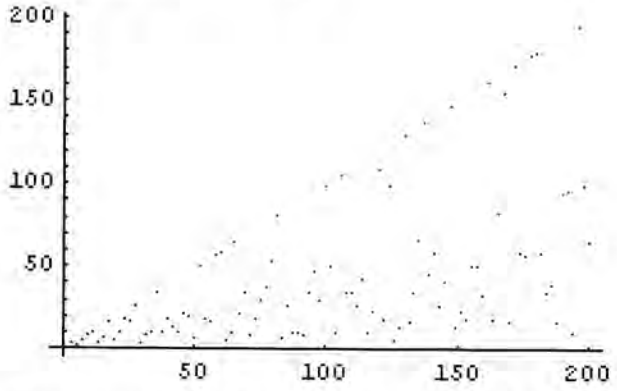
one-line **PerfectShuffle** function. The **?EvenQ** checks that the input is even. The **Length** of the perfect shuffle matrix is one more than the Perfect Shuffle Number.

```
PerfectShuffle[n_?EvenQ] := Length
[NestWhileList[shuffle, Range[n], Unequal,
All]] - 1
```

Now we apply this function to decks of size $n$, where $n$ goes from 4 to 200 in jumps of size 2, and graph the results:

```
PSNumbers = Table[{n, PerfectShuffle[n]},
{n, 4, 200, 2}];

ListPlot[PSNumbers]
```



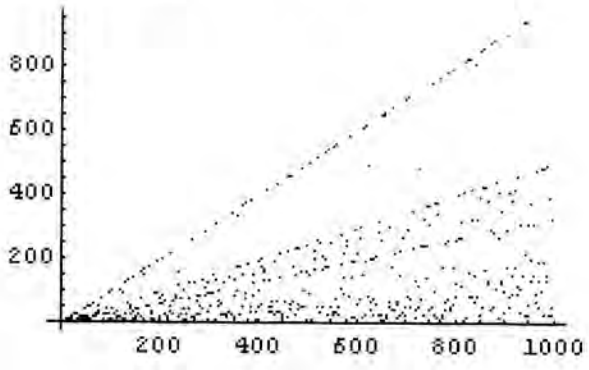

Clearly, no deck takes more perfect shuffles than its size to get back to its original state, and most take considerably less. But to investigate larger decks, we will need to find a faster algorithm. This one is slowed down by the fact that it keeps track of all shuffles. We need to keep only the most recent shuffle. This fact is reflected in the following, much faster algorithm.

```
FastPerfectShuffle[n_]:= Module[{x, y, i},
x = Range[n]; y = shuffle[x]; i = 1;
While[x ≠ y, y = shuffle[y]; i++]; i]
```

We can quickly examine the Perfect Shuffle Number for $n$ up to 1000.

```
PSNumbers = Table[{n,
FastPerfectShuffle[n]}, {n, 4, 1000, 2}];

ListPlot[PSNumbers]
```




## A mathematical solution

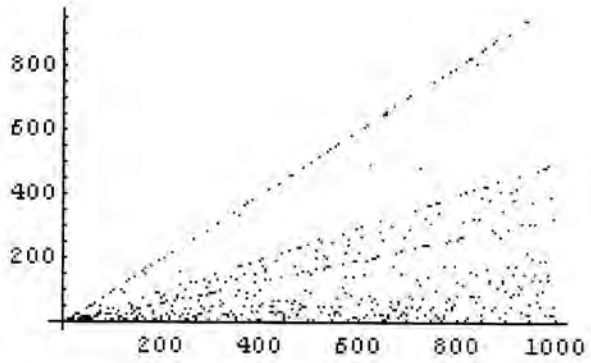
Is there a mathematical expression that computes Perfect Shuffle Numbers? Yes—according to Martin Gardner in his book *Mathematical Carnival*. The number of perfect in-shuffles needed for a deck of size $n$ ($n$ is even) is the smallest $x$ where $2^x = 1$ modulo $(n + 1)$. So let's create a **MathPerfectShuffle** function that uses this expression to compute the Perfect Shuffle Number.

```
MathPerfectShuffle[n_] := Module[{x}, x
= 2; While[Mod[2^x, n + 1] ≠ 1, x++]; x]

MathPSNumbers = Table[{n,
MathPerfectShuffle[n]}, {n, 4, 1000, 2}];
```

The resulting graph for $n$ up to 1000 is identical.

```
ListPlot[MathPSNumbers]
```




## Your turn

What happens if you use the out-shuffle instead of the in-shuffle for each shuffle? Go through a similar analysis. Can you come up with a mathematical expression that produces a mathematical solution?

## USACO

The USA Computing Olympiad (USACO) will begin its 2000–2001 season with the first Internet competition to be held November 8–15, 2000. The fall competition is a programming contest open to all pre-college students throughout the world. You may sign up by joining the mailing list hs-computing@delos.com.

Problems and rules will be posted to the hs-computing mailing list. Solutions must be written in C/C++ or Pascal and must be returned via e-mail by the contest completion deadline.

Students can work on problem solutions anywhere they wish during the one-week period. No entry fees are charged, and all winners will receive an official award and have their names immortalized on the USACO Web pages.

To find out more about the USACO and the results of our USA 2000 team at the 2000 International Olympiad in Informatics (IOI) in Beijing, China, go to our Web site at www.usaco.org and click on 2000 and then IOI. To get started using our training materials, go to ace.delos.com/usacogate. ◙